# ICT FOR AGRICULTURE, RURAL DEVELOPMENT AND ENVIRONMENT

## Where we are?
## Where we will go?

edited by Tomas Mildorf and Karel Charvat jr.

# ICT FOR AGRICULTURE, RURAL DEVELOPMENT AND ENVIRONMENT

**Where we are? Where we will go?**

# ICT FOR AGRICULTURE, RURAL DEVELOPMENT AND ENVIRONMENT

## Where we are?
## Where we will go?

edited by Tomas Mildorf and Karel Charvat jr.

**Czech Centre for Science and Society**
**Wirelessinfo**
*2012*

*Editors:*
**Tomas Mildorf**
České centrum pro vědu a společnost
*mildorf@centrum.cz*

**Karel Charvat jr.**
Wirelessinfo
*charvat_junior@wirelessinfo.cz*

*Reviewer:*
**Maris Alberts**
Institute of Mathematics and Computer Science, University of Latvia
*alberts@latnet.lv*

*Authors:*
Balogh S., Baptista F., Benda P., Bojko J., Brun F., Buitendag A.A.K., Cartolano jr. E.A., Castro Neto M., Cerba O., Cesaretti G.P., Champomier J.C., Charvat K., Charvat jr. K., Cheruy G., Claupein W., Conesa-Munoz J., Crehan P., Csoto M., Csukas B., Dreux B., Dvorak V., Fusai B., Gelb E., Gerhards R., Gnip P., Gravalos I., Havlicek Z., Holster H, Horakova S., Jezek J., Kahuda D., Kateris D., Keller M., Kepka M., Kozel J., Krivanek Z., Kubicek P., Link-Dolezal J., Lohr V., Lokers R., Loutridis S., Lukas V., Masselin-Silvin S., Maurer L., Misso R., Molinari F., Moshou D., Musil M., Natsis A., Navarro M., Navas L.M., Oberti J., Oskera M., Pesonen L., Prepin B., Prochazka F., Ribeiro A., Ruiz G., Samie J.F., Saraiva A.M, Sawalhi N., Seaz A., Schmitz M., Silerova E., Svehla M., Teye F., Vach M., Van Der Walt J.S., Varga M., Veiga A.K., Verdouw C., Vohnout P., Waksman G., Weis M., Wissocq A.A.K., Zech C.

**ISBN: 978-80-905151-0-9**
Photo cover: www.shutterstock.com
First published, 2012

# Table of contents

# Introduction

Karel Charvat

What is the aim of this book? Why we have decided to put together contribution from different authors? We don't want only to describe the status of current development and research in ICT for Agriculture, Food and Environment. We would like to analyse the period of last 10 years in terms of the results achieved in the research area, implementation in practice and success of the implementation. The book analyses the technological progress during this period and the progress in the application area. What are the current challenges having the influence on ICT for agriculture and environment? What is the current status of research and what are the future trends?

The book is based on two pillars:

- Results of current research in Czech Centre for Science and Society, WirelessInfo and our members;
- Articles of invited authors from around the world.

The book was accomplished thanks to the support of the AgriXchange, COIN IP, VlitNod, LernSens and Habitats projects with use of results of previous projects including Rural Wins, A Bard, Aforo, Ami@NEtFood. Other projects supporting single articles are mentioned within these articles.

The book is divided into five chapters:

- Where we are – overview of previous activities and description of the state of the art
- Challenges for future ICT for agriculture
- Where we go
- Vision of the future
- Conclusion

The book and materials should help to build and publicly discuss the Strategic Research Agenda (SRA) of agriXchange. The draft of agriXchange SRA is part of this book.

# Section I
# Where we are?

# Overview of previous activities

Karel Charvat
*Czech Centre for Science and Society*
E-mail: charvat@ccss.cz

**Introduction**

ICT for agriculture, environment and rural development is not a new issue. There were many activities and initiatives during last year. New research projects are very often starting from scratch without taking into account results from previous activities. Therefore, we decided to include this chapter to show some project results performed mainly during the first decade of the 21st century. We expect that this chapter can provide a good overview how ideas have changed, what was reached and where invisible progress is. This chapter is divided into three parts:

- Analysis of supporting and coordination actions and key declarations with focus on future strategies in the area of ICT for agriculture, rural development and environment.
- Analysis of surveys about ICT adoption provided at the EFITA and WCCA congresses.
- AgriXchange analysis of the state of the art of ICT for agriculture and interoperability in countries of EU 27 and Switzerland.

Analysis of supporting and coordination actions and key declarations with focus on future strategies in the area of ICT for agriculture, rural development and environment includes a short overview of the results or recommendations from the projects and studies and a list of previous declarations:

- Aforo Roadmap
- Rural Wins Recommendation
- Valencia Declaration
- eRural Brussels conference conclusions
- Prague Declaration
- aBard vision
- ami@netfood Strategic Research Agenda
- Study on Availability of Access to Computer Networks in Rural Areas
- Future Farms recommendations
- Cologne Declaration

This chapter is prepared as an overview and a review of existing outputs and publications from the above mentioned projects.

### Aforo

The objective of the AFORO [1] project (running in 2002 and 2003) was to provide a vision and work plan to implement future RTD trends for the transformation of agrifood industries into digital enterprises". The AFORO consortium split the agri-food domain into several more "manageable" sectors. Each of them had its own roadmap. The selected sectors were:

(a) primary sources,

(b) processed food products,

(c) beverages, and

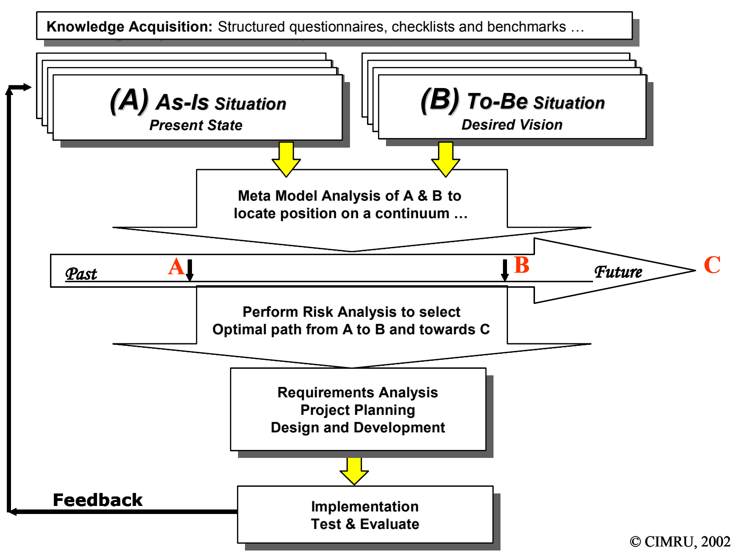
(d) additives, conservatives & flavours.

For every sector the main objectives were:

- to define business needs,
- to identify the main constraints to be taken into account,
- to define technology independent roadmap based on business demands.

The AFORO methodology defined an approach on how the business needs of the agrifood domain can be established in an ordered way. The process was divided in the following steps:

- data gathering,
- ascertaining of current and future objectives,
- analysis of these objectives,
- listing of key drivers [1].



The important output from AFORO roadmap is consensus of business needs. The following business needs were recognized by AFORO [2]

- To support food traceability and safety over the whole European market chain.
- To implement interoperability technologies enabling networked organizations and forming Single Food European Market.
- To develop market knowledge supporting innovative value added products, processes and business strategies.
- To design and develop interoperability tools for European logistic and services.
- To design European Agrifood Information System including materials, technologies, and results of RTD activities.
- To support transformation of agrifood business into an effective collaborative organization.

**Rural Wins**

The objective of Rural Wins [3] (2002 – 2003) was to build a strategic RTD roadmap developing an information and communications technologies' vision to ensure the economically and technically feasible deployment of information and communication solutions for rural areas including also maritime regions and islands.

The project provided analysis of:

- the trends in technology development,
- the needed equipment,
- the deployment of services.

Different scenarios of joint public and private initiatives and business models were analysed in order to reduce the discriminatory gap that nowadays exists between rural and urban areas with regard to broadband accessibility and applications' deployment.
RURAL WINS used the European qualitative system of DG REGIO to define and structure its analyses three types of rural areas:

Integrated rural areas - territories with growing population, employment basis in the secondary and tertiary sectors, but with farming being still important land use. The environmental, social and cultural heritage of some of these areas, relatively close to big cities, may be under pressure of "rurbanisation". The rural character of some of these areas is at risk of becoming predominantly dwelling suburban areas.

Intermediate rural areas - relatively distant from urban centres, with a mixture of primary and secondary sectors; in many countries larger scale farming operations can be found.

Remote rural areas - areas with the lowest population densities, the lowest incomes and older population. These areas depend heavily on agricultural or fishing employment and generally provide the least adequate basic services. Isolated features are often of topographic character, e.g. mountains, or remote places from transport and communications' networks.

For all three types of rural areas, the barriers that need to be addressed by Broadband ICTs are identified as:

- Distance barriers - access to administrative and governmental services and structures (taxes, subsidies etc.)
- Economic barriers - access to wider business and labour markets (suppliers, customers, opportunities)
- Social barriers of rural inhabitants to information, education & training facilities, health, social services etc.
- Information barriers – currently the amenities of many rural areas are "invisible" to the "outside world" (inhabitants of other areas, urban centres or citizens of other states - rural tourism, local products etc.)

Based on the analysis of socio-economic and technology trends, the project has concluded that for Rural Broadband access is needed:

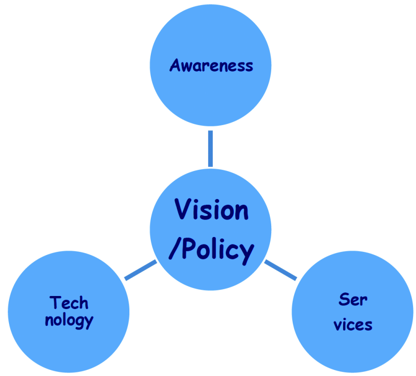
- Public/Private Partnerships
- New access technologies

The project developed a strategic roadmap template that provides a simple, coherent and complete framework to describe the ICT requirements and issues for rural and maritime areas. The framework consists of 4 dimensions:

1. Vision/Policy,
2. Awareness,
3. Technology/Infrastructure
4. Services/Applications

These were identified for each of the RURAL WINS' (a) Integrated, (b) Intermediate & (c) Remote types of rural and maritime areas.

Using this framework, the following concrete scenarios are identified for each type of rural area:




- Integrated areas
  - ICT needs – similar to urban.
  - Recommended – "standard" fibre/wired/mobile/WLAN
  - Objectives – full parity and use with urban areas
  - Implementation – commercial
- Intermediate areas
  - ICT needs – distributed "economies of scope"
  - Recommended – some fibre/wire/mobile/WLAN to towns, satellite/WLAN elsewhere
  - Objectives – competitive SMEs & access by all to all services
  - Implementation – public/private partnerships
- Remote areas
  - ICT needs – part of regional economic, social & cultural development
  - Recommended – satellite/WLAN - new access approaches are required
  - Objectives – ubiquitous fixed & mobile services to overcome barriers
  - Implementation – public funding, public/private & community partnerships

The use of bi-directional broadband satellite links and local/community owned and operated wireless LANS (particularly Meshs of Wi-Fi (802.11a/b) wireless cells) are identified as being particularly relevant eRural access technologies for remote rural areas.
Rural Wins promotes "eRural" policy. The eRural IP has to adopt an integrated and multidisciplinary approach across the whole value chain from technology to services and awareness that will cover all RTD&D to eliminate the Urban/Rural Digital Divide. [4]

**Valencia Declaration**

The Valenica Declaration was the output of the European Conference "Information Society as Key Enabler for Rural Development" organised in Valencia, on 3rd and 4th February 2003. The Valencia Declaration addressed the following items [5]:

- INFRASTRUCTURES AND SERVICES - telecommunications infrastructures must provide the same level of information transmission and of technology and knowledge transfer. Due to low population density and weak economic activity in rural areas, it is necessary to design an optimal convergence model in order to deploy broadband telecommunications networks. Public administration has to support the deployment of such an infrastructure, offering a strategic framework of cooperation and reinforcing the process of telecommunications liberalisation. Member States should address communications infrastructures and their contents in rural areas as a strategic priority. This will require implementation of European Authorities' suggestions regarding the development of information society national and regional plans, through the development of infrastructures (broadband) as well as services. With regards to services to citizens, it must be taken into account that one of the reasons for the depopulation of rural areas is that urban areas provide services to which rural population has no access. The use of ICT could be the key to overcome this situation. Already available but needing to be enhanced, is direct access to administrations (e-Government), particularly the remote delivery of traditional public services, such as health (e-health), social assistance, education and lifelong learning (elearning).

- TRADITIONAL SECTORS AND NEW BUSINESS OPPORTUNITIES – it is necessary to use ICT in order to gain major benefits through becoming a part of the information society. With regards to the traditional sectors in rural areas (agriculture, live stock, fishing, forestry and food industries) it is worth mentioning the following benefits:
    - Concerning food safety and quality: the use of more efficient quality management and assurance tools as well as training, communication and information applications and government follow-up models.
    - The increase in clearness, transparency and efficiency in agricultural markets as well as in inside sector relationships and in relations among different sectors.
    - The increase in competitiveness by means of a faster adaptation to new market trends.
    - The increase in consumer confidence through the use of tools which enable food traceability, in a way that they will be able to check food safety and quality and that throughout the production process, environmental and animal welfare rules have been respected.
    - The use of e-business tools to concentrate offer and to allow producers to increase their share in the end value of the product.
- KNOWLEDGE SOCIETY FOR ALL - To get rural citizens and companies into the information society they must participate in the general cultural change process, required by this new society model. To this end, the first step to be taken is the development of “digital literacy” initiatives, which will promote the day-to-day use of computers and the Internet. ICT usage levels in companies and homes vary between countries, regions and even counties, but are always significantly lower in rural compared to urban areas. However, it is worth bearing in mind that rural areas are very diverse one from another and show varying levels of socio-economic development. Another factor which must not be disregarded is language. The respect for European multilingual diversity demands actions which contemplate this reality. Overcoming the psychological resistance to technological change must not be underestimated, and will require awareness, training and capacity-building initiatives, while, at the same time, the communication interfaces improve their user-centered and user-friendly character. Moreover, ICT can provide rural inhabitants

with cultural and leisure services, that could help to overcome the "feeling of isolation" that is often considered a major cause of young people leaving rural areas. The creation of virtual communities and the recreational uses of the Internet can open a window to the world for rural citizens, while increasing the feeling of belonging to their rural community and identity.

The conclusion of the Valencia Declaration stressed as important to invite all public and private agents to create a new strategic framework of cooperation from which to promote:

- the development of telecommunications infrastructures on an equal basis in all European territories, within the confines of current regulatory legislation;
- the location of new activities in rural areas, as well as the development of value-creation actions for all businesses and services;
- the formation of active policies by those responsible for public administration at European, national, regional and local level in order to ensure appropriate action for the achievement of these objectives;
- multidisciplinary research (technical and socio-economic) to enable a better understanding of the fundamental new drivers of the information society and their impact on rural areas;
- the continuity of discussions, to share expertise and best practices, and to facilitate the progression from a model based on pilot projects to a model that allows successful initiatives to be widely adopted.

**eRural Brussels conference conclusions**

As a follow-up of the Valencia conference, the European associations EFITA together with the Czech Centre of Strategic Study (CCSS) and with support of DG INFSO of the European Commission organised one day @rural conference in Brussels. The mission of the @rural conference was to

- facilitate the support for establishing of new European @rural policy,
- exchange information and experience,
- the development of knowledge in the area of ICT in rural development in order to enhance the competitiveness of Europe,
- to promote the awareness of ICT in rural areas of Europe.

The Brussels @rural conference joined representatives of national, regional and local governments, ICT researchers and developers, ICT industry, consultants, specialists for rural development with representatives of the European Commission.

@rural conference conclusions stressed that IC technologies offer the possibility to bring new activities, services and applications to rural areas, or to enhance those already existing, providing thus a chance to overcome the barriers and to bridge the widening rural-urban divide. [6].

The development of the information society in rural areas fosters European development and integration, and increases competitiveness of European companies. A region needs a solid foundation for innovation, particularly through an innovation infrastructure with telecommunications and effective use of knowledge. Communications technology enables increased interconnectivity between knowledge workers and companies through virtual networking. Shortages of skills and qualified staff emerge as a major obstacle to innovation. Regions should therefore pay more attention to lifelong learning to facilitate the adoption of new technologies.

The discussed @rural vision supported the stimulation of services, applications and content including modern online public services, e-government, e-learning, e-health services and dynamic business environment. It is important to fund more services, applications and content. Broadband infrastructure provision depends on the availability of new services to use it. This problem is particularly acute in rural areas, where competitive infrastructure provision is not emerging as rapidly as in more central or urban areas. There is thus need for a more pro-active and holistic approach that harnesses latent demand in rural areas for service provision. Getting affordable broadband to areas currently regarded as commercially unviable was mentioned as a "challenge". Such an approach would go a long way to ensure "wider adoption, broader availability and an extension of IST applications and services in all economic and public services and in the society as a whole"

Conclusions mentioned the importance of tacit and specialised knowledge calls for greater mobility of knowledge workers in rural regions and investment in training and education. Traditional approaches to the production and use of knowledge have to be adapted to this system view of the innovation process in a knowledge society for rural regions. To this end, new relationships must be established between public support organisations, universities and enterprises. In addition to their traditional roles in education and research, universities should promote the transfer and diffusion of knowledge and technologies using ICTs, especially towards their local regional business environment.

As priority was discussed the concept of ambient intelligence that provides a vision of the information society where the emphasis is on greater user-friendliness, more efficient services support, user-empowerment, and support for human interactions.

To assert implementation of European Research Area principles into rural regions, it is important to support research focusing on ICT implementation in rural areas with a critical mass of rural actors, effort and public-private-partnerships across Europe. The eRural research needs to adopt an integrated and multidisciplinary approach across the whole value chain from technology to services and awareness that will cover all RTD&D to eliminate the Urban/Rural Digital Divide.

Among the application topics, main future challenges were mentioned: eGovernment, environment monitoring and emergency services, food safety, precision farming, controlling of subsidies, application for food industry, ICT solution for forestry and wood industry, eBusiness, teleworking, added value chain, e rural tourism, eHealth, eLearning with special needs for remote areas, the eContent solutions, entertainment, eCulture.

In communication area, the important research topics were recognised: mobile systems and mainly WLAN, satellite communication, hyperlan, terminals able to work in different networks.

WEB technologies became an important topic – knowledge management, semantic WEB, personalised WEB, multilingual WEB, distributed system and interoperability including grid technologies (mainly semantic and knowledge grid), data and network security, location

based services, GIS and earth observation, visualisation, image processing simulation, embedded systems.
It was recognised, that the current level of development of technologies enables their direct implementation. It was mentioned that support has to be focused not only on research itself but also on the implementation of the research results. The relation between RTD support and implementation support (structural funds) is necessary. It was recognised as important to coordinate activities of DG INFSO, DG Research, DG Agri and DG Regio on European level.

**Prague Declaration**

The Prague Declaration was discussed and published at the European Conference Information System in Agriculture and Forestry organised in Prague between 15th and 17th May 2006. The document in the context of a proposed eRural Policy for Europe notes that the current EU Rural Development Strategy for 2007-2013 represents a welcome opportunity for the creation of a Rural Knowledge Society.
The participants declared that the overall eRural Policy Strategy should be based on the combination of a bottom-up and top-down Information and Communications Technology (ICT) development - always rooted in the cultural heritage, needs and desires of the rural communities. This will involve ICT decision makers, developers, providers, end users and other stakeholders.

The imperative aspects are the following:

- Policy and Leadership
- Research and Innovation Related to the Development of a Rural Knowledge Society
- Technology Access, Applications and Content
- New Working Environment Business Models
- Social and Human Aspects
- Environmental Aspects

The Prague Declaration devotes the following objectives:

- Sustainable eRural policies by those responsible for public administration at European, national, regional and local level.
- The support and take up of affordable telecommunications infrastructures and services on a unified basis in all European and other relevant territories.
- The development of value-creating activities supporting sustainable development of new growth areas in rural communities.
- Technology platforms supporting multidisciplinary and collaborative research and development on social, economic and environmental aspects to enable fundamental new drivers of the Rural Knowledge Society. These include assessment and monitoring of their impact on rural areas.
- Further discussions to share expertise and implementation of best practices. [7], [8]

**aBard vision**

A-BARD (2005 – 2006) was a Coordination Action that was researching rural broadband provision and use. It was targeted to the needs of rural communities. A-BARD addressed questions close to eRural actors including [9]:

- Rural broadband deployment in rural Europe: issues, models, best practices, affordability and accessibility?
- Broadband applications and services: what is emerging and can some of those applications and services directly address the digital divide?
- A-BARD is documenting emerging results and experience in order to focus and leverage emerging results from on-going RTD, mobile applications and services deployment and ICT take-up.

The aBard recommendations [8] were based primarily on the assessment of work that was undertaken in the A-BARD project, and a synthesis of technological possibilities, market opportunities and the capacity of actors in the regional authorities (to develop and expand ICT based services and applications). This synthesis takes into account the experience of rural areas as a whole and of possible development scenarios in the context of wider EU developments. It provided a perspective on market segmentation in terms of both existing and foreseeable telecommunications products / services. The A-BARD recommendations were structured into four categories:

- Policy aspects
- Strategic actions
- Standalone initiatives
- Further research & innovation

These needed to balance top-down and bottom-up approaches. They are summarised as follows:

1. Define an ambitious European Rural Broadband Strategy as an integral part of Sustainable Rural Development Policy
    - Allocate public funding where there is "market failure"
    - In i2010 and FP7, include specific infrastructure, ICT use and RTD initiatives for rural areas
2. Stimulate business and technical competition in the Rural Broadband Market
    - Every user should have a choice of 2 or more broadband access options
    - Stimulate Public Sector Demand aggregation in rural and remote areas
3. Develop sustainable Connected Rural eCommunities to stimulate demand and broadband take up
    - Enhance Regional Leadership and Local Champions
    - Promote and support awareness ("know what") and training ("know how")
4. Provide services and content that rural users want ("killer applications").
    - Local content
    - Entertainment
    - As well as eBusiness, eLearning, eHealth, eGovernment

**ami@netfood Strategic Research Agenda**

The objective of the AMI@Netfood project (2005 -2006) was to support the implementation of the IST Research Priority and Framework Programme, providing a long-term vision on future trends on cientific and technology research oriented to the development and

application of ambient intelligence technologies for the agri-food domain. [10]
The Strategic Research Agenda (SRA) outlined activities necessary to support both rural development and in particular agri-food industries. Concerning agri-food businesses, SRA intended to provide a path to facilitate the sector to retain their position as world leaders in providing safe and healthy food products at reasonable cost. The approach taken was to draw upon ICT to support businesses and industry in the agri-food sector and transform it into a Collaborative Working Environment (CWE). In relation with rural development domain, SRA described the needs of the sector and proposes measures in order to implement ICT solutions in rural areas to support their development. The approach selected is not only focused on the development of applications and infrastructures, but also as a means to promote the diversification of rural activities and the promotion of new services through the wide adoption of information and communications technologies [11].

The ami@netfood SRA defined the following challenges:

- Support the European agri-food industry, especially SMEs, to be a worldwide leader in the supply of high quality and safe food products.
- Increase the level of involvement of consumers in the agrifood value chain by means of the wide adoption of relevant IC technologies and applications.
- Increase the areas in which European citizens find collaborative working environments assisted by ICTs by extending them to agri-food industry and rural domain.
- Open new business opportunities for the European ICT industry through development of new applications and tools to support the European agri-food and rural sector.
- Contribute to trigger the investment in ICT and telecommunications infrastructure by means of creating new business models in rural areas.
- Making rural Europe a more attractive place to live, invest and work, promoting knowledge and innovation for growth and creating more and better jobs.

The research and technology development (RTD) domains selected were:[12]

- ICT applications for the complete traceability of products and services throughout a networked value chain.
- Collaborative environments in agri-food and rural areas.
- ICT applications supporting the management of natural resources and rural development creating value for citizens and businesses.
- Innovative ICT applications in rural areas using broadband infrastructure.

The defined RTD objectives were:

- Developing of interoperable integrated intra- and inter-enterprise applications.
- Improving network collaboration.

- Increase the effectiveness and efficiency of knowledge sharing.
- Improve the customer orientated business model.
- Supporting the dynamic network management.

**Study on Availability of Access to Computer Networks in Rural Areas**

The DG AGRI "Study on Availability of Access to Computer Networks in Rural Areas" provided policy makers, stakeholders and others with guidance how to maximise the benefits of Information & Communications Technology (ICT) for growth and jobs in all rural areas of Europe, using the support of rural development programmes. The study prepared:

- a database of best practices,
- a review of existing policies and literature Future Farms recommendation [13].

On the basis of the analysis a set of recommendations with focus on maximisation of the benefits of ICT for rural development was prepared:

- A coherent eRural strategy as an integral part of sustainable rural development policy, focusing on building capacity, even though this often produces "softer" outputs;
- Improvement within the eRural strategy of control and monitoring of ICT indicators, policies and initiatives including the collection of coherent statistical data;
- Measures which stimulate business and technical competition at different levels of scope and sophistication within the rural broadband market;
- Developing sustainable connected rural eCommunities to stimulate demand and ICT take-up – particularly by enhancing regional leadership and local champions to ensure that 'bottom up' projects flourish, and by supporting awareness ("know what") and training ("know how");
- Providing services and content that rural users feel are pertinent to them, especially entertainment and local content, as well as policy priorities such as eBusiness, eLearning, eHealth and eGovernment services;
- Encouraging initiatives which promote the theme of eCommunity, particularly by way of a common eRural agenda;
- Adopting a rubric of best practice at the interface between LEADER and those seeking access to funding;
- Extending investment in broadband infrastructure to all local public sector agencies and schools;
- Investing and developing the content of local networks;
- Raising the digital e-skills of local businesses and citizens;
- Introduction of an eProcurement process with appropriate safeguards and

innovative proactive online support to fast-track ICT projects in rural areas;

- Explicitly encourage the role of local authorities in laying ducts and then renting them to operators on an open and non-discriminatory basis, and promoting indoor pre-cabling for all new buildings in their regions.

Review adds new factors to those identified in previous research:

- A shared sense of lagging behind, which can be stimulated constructively by a local 'champion';
- Being spurred on as a consequence of a successful local enterprise;
- Being encouraged by the experience or ICT familiarity of others;
- Following a targeted intervention which demonstrably has improved the local quality of life;
- Emotional responses for local, personal reasons;
- Local resistance to an imposed agenda; [13]

**Future Farm vision**

Future Farm was a European project funded by the EU as part of the Seventh Research Framework Programme. It ran between 2008 and 2010. The main aim of the project was to make a vision of future arable farming. [14]

The project recognised that the future farming and also future farming knowledge management system will have to solve many problems, where there will be very different requirements on production, but also on strategic decisions. There exists a list of influences and drivers that will also influence farming. For example requirements on food quality and safety in opposition with requirements coming from growing population and on renewable energy production technologies. The project also highlights that in some cases production on renewable production energy can have negative influence on the environment. It is important to introduce a new knowledge system that will help to solve such problems.

By 2030 the importance of biotechnology will grow and research results could be transferred on farm level. There will be two important trends: requirements on quality of production (high quality food, vegetables and fruits and also growing demand on special food) and on environmental friendly production on one side and increasing demand on amount of production on the other side.

The project expects that two main groups of farm will exist in 2030:

- Multifunctional farms.
- Industrial farms with focus on high efficiency and high quality of production.

The focus of the multifunctional farm will be on efficient agriculture from an environmental and socioeconomic point of view. But the future of the multifunctional farms will depend on public dialog and valuation of non production goods.

The focus of the industrial farm will be to produce enough food and energy in a sustainable way which meets the consumer's demands. The quality of food will be important in Europe and it is expected, that also bio production will be industrialised or will become knowledge intensive. The expectation is that the industrial farm will be able to exist without subsidies, but it will dependent on the level of restriction. Science will be the key driver.

In 2030 agriculture will become fully knowledge driven. This will require full adoption of ICT. New sensors and nanotechnologies will become part of management. ICT facilitated the development of robotics and automation now used in many industries including the agri-food sector.

On architecture level of information systems, we recommend to be focused on service oriented architecture, which could guarantee better connection and interoperability of future systems. It will influence used GIS systems, better acceptation of XML standards, but also importance of robotics will grow.

For adoption of new technologies, it will be important to focus on two horizontal issues: education and standardisation. Without educated staff it will not be possible to introduce new knowledge intensive methods. Standardisation importance will grow for interconnectivity of different levels of farming knowledge systems. [15], [16].

**Cologne declaration**

The Cologne Declaration was prepared on European conference GeoFARMatics 2010 in Cologne in Germany, which was organised by the EU-funded projects FutureFarm, AgriXchange and CAPIGI. The Cologne Declaration recognises that agri-food. New changes in the global food supply, growing demands on food quality and quantity, energy demands and environmental aspects introduce new demands on future knowledge management. The objective of future knowledge management has to be focusedon the agri-food and rural communities to be able to react adequately to these changes. For this reason, knowledge management has to become part of the agri-food and rural production system. This requires a clear vision, which has to be based on discussions with farmers, experts, politicians and other stakeholders. The single digital market, Future Internet, sharing of knowledge, social networks, protection of data and open access to information will be essential for farming and rural communities.

The imperative aspects are the following:

- Policy and Leadership
    - including representatives of ICT for agriculture and rural development specialists into the legislative processes of the European Communities leading to the definition of priorities of the Digital Agenda;
    - raising awareness and establishing a social platform for exchange of

information among key participants, especially rural communities;

    - supporting a platform for standardisation of information inside rural communities and the agri-food sector but also between the agri-food sector and other sectors;
    - including rural knowledge management as an essential part of the future European Strategic Development plan;
    - building a coherent strategy for rural public sector information management.

- Research and Innovation - The development of knowledge-based systems for the farming sector has to be supported by ICT focused on:
    - Future Internet and Internet of things including sensor technology and machine to machine communication;
    - Service Oriented Architecture as a key element of architectures for future knowledge management systems;
    - The power of social networks and social media or so called knowledge internet;
    - Management and accessibility of geospatial information as a key information source for any decision;

  On the application side the focus has to be on:

    - ICT applications for the complete traceability of production, products and services throughout a networked value chain including logistics;
    - Collaborative environments and trusted sharing of knowledge and supporting innovations in agri-food and rural areas, especially supporting food quality and security;
    - ICT applications supporting the management of natural resources and rural development.

- Future Technological Solutions - Future Internet architectures must reflect the needs and specificity of rural communities. It has to be resilient, trustworthy and energy-efficient and designed to support open access and increasing heterogeneity of end-points. Networks should sustain a large number of devices, many orders of magnitude higher than the current Internet, handle large irregular information flows and be compatible with ultra high capacity, end-to-end connectivity. Service Oriented Architectures have to provide methods for systems development and integration where systems group functionality around business processes and package these as interoperable services. The future development of technology has to be based on a broader utilisation of social networks. It is important to support the development of machine-readable legislation, guidelines and standards to integrate management information systems with policy tools.

- Standardisation - Integration and orchestration among services based on semantic integration of collaborative activities, including semantic compatibility for information and services, as well as ontologies for collaboration will be a major priority for future solutions.
- Social organisation of knowledge management - The concept of Trust Centres has to represent an integrated approach to guarantee the security aspects for all participants in the future farm. There will be a growing importance of protection of privacy and IPR because trust of information is one from the priorities for all rural communities. Pan European Social Networks have to support trust centres and enable such technologies as cloud applications and which will have to guarantee knowledge security.
- Social and Human Aspects - Rural businesses are usually small or medium size businesses according to the number of people they employ and so knowledge management and internal processes are different from large companies. Future knowledge systems have to be based on each community's own concepts of value, cultural heritage and a local vision of a preferred future. The objective is to develop human-centred reference models of sustainable rural life-styles that overcome social divisions and exclusion and include unique rural features and create new rural businesses and social infrastructures and attractive computer-based education.
- Environmental Aspects - Societal and political pressures for increased environmental standards are expected. It will be necessary to discuss such aspects as the utilisation of GM production and environmental aspects of bio fuel production. These aspects will, on one side, require exact economical, health and environmental mathematical analysis and, on the other side, public discussion and e participation on such decisions.

The Cologne Declaration entreats all public and private agents to create a new strategic framework of cooperation to promote:

- Defined key objectives for the agri-food and rural Digital Agenda for 2020;
- Support for the development of future internet technologies and the internet of things;
- Support for social networking to reach consensus about future environmental, economical and social priorities of agri-food and rural strategies;
- The initiation of a continuous e-conference of all stakeholders to define a clear future strategy and priorities;
- Support for new ICT and knowledge-based solutions supporting the development of future generations of rural businesses;
- Support the development of knowledge technologies to guarantee, in the future, high quality food production;
- Open public discussion on how to solve future problems and to secure food

production versus the production of energy and environmental benefits. [17].

## Conclusion

If we compare the changes in recommendations during last ten years, we can see the biggest progress in deployment of communication infrastructure. The progress in deployment of infrastructure in Europe was very fast and changes from one year to other were visible. Currently we can see similar trends also in developing countries and mainly development of mobile network is now very fast.

The take up of new solutions into practice and also research in agri-food and rural applications is slower. Some priorities including support for networking of organisations and support for traceability can be sees as a major priority for the whole period.

We can consider that there is necessity to support:

- To better implement RTD results.
- Long time suitability of research and support for long time vision for RTD development in the agrifood sector.

## References


Dragan Stokic, Ana Teresa Correia, Martin John, Brian Wall, Fernando Ubieta, Vania Bicocchi, Massimo Garuti, Alvaro Oliveira, Juan Antonio Alvarez, Itziar Cuenca,Julio Font, Maria Acilu, Dimitris Kiritsis Dragan Stokic, AnaTeresa Correia, John Martin Agri-Food Roadmaps, IST-2001-37258, Roadmap based on business demands for each sector

Fernando Ubieta Aforo Information Society and Rural Development Evaluation Scenario, http://www.ruralwins.org/images/FernandoUbieta.pdf http://www.ruralwins.org/objectives.htm

John J O'Flaherty at all: Roadmap for ICT solutions for rural areas and maritime regions IST-2001-39107 D5.3:The Roadmap for Rural ICTs http://www.ruralwins.org/RW_D5.3.pdf

The Valencia Declaration, http://europeandcis.undp.org/uploads/valencia%20declaration%20rural%20info%20society.pdf

@rural conference conclusions, http://www.ami-communities.eu/www.amiatwork.com/europa.eu.int/information_society/activities/atwork/erural_at_work/documents/at_rural_conf_concls.pdf

THE PRAGUE DECLARATION, http://www.ccss.cz/en/?menuID=50&articleID=105&action=article&presenter=ArticleDetail

Karel Charvat at all, A-BARD – Analysing Broadband Access for Rural Development, Coordination Action, Priority SSP-2003-8.1.B.3.5 - Information Society Issues, D7.3 Final A-BARD Recommendations ftp://ftp.cordis.europa.eu/pub/ist/docs/directorate_d/cnt/a-bard_en.pdf http://www.ami-netfood.com/

Dr. Bjarki A Brynjarsson, Dr. Dragan Stokic, Harald Sundmaeker, Jesús de Juan. DRAFT Strategic Research Agenda. Collaborative Work Environments – ICTs supporting Agri-food businesses and Rural development

Strategic Research Agenda – Summary Collaborative Work Environments - ICT supporting agrifood businesses and rural development

Study on Availability of Access to Computer Networks in Rural Areas, Contract No: 30-CE-0099278/00-78, http://ec.europa.eu/agriculture/analysis/external/networks/fulltext_en.pdf http://www.futurefarm.eu/about

Karel Charvát, Pavel Gnip, Walter Mayer FutureFarm vision, AGRIS on-line Papers in Economics and Informatics
Charvat at all D1.2.3 Visions and recommendations for knowledge management, Report from FUTUREFARM project
Cologne Declaration, http://www.wirelessinfo.cz/?menuID=1&articleID=5&action=article&presenter=ArticleDetail

# The EFITA[1] ICT Adoption Questionnaire – Priority Indicators for the Future

E. Gelb
*Samuel Neaman Institute for National Policy Research Hebrew University, Center for Agricultural Economic Research*
E-mail: gelb@agri.huji.ac.il

## Introduction

The EFITA ICT Adoption Questionnaire program was initiated in 1999. The questionnaire itself was designed to provide a unique overtime set of reply convenience sets. These sets in turn were expected to provide an indicative tool to identify, relate and compare Adoption of ICT in Agriculture trends - in the EFITA member countries***. This Questionnaire initiative, by now reiterated in seven consecutive EFITA conferences since 1999, succeeded in providing a generalized review of ICT Adoption constraint trends over time. Questionnaire summaries enable comparison of indicated trend results with relevant non EFITA originating references as well. Those include the USDA Computer-use survey results since 1997 and Questionnaire replies from other ICT dedicated conferences.

The Questionnaire replies have identified and quantified a critical need to improve ICT Adoption proficiency. This over and above the prevalent need to minimize competence disparities in ICT utilization. Questionnaire replies stressed practical measures to alleviate this situation, indicated a need for enhanced agricultural extension and justified public funding of ICT services for agriculture. Additional competencies identified for future prioritizing included social networks, multi directional feedback between farmers, extension and research and enhanced integration within agricultural production chains. These were considered as relevant for non EFITA countries as well.

Continuation of this sequence of surveys can provide an ongoing, practical indication of ICT Adoption constraint trends and priorities to alleviate them. Such a baseline tool was considered as useful for ICT policy and development decision makers, services for the agricultural and rural sectors, farmers and their associated extension, research and services.

## ICT adoption

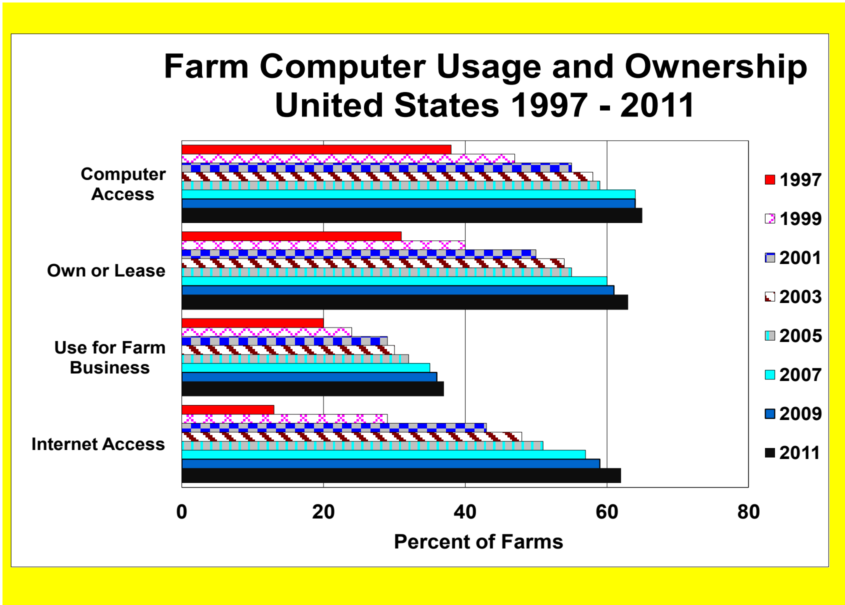
Effective adoption of ICT by farmers in Agricultural production and for the production chains till the end consumer remains a challenge since the initial use of Linear Programming for Agriculture in the late 1950s. The 1986 DLG sponsored congress in Hannover "Microelectronics (ICT at the time) in Agriculture" directly addressed this issue Gelb, 1986. A major focus related to the concern: "… how can farmers utilize ICT in Agricultural Production and what is the role of extension in attaining this goal?. By 1986 it was convincingly realized that although it might be feasible to adopt ICT and be cost effective ICT adoption is not necessarily straightforward, can initially even be counter-productive and inferior to established traditions and alternatives. Agricultural extension can, and is expected to go a long way to remedy such negative eventualities. To better understand the complexities

[1] European Federation for Information Technology in Agriculture

of the issues involved INEA (the Italian National Institute of Agricultural Economics - www.inea.it) and the Toscana Extension service sponsored in 1999 an international workshop in Alberese, Italy. The specific focus was on the adoption of Internet by Extension and farmers. It was a comprehensive attempt to identify and evaluate the critical success factors of this appreciated potential innovation. There was general agreement on critical success factors such as:
Internet accessibility on individual farms or at farmer-gathering locations; inputting of information that farmers/extension want and/or need; identifying a tangible benefit to information users; defining and serving target audiences; packaging information in a way that it can be understood and applied; a simple, user-friendly search engine and interface design; responsibility for the information quality and reliability; revealed preference as a guiding factor for Internet development.

Going deeper into the summary of detailed participant's opinions provided a practical, baseline, reference for Internet adoption planning, program implementation and goal achievement evaluation detailed in Gelb, Bonati 1999.

The Alberese workshop was soon followed by an intensive multi-national effort to further identify and quantify more generalized ICT adoption constraints – over and beyond the explicit Internet expressed insights. The effort was initiated and coordinated in the EFITA conference in Bonn in 1999. It was reiterated in the following EFITA Conference in Montpellier, France in 2001, in Debrecen, Hungary in 2003 and thereafter bi-annually during the EFITA conferences till and including the 8$^{th}$ EFITA conference in 2011 in Wageningn, Holland. The convenience-set tool to quantify this evaluation was initially designed in Alberese and refined for the Bonn Conference in the form of a Questionnaire: "*The EFITA ICT Adoption Questionnaire*". Each conference participant in turn was requested to elaborate their ICT Adoption experience as a specific assessment of constraints and their source. The responses were collated bi-annualy and reviewed over time. Eventually the accumulated sets enabled indicative comparisons from 1998 till the recent EFITA Conference in 2011 - Gelb, et.al. 2004, Gelb, Voet 2009. Comparisons between the initial Bonn, Montpellier and subsequent Debrecen replies could be associated with "short term" ICT Adoption constraint references. The Bonn baseline replies related to the 2011 Wageningen replies, more than a decade later can indicate longer term trends. This longer scope can expand the inclusion not only of additional ICT adoption constraints but responses to ICT innovations beyond initial familiarization difficulties. These include ever-changing quality integration of ICT infrastructure systems and technological innovations, direct and indirect ICT end-user benefit awareness, over time end-user proficiency, incorporation of advanced human engineering, enlightened farmer sophistication, and many more. The questionnaires do not provide a rigorous random sample analysis. However by maintaining their format over time the replies can be viewed as complementary convenience sets with the replies uniformly shared by similarly involved "ICT in agriculture" professionals. These provide a rough indication of generalized trends and their relative revealed importance. Furthermore the consistency of the replies enables relating them to perceived trends elsewhere. These for example could be indicators derived from results from the USDA computer-use surveys USDA, 2011, questionnaires reiterated with "ICT interested" farmers in Germany Rosskoff 2003, specific symposium presentations Taragula, 2005 and replies from respondents from non-EFITA member countries and related conferences. A long list of references and evaluation of ICT Adoption trends, Questionnaire replies and references can be accessed at Gelb, Offer 2006,. Gelb 2006, 2007 and Gelb et.al. 2008. The references are not limited to EFITA countries and reflect current increased growing international interest and concerns as well e.g. Pehu et.al. 2011.




*Farm Computer useage and ownership Agriculture in the USA, Source:USDA, 2011*

A useful example of such a trend comparison exercise and the ability to define issues relevant for EFITA countries would be to review the Alberese Internet adoption evaluation Gelb, Bonati 1999 as compared to the Questionnaires and USDA survey results. We can ponder - for example: "why is the Internet penetration in the USA since initiation in 1997 till 2011 less than 70%; is this quantified adoption pattern a relevant prognosis for Internet Adopting EFITA countries; a problem of current and future public interest, of agricultural consequence and an issue of national concern…."? A more detailed description follows.

**Evaluation of ICT Adoption questionnaire revealed trends**

While being useful in an attempt to understand and explain past ICT Adoption issues the questionnaire revealed trends might possibly suggest constraints to be expected. Being forewarned, and more important, enabling identification of specific problems in advance can carry significant benefits. The usefulness of predicting an eventuality (e.g. a decade long Internet Adoption delay since its introduction - relevant to >30% of the farmers) cannot be overemphasized. Relevant target groups within their concerns would include policy and decision makers facing regional and national ICT investment and priority choices, all those involved in services to agriculture including the production chains and last but not least the farmers with extension and research serving them. Combining these target group efficiencies and the significance of food production and rural viability as current National priorities

suggests an urgency to view the ICT Adoption trends and constraints from a public interest point of view.
The first two issues that come to mind are:

a. Is ICT adoption after all these years and accumulated experience still an issue of unique importance?

b. Is public funding of ICT services for agriculture justified?

Gelb and Parker 2007 evaluated in detail the replies to this first question namely: "Is ICT Adoption for Agriculture Still an Important Issue"? Table 1 adds an additional 5 years to the Gelb, Parker observations. It summarizes the replies collated over the whole 1999-2011 period.

| | **Montpellier 2001** | **Villa Real 2005** | **Glasgow 2007** | **Wageningen 2009** | **EFITA 2011** |
|---|---|---|---|---|---|
| **Are there problems with ICT uptake in Agriculture?** | **72.0** | **96.7** | **94.4** | **90.3** | **90.0** |
| **Is public funding for ICT services for Agriculture justified?** | **67.8** | **88.1** | **88.2** | **77.4** | **>90.0** |

*Table: Is ICT Adoption still an issue and is public funding for Agricultural ICT Services justifies? (% of "Yes" replies)*

The trends indicated in both questions are clear – even though quantitatively they are imprecise. Since 2001, over six consecutive EFITA conferences, ICT Adoption is recognized by a constant majority of conference participants as an ongoing and recognized problem. Gelb, Parker and Gelb, Voet confirm and detail this observation. The questionnaire continues to ask in Table 2 re this acceptably "recognized problem": *what are the main factors limiting the use of ICT by farmers over time?*

| | **Bonn 1999** | **Villa Real 2005** | **Glasgow 2007** | **Wageningen 2009** | **EFITA 2011** |
|---|---|---|---|---|---|
| **Inability to use ICT** | **29.3** | **45.0** | **12.5** | **45.2** | **<45.0** |
| **Infrastructure issues** | **18.9** | **35.0** | **28.6** | **25.8** | **0.0** |
| **Cost** | **17.6** | **25.0** | **24.2** | **29.0** | **<25.0** |
| **Lack of training** | **8.6** | **16.7** | **17.9** | **58.0** | **>65.0** |

*Table: What are the factors limiting the use of ICT by Farmers (% of "Yes" replies)*

The replies quite reasonably offer the following trend interpretation: Cost and infrastructure, since 1999 were not observed to be an increasingly dominant constraint. Infrastructure issues appear to have been resolved to a large degree - to the extent of being a non issue or as being ultimately solved in the broader context of regional and national communication services. Costs seem to be contained at a relatively low level – *but more important not as a significantly increasing issue.* As a trend for the future it can perhaps be assumed that the cost of the future

ICT innovations will be offset by lower production costs and greater efficiency of equipment, systems, communications and human capital involved (e.g. automation).

A consistent trend of a constant lack of substantial improvement in farmer ability to use ICT is clearly pronounced in the Questionnaire replies. Complementarily they point out that lack of training is a significantly growing and influential constraint. Within the confines of accuracy this comment could actually be phrased such that "the need for training has not diminished over time". Blame can be allotted to the increased complexities of ICT since 1999. A counter claim however would point to enhanced real life ICT competence visible in agricultural production. This competence would be due to achieved ICT utilization experience, updated education of the "new" agricultural graduates, "industrialized farmers" and innovative human engineering. All these should have more than compensated for the Questionnaire "lack of training" replies. In reality - they do not. In this case interpretation of this trend result by the mentioned target groups justifies, and even dictates, serious consideration of what to do, costs and "what if" impacts.

At this point it seems reasonable to test the validity of "Trend seeking and their evaluation". Specifically if ICT Adoption trends are not unique case study of technological innovation adoption much can be learned from comparison to "others". Namely is the ICT (and creation of the "Knowledge Based" society) adoption pattern comparable to the history of other such major category innovations – e.g. electricity. Standage, 1998 focuses this question on the history of Internet. Following this evaluation by looking back at the Alberese Workshop summary, Rogers' 1962 Innovation pattern as applied to the Internet Adoption results of the USDA survey combined with and the Questionnaire replies Gelb, Voet 2009 suggests that there are ICT Adoption similarities with other Innovations. And yes there is if not much to learn an outline for evaluating future trend priorities.

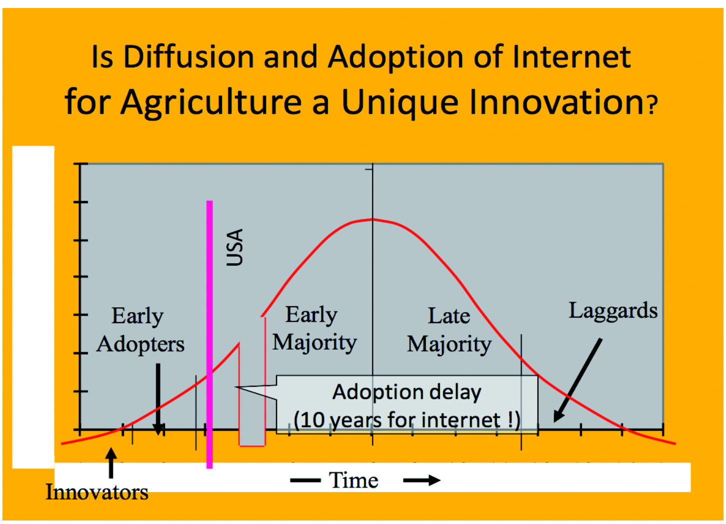



*Figure 1. Is Diffusion and Adoption of Internet for Agriculture a Unique Innovation? Source: USDA and Rogers 1962f*

In terms of future constraint trends such a formulation confirms the trend indicated by Questionnaire responses - namely that: " ICT adoption after all these years and accumulated experience is still an issue". Specifically in Table 1. it is recognized in the replies over time.
In terms of practicality consideration of priorities a significant dilemma emerges for the target groups mentioned above. For example from Policy and Decision maker's point of view. By accepting indication of trends for the future as a basis for priorities allocation it remains to be decided: should priorities and efforts focus on rescuing the "Late Majority" by bringing it up to par?

(Provide elementary extension and production subsidies for underperforming farmers). Alternatively should straightforward subsidies subsidize and support the "Early Majority" and/or encourage the "Early Adopters and Innovators". They in their achievements would demonstrate as act as "locomotives" to haul specific innovations into the mainstream agricultural practices?

The Questionnaire identified trends do not attempt to assume specific probabilities, suggest a comparison of estimated results or case by case recommendation. Suffice it to note that the trends, as a reference tool, could be identified, considered and shared quantifiably with relevant stakeholders.

**Discussion**

ICT and ever present ICT innovations are now undisputedly a dominant feature of numerous known and yet to be discovered components of our Agricultural production routine and Knowledge Society. Undertaking major decisions to adopt and promote these innovations in Agricultural production and Rural development - based on sketchy projections of future problems is risky at best. The benefit of early identification of adoption constraint trends cannot be overemphasized. Currently Agricultural Production and Rural Viability in all countries are of critical importance. The missed opportunity costs caused by non optimal ICT development and inferior adoption priorities can be prohibitive. These sub optimal priorities can result in inefficient allocation of production inputs, misdirected development funds, loss of comparative advantages, competitiveness, squandered rural futures and in worst cases social disruption. The modest contribution of the EFITA ICT Adoption Questionnaire to avoiding such outcomes is useful.

EFITA, among other vehicles is at the forefront of following ICT innovations and their adoption in agriculture, food production and rural development. A unique ICT adoption feature is the exponential increase of "Sources of Innovation" characterized and enabled by the almost universal adoption of ICT by stakeholders. Multi-directional interactions between stakeholders is currently intensifying and accelerating these processes. The EFITA questionnaires, in lieu of their consistency and focus, over time can assist in identifying past ICT Adoption constraint trends in the widest Agricultural production and Rural viability contexts. More important they can roughly indicate future constraint tenacity. This alone justifies the effort involved in continuation of reiterating the questionnaires beyond maintaining a valuable planning reference baseline. Beneficiaries of these indications range wide and far: Policy and decision makers, Extension, Research and service providers, additional sectors and the public at large – nationally and across national borders. As baselines they can also assist in monitoring and better understanding of current ICT oriented initiatives such as agriXchange, "ICT-AGRI ERA-NET" **(ict-agri.eu)** , FutureFarm, Social Networks and more.

At this point it seems to be too early to uniformly superimpose the identified adoption trends on e.g. activities to increase the impact of social networks on agricultural productivity, the future of an Open Farm Map, structured centralized information sharing, facilities for tracking and tracing agricultural products, real time production models, innovative automation and other ubiquitous ICT innovations. The generalities of the trends identified however do indicate the need to focus on elaborate training to minimize the time lag in adoption of an ICT innovation which in turn also justifies public funding of involvement in this effort.

**References:**


Gelb, E., 1986. Microelectronics in Agricultural Extension. Proceedings - DLG Congress Microelectronics in Agriculture Facts and Trends. Hanover, Germany.

Gelb, E., Deman, B., 1992. Users in Search of Farm Computer - Technology Applications. Proceedings, 4th International Congress on Computer Technology in Agriculture, Versailles.

Gelb, E. 1999. Internet for Extension in Agriculture - Review of Alberese Workshop. EFITA Papers.

Gelb, E., Getz, D., Obermann, G., 2003. Technology Spillovers in Rural Areas Proceedings, 3d EFITA Conference, Debrecen, Hungary.

Gelb, E., Scheifer, G., Wagner, P., Parker, C., Roskoff, K., 2004. ICT Adoption - A Summary of the EFITA (European Federation for Information Technology in Agriculture, Food and Environment) Questionnaires. Proceedings - AFITA/WCCA, Bangkok, Thailand.

Gelb, E., Offer, A., 2006. Information and Communication Technology in Agriculture: Perspectives of Technological Innovation. Editors for EFITA, Samuel Neaman Institute for Policy Research and Center for Agricultural Economic Research
http://departments.agri.huji.ac.il/economics/gelb-table.html

Gelb, E., 2006. Priorities for Agricultural Research, Development and ICT Adoption Remains a Challenge. Proceedings, 5th AFITA Conference , Bangalore, India.

Gelb, E., Parker, C., 2007. Is ICT Adoption for Agriculture still an Important issue? http://departments.agri.huji.ac.il/economics/gelb-gelb-parker.pdf

Gelb, E., 2007. ICT as an Enabler for Rural Economic and Social Viability: Bridging the Urban-Rural Digital Divide. Proceedings 18th ESEE Seminar, Prague.

Gelb, E., Public Funding of Agricultural Research, 2007. In Agricultural Research, Lowenstein (Ed.)

Gelb, E., Maru, A., Brodgen, J., Dodsworth, E., Samii, R., Pesce, V., 2008. Adoption of ICT Enabled Information Systems for Agricultural Development and Rural Viability, 6th AFITA Conference Summary. Atsugi, Japan.

Gelb, E., Aurenhammer, H., Wagner, P., Blackmore, S., 2009. Technological Innovation Adoption - Constraint Commonalities, Summary of the 2009 EFITA/JIAC plenary session.

Gelb, E., Voet, H., 2009 ICT Adoption Trends in Agriculture: A summary of the EFITA ICT Adoption Questionnaires (1999 – 2009)
http://departments.agri.huji.ac.il/economics/voet-gelb.pdf

Gelb, E., 2009. Constraints and Innovations of Agricultural Informatics. Proceedings Applications of Agricultural Informatics, Taiwan Seminar.

Gelb, E., Gal, B., Wolfson, D., 2010 - Information and Communication Technologies (ICT) for Agricultural Extension – An Israeli Case Study of ICTs for Agricultural Extension: Global Experiments, Innovation and Experience. New India Publishing Agency

Pehu, E.,et, al. 2011. ICT in Agriculture - Connecting Smallholders to Knowledge, Networks, and Institutions. Report Number 64605, IBRD, ARD, INFODEV.
http://www.ictinagriculture.org/ictinag/content/ict-agriculture-sourcebook

Rogers, E., (1962). Diffusion of Innovations. Free Press ISBN 0-7432-2209-1

Rosskoff, K., Wagner, P., 2003 Requirements for Agricultural Software and reasons for Adoption Constraints - Results of Empirical Studies. Proceedings EFITA Conference, Debrecen, Hungary.

Standage, T., 1998. The Victorian Internet.

Taragula, N., 2005. Information and Communication Technology (ICT) Adoption in Horticulture: A comparison to the EFITA Questionnaire's Baseline. Proceedings, International Society for Horticultural Science', Berlin

USDA, 2011. Farm computer use and ownership, 2011.
http://usda.mannlib.cornell.edu/usda/current/FarmComp/FarmComp-08-12-2011.pdf

# Current situation on data exchange in agriculture in the EU27 and Switzerland

Frederick Teye[1], Henri Holster[2], Liisa Pesonen[1], Sarka Horakova[3]
*[1] MTT Agrifood research Finland, [2] Wageningen UR Livestock Research, [3] Wirelessinfo*
E-mail: Frederick.teye@mtt.fi

**Introduction**

Within the knowledge-based bio-economy, information sharing is an important issue. In agri-food business, this is a complex issue because many aspects and dimensions play a role. Information systems for agriculture lack standardization, which hampers efficient exchange of information (Nikkilä, Seilonen & Koskinen 2010). This issue subsequently leads to inefficient business processes and hampers adaption of new knowledge and technology (Sørensen, Fountas, Nash, Pesonen, Bochtis, Pedersen, Basso, & Blackmore 2010). The exchange of information at whole chain or network level is poorly organized. Although arable and livestock farming systems have their own specific needs, there are many similarities in the need for an integrated approach. Spatial data increasingly plays an important role in agriculture (Gebbers & Adamchuk 2010).

CGI (2006) described meaning of information sharing and standardisation with the following words: "Companies must be prepared to share standards-based data free of charge. Sharing information between trading partners will result in an improved information flow and, as a consequence, improved collaboration to better serve the consumer. A resulting collaborative information platform could become the basis for further supply chain solutions". The publication of the Technology Platform (2010) on strategic European research agenda, noted that ICT will play a key role in particular in providing new information management systems and better communication between the different actors in addition to increasing productivity of farming systems. The publication also refers to future ICT for communicating values and providing tools for consumers to enable ethical decision making concerning food.

In order to contribute to a better harmonization of ICT development in European agri-business, the EU-funded project 'agriXchange' was started in 2010. The overall objective of 'agriXchange' is to coordinate and support the setting up of a sustainable network for developing a system for common data exchange in agriculture. This will be achieved by 1) establishing a platform on data exchange in agriculture in the EU, 2) developing a reference framework for interoperability of data exchange and 3) identifying the main challenges for harmonizing data exchange.

As a first step in this project, the state-of-the art of ICT and data exchange in agriculture was examined. This paper highlights the set up and results of research carried out into the state-of-the art of ICT and data exchange in agriculture in the EU member states (incl. Switzerland). Additionally, results on specific topics are discussed and general recommendations are defined.

## Materials and methods

The aim of this paper is to obtain an overview of state of the art of current data exchange in general within the EU and specifically within each EU region. Research on data exchange and data integration was carried out in 27 EU member states and Switzerland; in focus groups A to F.

| Responsible consortium | Focus group: Countries |
|---|---|
| Institut de l'Élevage, France | A: France, Switzerland |
| Altavia Company, Italy | B: Italy, Spain, Portugal, Greece, Cyprus, Malta, Bulgaria |
| Wageningen UR Livestock | C: UK, Ireland, Netherlands, Belgium, Luxembourg, |
| KTBL, Germany | D: Germany, Austria |
| MTT Agrifood Research Finland | E: Latvia, Estonia, Lithuania, Sweden, Finland, Denmark |
| Wirelessinfo, Czech Republic | F: Poland, Czech, Slovenia, Slovakia, Hungary, Romania |

*Table 1. Responsible consortium partners and their focus countries*

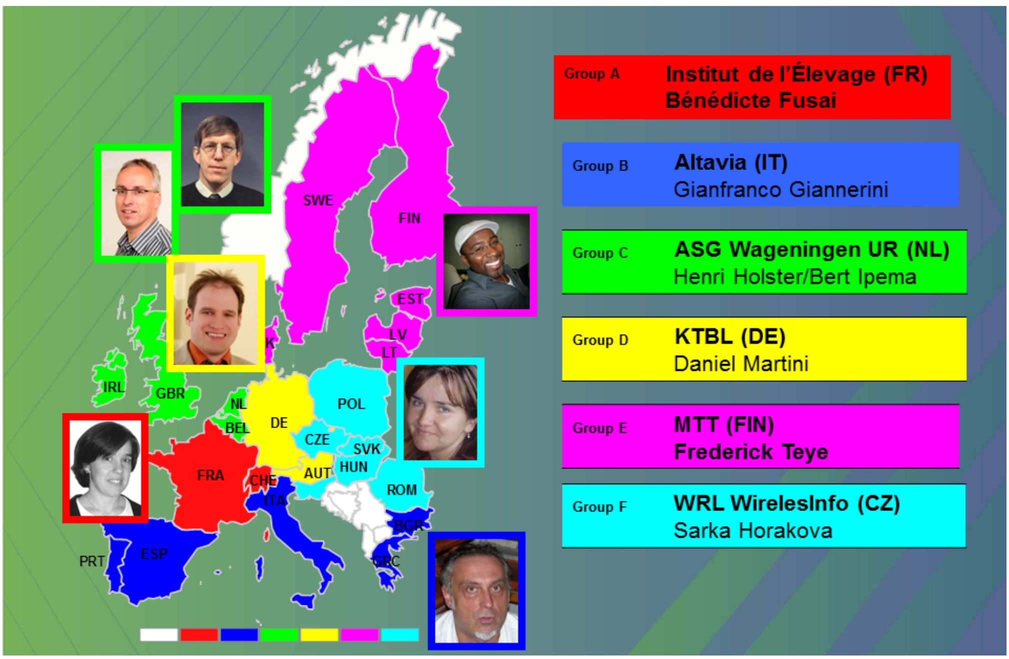



*Figure 1. Responsible consortium partners and their focus countries on the EU map*

The research has focused on farmers in connection with internal (on-farm) and external (business-chain) processes. For the research, semi-structured questioning, followed by telephone interviews were used in this study. The semi-structured template for the questionnaire was developed based on the framework by Giachetti (2004) to facilitate enquiring and analyzing information on network and integration between enterprises. The

investigation employed experts for a quantitative and qualitative inquire. The experts located in each country provided information about agricultural data exchange in their respective countries. For each country separate reports were prepared. In the analysis of the results, the Giachetti framework was used to "map" the current state of system and information integration in the research countries.
Information about country wise agricultural characteristics was obtained from Eurostat (Anon 2010).

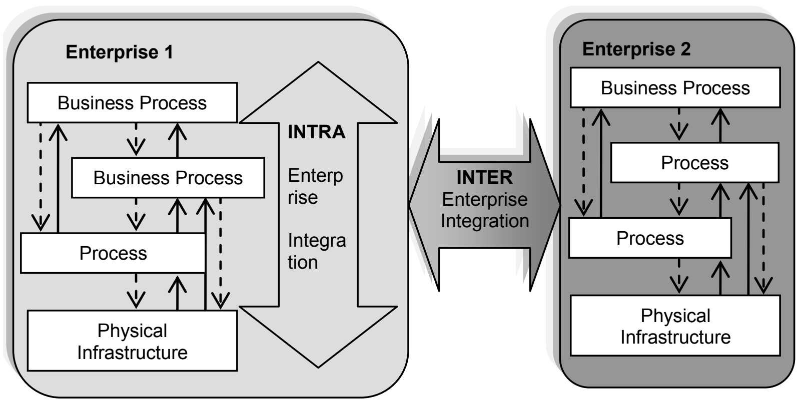



*Figure 2. Generic Integration Framework (adapted from Giachetti 2004)*

The mapping covered processes, applications, data, and physical infrastructure in agricultural data exchange with particular reference to characteristics of farms, the level of automation, data integration, ICT and technology usage. A summary of the key results are given in the next section.

**Results and discussions**

<u>Trends in farm characteristics:</u>

According the results of the research from 2003 to 2007 there was small decrease in the total utilized agricultural area (UAA) in the 27 EU member states. However, the number of agricultural holdings decreased in the same period by 1.3 million (more than 8%). This resulted in an increase in the area per holding of more than 9%. Arable farms are largest in the Czech Republic, Denmark, UK and France. In Bulgaria, Romania, Slovakia and Hungary more than 50% of the arable holdings have land areas of less than 2 ha (Figure 3). The largest dairy farms were in Denmark, Cyprus, Czech Republic and the UK. In Romania, Bulgaria, Lithuania, Slovakia, Latvia, Poland, Estonia and Hungary, more than 80% of the dairy holdings have less than 10 cows.

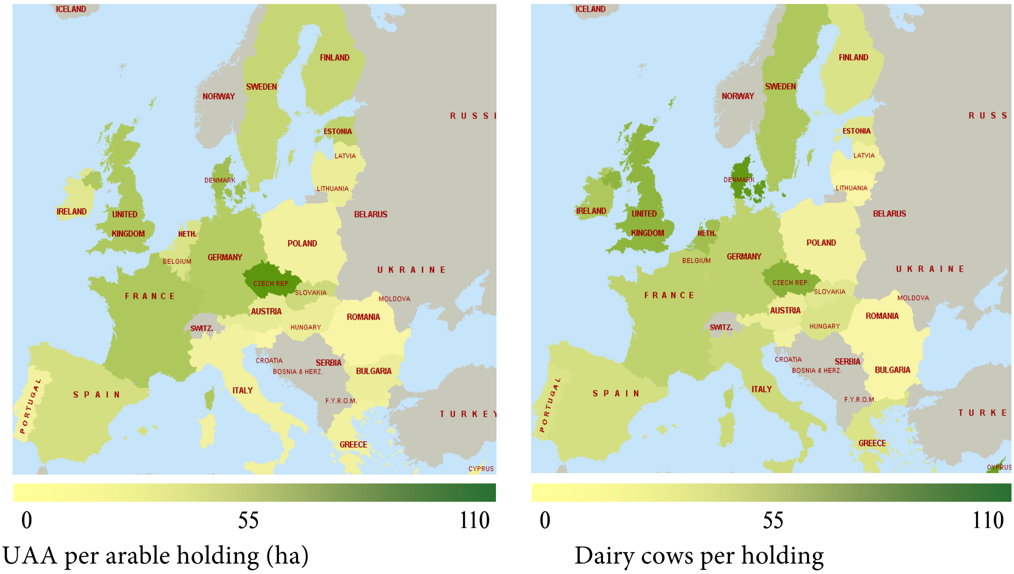



*Figure 3. Size of farms*

Farm automation level:

Detailed statistics about the automation level in the agricultural sector of all EU-27 states is hardly available. From the country reports available, it can be concluded that in agricultural enterprises, farmers internet access ranges from country to country is between 20 and 95%. Comparable statistics show that in January 2009 in average 93% of all enterprises in the EU had access to internet (Smihily & Storm 2010). Mobile internet connections were used by 28% of all enterprises in the EU with lowest values for Greece, Cyprus and Romania.

Some countries mentioned that the adaption of new technologies is lower in older farmers than in younger farmers. Lööf & Seybert (2009) also observer that generally more than 66% of the female individuals and 70% of the male individuals in the age groups below 55 years used the internet at least once a week, while in the group above 54 years these percentages were only 26 and 38%.

Looking at the basic level of access to mobile data infrastructure from several countries it is reported that the usage of mobile phones and even smart phones (with real internet application –apps- capabilities) is rather high or expected to increase drastically in the coming years. The information received showing that more than 90% of farmers using 'mobiles' in countries like Italy, Ireland, Spain are rather contrary compared to the figures of internet access by PC in those countries. It should be noticed that in these countries, mobile devices are far more used as cell phones than for mobile applications based on data communication. Experts estimate the usage of mobile application for business purposes between 2-5% in these countries.

By characterizing precision farming (PF) as a measure of farm automation, in most EU countries PF is only used to a small extent by farmers. A lot of experts reported the existence of PF and the usage of Geo spatial data only in experimental (research) projects. However, there is a significant difference in areas across Europe, in Western and North Europe and for example in Czech Republic there is more progress in PF development. Manufacturers of agricultural machines are the main booster for adaptation of PF techniques in developed countries such as Germany, the Netherlands, Denmark and Finland.

Data integration:

In general big differences all over Europe can be seen in data integration at process level. In a lot of countries in Europe data integration at process level is hardly organised at intra-enterprise and inter-enterprise level. The availability and accessibility of (broadband) Internet in rural areas is an issue in most countries. Except from some countries like Germany, France, Denmark, Belgium and the Netherlands, no (private) unions or bodies are reported who take care of the organization of dataflow or standardization. Collaboration between private and public organization to advanced infrastructure is also low in countries like Romania, Slovakia, Czech Republic and Lithuania.
In many EU countries data definitions (semantics) have only public standards (XML schema's and web services for example) mentioned. Standards definitions such as ISOBUS are available for example machinery (ISO 11783), milking equipment (ISO 11788 ADED), electronic animal identification (ISO 11784/11785/14223 and 24631) or forestry (ISO 19115). Syntaxes for EDI messaging from agroXML (Germany, some other countries), ISOagrinet (international), Agro EDI Europe and E-daplos (France), AgroConnect (The Netherlands) were reported. However, data integration along the whole food chain from farm to consumer is still lacking.

ICT and technology usage:

Agricultural technology adaptation and developments are not always positive because of lack of young people in agriculture (Slovakia, Bulgaria, Italy, others). Countries having a lot of small (probably poor as well) farmers are facing severe problems in the capabilities of investing in automation. Fast developing agricultural countries like the Baltic States have high potentials concerning the building of new ICT infrastructure as they are not bothered by old systems and structures.
Availability of broadband internet in rural areas is very often mentioned as a big issue that hampers ICT adaptation in agriculture. A summary of the level of ICT and technology adaptation in some of the EU countries and Switzerland is presented in next Table.

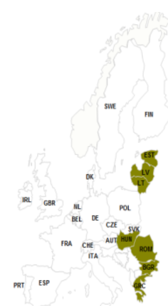
| | Farm | | Farm | Phones/ | | Geo - | | Data Exchange |
|---|---|---|---|---|---|---|---|---|
| **Country** | **PC** | **Internet** | **Info. Sys.** | **Handheld** | **LPIS relevance** | **Fertilizing** | **Animal Registra - tion** | **Level of Development** |
| BGR | Low | Low | Low | - | Average | - | - | Hardly any |
| CZE | High | High | High | Low | Average | Average | - | Averagely |
| DNK | High | High | - | High | High | Average | High | Well |
| EST | High | High | Average | - | Average | Low | Average | Poorly |
| FIN | High | High | High | High | High | Average | High | Well |
| FRA | High | Average | Average | High | High | Average | High | Well |
| HUN | Average | Average | Low | Low | Average | Low | Average | Poorly |
| ITA | Average | Low | Average | Average | Average | Average | High | Average |
| LVA | Low | High | Low | - | Average | Low | High | Poorly |
| NLD | High | High | High | High | High | Average | High | Well |
| POL | Average | Average | Average | - | Average | Low | Average | Hardly any |
| ROM | Low | Low | Low | Low | Average | - | Average | Hardly any |
| SVK | High | Average | Low | Low | Average | Low | Average | Poorly |
| SVN | Low | Low | Low | Low | Average | - | Average | Poorly |
| ESP | High | - | Average | Low | High | Low | High | Averagely |
| SWE | High | - | - | High | High | Low | High | - |
| CHE | High | Average | Average | Low | Average | Low | High | Averagely |

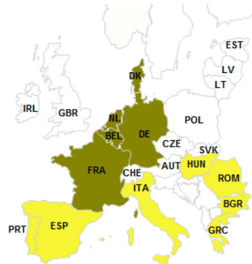
*Table 2. Level of ICT and technology adaptation in the EU countries and Switzerland*

<u>Different European regions:</u>

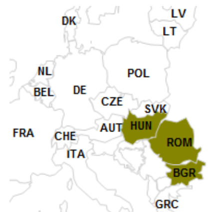
Looking at data integration in Europe it can be mapped in several regions following some specific characteristics which carry their own challenges and recommendations.

*Region with most small farms and poor farmers.*
In this region there is no standardization and hardly any ICT. Here data structures should be organized by public services to get developments started. The import of systems and data standards through private business (through multinational trade) will help as well. Last recommendation for this region is to copy knowledge (learn) from obligatory public services from other countries with structured and standardized ICT systems for agricultural data transfer.

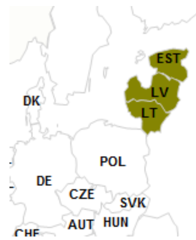




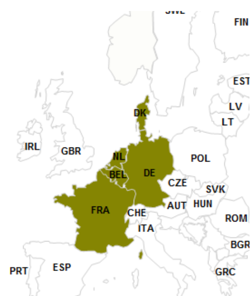

*Region where a focus on ICT is highly related to basic local challenges.*
For example water management, erosion (in Southern parts) or trade (more in Western- and Northern Europe).
Some recommendations concerning the trade group are that business and governments should work on cross-border trade support by international adapted data integration systems. The setup of new standards upon or through international (not national) supported bodies is another helper. Regarding water management the implementation of best practices from other countries and the implementation of integration GPS and sensor data, smart phone apps for easier data access are mentioned.




*Region with problematic aging and so adapting ICT by farmers.*
In these countries the effect of adapting new technology should be demonstrated, extended by education, not only for learning purposes but also to create a new and enthusiastic working environment for younger aged workers.




*Region with fast upcoming production areas which are relatively new countries in agri IT as well.*
Here stimulating the building of new internet infrastructure, especially mobile broadband will is crucial. Parallel to this, new and existing international standards should be adapted in a smart way. A last recommendation is to work on cross border trade and data integrated systems.




*Region with a standardization past ('old fashioned' structures).*
Old systems and structures on data integration can hinder progressive development. Based on this assumption, recommended solutions can be achieved through the renewing and redesign of old structures by private-public (cooperated) investments. Investments in open data infrastructures, stimulating by governments and stimulating of open innovation environment is worthwhile. This should be done by private-public investment on new common data exchange structures.

Data exchange – levels on data integration in the EU:
Having a look at the development stage of countries in data integration, Europe can be split into 4 levels. The levels and short characterizations are shown in Figure 4.
Level 1 - None or hardly any data integration (like BGR, ROM, MLT, CYP)
No private initiatives, public (CAP) providing rather closed (registry-) databases.

Level 2 - Poorly developed (mostly Southern and Eastern, Baltic States)
A move towards data integration has been initiated by CAP/Governments through interfaces. Some shared databases and portals are showing up. But still hardly any integrated private systems exist.

Level 3 - Rather well developed (Scandinavian states, CZE, GBR, IRL, BEL)
Involvement by private organization is evolving. A few data dictionaries are showing up and used.

Level 4- Quite well developed (FRA, DEU, NLD, DNK)
System assessment and move towards open/shared communities. Existence of private standardization bodies. Usage of national, private owned and global standards. Infrastructure based on hub structures (communicating and transporting systems). Further developing towards integrated models.

Referring to data integration model, presently even level 4 is far from perfect with regards to data integration. Data integration at inter-enterprise level is often well organised but from intra-enterprise point of view there is hardly any efficient and structured infrastructure. Data exchanging processes are drawn like a spaghetti mess. Future integrated models are evolving towards an infrastructure more or less made up of open hub systems where those hubs have the functionality not only to transport data in a highly efficient way but also to be the place to add value and distribute it through services.

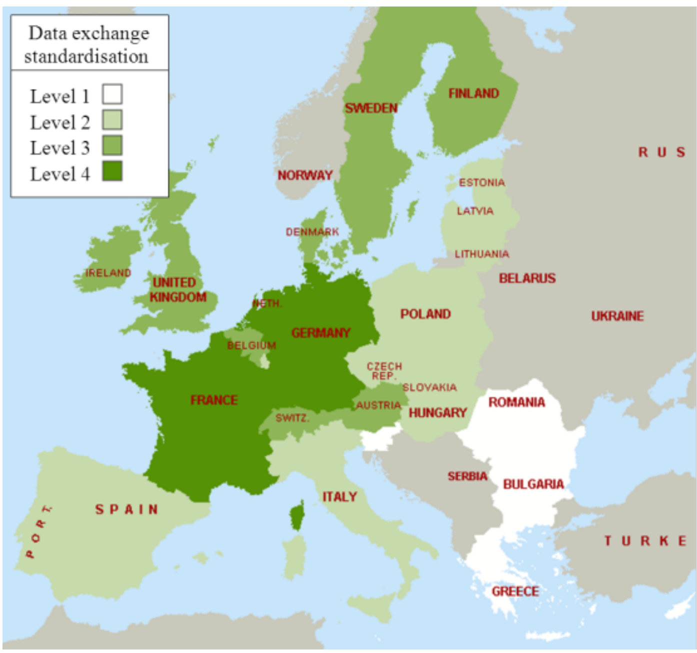



*Figure 4. Levels of data integration in the EU member states*

**Recommendations per level on data integration:**

Regarding the levels on data integration in Europe next recommendations per level can be given to policy and business.

Level 1: None or hardly any data integration

I. Invest in mobile broadband infrastructure
II. Build reliable public (CAP) services and extend them with webservices to provide private business with development opportunities
III. Invest in new technology adaption through education & demonstration
IV. Implement the (most easy) best practices from other EU regions
V. Connect to existing global standards

Level 2: Poorly developed data integration

I. Invest in mobile broadband infrastructure
II. Stimulate developing of shared databases
III. Demonstrate best practices in IT services (through local/EU subsidies
IV. Organize private business in setting up integrated systems

Level 3: Rather good data integration

I. Invest in mobile broadband infrastructure
II. Activate and organize private business in collaboration on shared and integrated systems
III. Work on adapting by demonstrating best practices in IT services
IV. Involve governments/public in setting up common shared structures
V. Get involved in European or global standard developments and implementations

Level 4: Fairly well developed data integration

I. Combine/redesign the best of several standards in different nations, like EDI-teelt-agroXML-E-Daplos
II. Take the lead in new international data integration initiatives and use international/global standardization bodies like UN/CEFACT
III. Initiate new private-public collaborations on redesigned shared data infrastructure. Direct this through private-public platform(s)
IV. Develop integrated business process models
V. Organize discussions and get directions how to become an open EU information society

**Meaning of standardization:**

Concerning the meaning and value of standardization experts have the opinion that standardisation should be done at the business service layers and not on processes, because of competitive surroundings. Focus should be put on demonstrating how processes can work, but keep them flexible and hence keep them out of the rigidity of (formal) standardization processes.

GCI (referred to in the introduction) envisions an open network, with flexible relationships between network partners, which implies less hierarchical or linear chain structures. This has consequences on innovations that will be developed within these open networks, as well as the

ever changing anonymous partners in the system. There will be less focus on the products themselves rather more emphasis on services.
These visions are leading to a new phase in the agriXchange project. Another needed research focus is to review what standardization enables or prevents under the afore-discussed contexts. In addition, the pros and the cons should be identified in the various phases of the project.

**Conclusions**

Based on the analysis of the state of the art in the present report, the challenges for future research and development of data exchange in the European agriculture were identified. Aging population of farmers which manifests itself through the lack of adaption and investments in new technology, especially in Southern and Eastern countries was noted. Lack of broadband availability in rural areas was reported. Furthermore, mobile infrastructures in most countries are not capable of sustaining the potential of use of mobile computing base for data communications which needed in emerging technologies for farm production. There is potential for countries that are developing quickly to adapt new data exchange infrastructural models and skip the complex and inefficient structures that currently exist in some EU countries. There are substantial differences across the EU in relation to the level of data integration and standardization ranging from none or hardly any data integration to quite well developed (France, Germany, The Netherlands and Denmark).
The importance of agricultural data exchange in the EU has been realized and new standardization approach has started. As a recommendation, now it is that open networks with flexible relationships between network partners will facilitate successful integration of the systems. But still technology service providers and users need to be convinced about the benefits.
Finally, focus should be on putting research information into practice to demonstrate data harmonization processes, however, this should be kept flexible and hence keep the rigidity of (formal) standardization processes minimal in agricultural data harmonization. Another research need is to identify what standardization enables or prevents in the exchange of agricultural data.

**Acknowledgements:**

The agriXchange project is funded by the EU within the FP7 programme for KBBE (Knowledge-Based Bio Economy). The focus groups (Table 1) and respective country experts who contributed to the information reproduced in this paper are acknowledged.

For information about the country reports:
(Focus group A) Institut de l'Élevage, France: benedicte.fusai@idele.fr
(Focus group B) Altavia Company, Italy: g.giannerini@altavianet.it
(Focus group C) Wageningen UR Livestock Research, The Netherlands: henri.holster@wur.nl
(Focus group D) KTBL, Germany: d.martini@ktbl.de
(Focus group E) MTT Agrifood Research, Finland: frederick.teye@mtt.fi
(Focus group F) Wirelessinfo, Czech Republic: horakova@wirelessinfo.cz
More information can be found at http://www.agrixchange.eu.

**References**


Holster, H, Fusai, B, Giannerini, G, Horakova, S, Martini, D, Teye, F & Ipema, B

2011. *State of the art data exchange in agriculture in the EU27 & Switzerland*. In: EFITA / WCCA 2011, 367-381

Anon 2010. Eurostat Website: http://ec.europa.eu/eurostat

Technology Platform 'Organics' 2010. *European Strategy and Action plan towards a sustainable bio-based economy by 2020.*

Lööf, A & Seybert, H 2009. *Internet usage in 2009 – Households and individuals*. Eurostat: Data in focus 46/2009, 6p

Smihily, M & Storm, H 2009. *ICT usage in enterprises 2009*. Eurostat: Data in focus 1/2010, 8p

http://www.tporganics.eu/upload/TPOrganics_ImplementationActionPlan.pdf.

Nikkilä, R, Seilonen, I & Koskinen K 2010. *Software Architecture for Farm Management Information Systems in Precision Agriculture*. In: Computers and Electronics in Agriculture 70 (2): 328-336.

Sørensen, C, Fountas, S, Nash, E, Pesonen, L, Bochtis, D, Pedersen, S, Basso, B & Blackmore, S 2010. *Conceptual Model of a Future Farm Management Information System*. In: Computers and Electronics in Agriculture 72: 37-47.

Gebbers, R & Adamchuk, V 2010. *Precision Agriculture and Food Security*. Science 327: 828-831.

GCI2006. *2016: The Future Value Chain*. Global Commerce Initiative, Capgemini& Intel Publication. 45p.

Giachetti, E 2004. *A framework to review the information integration of the enterprise.* International Journal of Production Research 42: 1147-1166.

# Section II
# Future challenges

# From external drivers to future challenges

Karel Charvat
*Czech Centre for Science and Society*
E-mail: charvat@ccss.cz

**Introduction**

What are the future challenges? What will influence the future rural development and farming sector? What could be the future trends?
Chapter II is divided into five parts:

- Analysis of external drivers for future agri-food sector and rural development and definition of challenges for future.
- Article with focus on territorial sustainability and vocational training for rural development.
- Using Grounded theory for knowledge discovery within Living Lab in African content.
- The next challenge - Exchanges of Environmental Information
- Use of environmental modeling for preventing the impacts of major accidents with dangerous substances
- Using methods of social validation for building regional geospatial infrastructure with focus on environment and tourism.
- The role of web technologies in rural and regional development.

**Future Farm**

Within the Future Farm project [5], [6] a trans-European investigation has led to the definition of the key objectives needed to realise this vision of a new concept of farming knowledge management respecting changing conditions and demands. As a result of the analysis provided by Future Farm, the project recognised the following groups of drivers, which will have influence on farm management and which could stimulate new demand on knowledge management [1]:

- **Climate change** – and its influence on crop composition and management methods.
- **Growing population –** will stimulate growing request on food and on energy.
- **Energy cost –** will generate new requirements for new methods of energy production.
- **Urbanization and land abandonment –** will lead to changes in society and land use.
- **Quality of food** - requirements of citizens and market on higher quality of food production.
- **Aging population and health problems –** will generate specific retirements on food production and diets.

- **Ethnical and cultural changes** – will generate specific requirements on food composition (growing muslim population in Europe, growing number of vegetarians).
- **Knowledge based bio economy** – will introduce new products and crops, including GMO.
- **Regulations and standards** – agreed government norms for the production and use of energy and protection of the environment.
- **Economic instruments** – market-based instruments (e.g., taxes, tradeable permits) to internalise externalities and promote the cost-effectiveness of energy and environmental policies and measures.
- **Subsidies** – phase-out of unproductive and distortive government subsidies (e.g. to energy, transport) and provision of transition supports where a need to ease environmental and social costs of change is necessary.
- **Investment**s – establishment of undistorted, cost-reflective prices in the energy market and conducive investment conditions to send the right signals to private investors.
- **Partnerships and voluntary agreements** – joint public/private programmes to develop and deploy sustainable energy approaches with industry.
- **Research and development** – government R&D and incentives to private R&D to promote innovation on energy for sustainable development.
- **Information and communications** – campaigns to promote better understanding by the general public of the national and international energy and environment situation and future challenges.
- **Assessments and scenarios** – sustainability assessments which identify synergies and trade-offs across the economic, environmental and social impacts of energy policy options.
- **Valuation of ecological performances** - as long as costs are mainly externalised into the direction of the environment (including agriculture and forestry) and no legal framework exists, necessary changes will be slow.
- **National strategies** – good governance approaches based on whole-of-government decision-making, transparency, and understanding of the political economy of promoting change in energy systems.
- **Politicians** - including their political awareness of the existing situation and the necessary changes in the future, including impacts and changes due to environment or technology.
- **Press** - due to their influence on politicians and their role of being intermediate between public and politics.

- **Education** - including training and know-how transfer and the awareness of the necessary speed of future changes.
- **Cooperation and integration models** - the complex chain in agriculture, biomass and environment as well as the complex structure of chain partners, their behaviour addressing new targets including control.
- **International organisations** - including Word Bank, FAO, CGIAR, etc. with their power/non power to influence necessary changes.

The experts in the Future Farm project analysed a these external drivers and grouped them according to their importance and potential influence on the future development of agriculture. The primary group of drivers, which will be the key issues for a future roadmap, consists of the following drivers [2]:

- Climate change;
- Growing population;
- Energy cost;
- Urbanisation and land abandonment;
- Aging population and health problems;
- Ethnical and cultural changes.

These drivers will have influence on each other and, in some cases, could have opposite influences on a number of factors.
The influence of the following group of drivers is not clearly evident, because in some cases it is difficult to predict, what will happen:

- Politicians;
- Press;
- International organisations.

Important drivers that are dependent on the primary drivers but relatively independent from other drivers are:

- Food quality and safety.

There are two groups of drivers, where all of them have strong synergies among single drivers. The first group is:

- Knowledge based bioeconomy;
- Research and development;
- Information and communication;
- Education;

- Investment.

The second strong synergy group comprises:

- Partnerships, cooperation and integration and voluntary agreements;
- Development of sustainable agriculture;
- Valuation of ecological performances.

Last group with relatively high synergies are:

- Subsidies;
- Standardisation and regulation;
- National strategies for rural development,
- Economical instruments.

Inside of Future Farm these external drivers were analysed also in cooperation with the project's pilot farms [3]. The opinion of stakeholders was discussed and analysed. There were some differences in stakeholders given by geographical conditions, size of farms, but also current experiences, used information systems, level of subsidies, etc. However, it was possible to find some trends common to all stakeholders. In principle, on the basis of these discussion and analysis, it is possible to divide the external drivers into four groups from the point of view of farming expectation and influences of farming production:

1. Drivers with mainly positive influence in farming;
2. Drivers with mainly negative influence in farming;
3. Drivers where could be expected both negative and positive influence;
4. Drivers, where farmers are not able to give a lear opinion or where it is not possible to predict their influence.

**Mainly negative**

- Energy cost is expected to have mainly negative influence on cost of production. Production of energy is not considered as important.

**Mainly Positive**

- Growing population is expected that it will support market growth.
- Investments are expected to have a positive influence. It is expected that it could exclude some farms from production.
- Partnerships, voluntary agreements, cooperation and integration models can bring higher stability.
- Research and development, information and communications and education will bring new opportunities, but they can exclude some farms from production.
- Valuation of ecological performances will be mainly positive, but it can bring different opportunities for different countries.

**Both**

- Climate changes - in some countries this can have positive influence, but mainly in south countries it can destroy production.
- Regulations and standards, economic instruments and subsidies can influence production in both directions.
- Urbanization and land abandonment will bring new possibilities for some farms, but there can be a problem with qualified employees.
- Quality of food - the main question is, if people will be able to pay for higher quality food.
- Aging population and health problems can bring problems with employers, but it can stimulate new products.

**Not clear**

Not clear drivers were difficult to understand for farmers. They don't have a clear opinion. Some drivers seems to be not predictable. This group includes knowledge based bio economy, assessments and scenarios, international organizations, politicians, press, national strategies, ethnical and cultural changes.

The provided exercises with the farmers demonstrate in some cases different opinions among farmers and expert's analysis. The analysis of farmers answers shows the influenced by immediate expectations rather than long time vision. The answers are usually influenced by current expertise, for example the answers related to the costs of energy. The growing costs for energy are direct risks for farmers. The future potential from energy production was not so visible.

Some external drivers are not fully understandable or predictable for farmers. It is important for farmers to explain what could be the influence of some drivers on their production. It is clear from the analysis that within the design of farm management system is necessary to take into consideration some external drivers. It should be possible to quantify the influence of the drivers on the farm and offer to farmers different possibilities on how to deal with concrete external drivers or how to adopt farm behaviour on the basis of external drivers. It will be important in future management systems to give to the farmers qualitative information about different potential scenarios based on external drivers. It is important to note that for many drivers there is no opinion, if they could bring positive or negative influences. In principle, they can have both negative and positive influences. Is seems important to have a chance to build different scenarios and present them to farmers. They could then adopt their management system on the basis of the new conditions.

**From drivers to challenges**

As it was already mentioned, in the next period rural Europe will be radically transformed in terms of the distribution of people and of economic activity and Gross its regions. These changes are inevitable. A common and future position of each important driver in reality can be different. In many cases two drivers can act against each other and their future influence on agri-production and food market depends on regulations and common policy. For example:

| | | |
|---|---|---|
| Food quality and safety | ⟺ | Food requirements for growing population. |
| Growing requirements for food | ⟺ | Renewable energy production technologies. |
| Renewable production energy demand | ⟺ | Demand on more environmentally friendly production. |

The combination of externals drivers will introduce new challenges for future agri production and also for all rural communities. We can recognise list of challenges, which has to be reflected by future farming knowledge management systems. We can name the following challenges for future:

- To include ICT and knowledge management for agri-food and rural communities generally as a vital part of the ICT policies and initiatives (for example Digital Agenda for Europe 2020). Previous analysis in Chapter I of this book demonstrated that threre exists ICT gap between urban and rural. This gap is more visible in developing countries, but also in Europe, North America, East Asia and other countries. Newly coming solutions as the single digital market, future internet, global of knowledge, social media and networks, protection of data and open access to information will be essential for farming and rural communities. Knowledge becomes one of the most important products and also material. If there is no equal access to knowledge by rural communities, the urban rural gap will grow and negative trends will continue e.g. abandonment of rural areas.
- **To find a balance between food safety and security, energy production and environment production** – As it was mentioned in previous chapter, current agriculture and food production is under many of different requirements. On one side, worldwide growing population brings new demands on food, but also on arable land, water, energy and other biological resources. There is growing demand for energy and growing costs of energy. On the other side, there exist requirements of citizens and market on higher quality of food production, but also on new products. It is related with aging population in some part of the world, but also with ethnical and cultural changes (migration of population). All of them have a strong influence on global ecosystem and environment generally. The influence is bi-directional and some issues like climatic changes, draughts, floods etc. have also large influence on production. We need to find a model of sustainabile agriculture production based on synergies and trade-offs across the economic, environmental and social impacts. This will require to find new methods of production, but also increasing the level of involvement of citizens in the agrifood value chain.
- Support better transfer of RTD results and innovation into everyday life of farmers, food industry and other rural communities - As the analysis in Chapter I demonstrated, there is well visible progress worldwide in roll out of basic

communication infrastructure. But the take up of new innovative rural solutions is not so fast and in many cases, there is a low transfer of RTD results and innovation into practice. To be able to react on future demands, the fast innovation transfer is necessary. It requires more parallel actions, like better interaction between support for RTD, innovation and implementation activities (in European scale Framework program and Leader activities for example), user demand on research priority, focus on long time of sustainability of founded research, but also support for training and awareness. Training of rural communities will be essential for adoption of new solutions and technologies. There is also necessary to stimulate open and public scientific debate about such controversial themes such as bio energy production or genetic modified production.

- **To accelerate bottom up activities as a driver for local and regional development** – aBard and DG Agri studies (Chapter I) recognised as the key success factors for development of rural regions local activities and existence of local champions (local actors who are not interested in technology but take up the role for the greater good of the communities they live in, younger and higher educated people moving to the rural areas, who wanted to have the a level of services as in cities, local business with clear innovative and knowledge based growing strategy that needs an access and sharing of knowledge, people moving to a rural area and wishing to engage in e-enabled, people who have holiday homes in the rural area, that want similar broadband access similar to cities and other activities). Existence of such champions has the same influence as subsidies on local and regional development. The experience demonstrated that in almost all projects there were no strong bottom-up interests. In order to support this we have to increase the areas in which rural industries and citizens find social spaces and collaborative working environments supporting local networking and initiatives.
- Making rural regions as an attractive place to live, invest and work, promoting knowledge and innovation for growth and creating more and better jobs - Abandonment of cultivated land has been increasing in many regions of the world. In several parts of Europe, the opposite trend exists with people moving back to countryside, but without returning to agriculture production. the number of people working in agriculture decreased in some countries to 2% of the country population. To attract people back to rural regions, it is necessary to improve infrastructure, possibilities of employment, culture, education, etc. It includes spatial planning, improvement of situation of local and regional SME industries with focus on agri-food industry but also on local production and tourism.
- **Build new ICT model for sharing and use of knowledge in rural regions**. Currently we can recognise more shift of technologies to web based and mobile solution, cloud computing, open access to context, social media, collaborative platforms and business intelligence. To build such solutions will help not only to rural communities, but will open also new business opportunities for the local and regional ICT industry through development of new applications and tools to support the European agri-food and rural sector. Participation of local ICT SMEs on development and implementation of local applications will play an important role in regional development.

## References

Charvat K. at all D1.1.1 List of external drivers, Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards.

Charvat K. at all D1.1.2, Analysis of external drivers, Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards.

Charvat K. at, D1.2.2, SWOT analysis of drivers and farms, Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards.

Charvat K. at all D1.2.3 Visions and recommendations for knowledge management, Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards.

Karel Charvat, Pavel Gnip Future Farm for Farm of Future, IST Africa 2011, Gabarone Botswana

Karel Charvát, Pavel Gnip, Walter Mayer FutureFarm vision, AGRIS on-line Papers in Economics and Informatics

# Sustainability of rural systems and Poles of excellence: the role of Information and Communication Technologies

Gian Paolo Cesaretti, Rosa Misso
*Simone Cesaretti Foundation, Department of Economic Studies, University"Parthenope"*
E-mail: misso@uniparthenope.it

**Introduction**

Today more than ever, the territorial systems with rural vocation demand new paradigms of development, able to respond both to internal than external pressures that are threatening to stifle their identity, the value and meaning they have in terms of economic, environmental, social and cultural sustainability.
Their projection in global markets, in particular, force the local systems, policy makers and their economic operators, to search continuously for innovative solutions able to face up the increasingly global challenges to the environment, economics and society.
In addition to the request of the usual managerial skills and high ability to adapt to market demands, the new model of "*rurality*" encouraged by the European Union, suggests a review of the components of the regions from a sustainability viewpoint.
In this context, the role that play the regional systems with rural vocation it becomes strategic to support the empowerment of its actors, to create value in a sustainable way, without setting aside the unbreakable bond that characterizes the territories, the society and its environment.
In the regions, in fact, social, environmental, economic and cultural variables are integrated and interrelated; or rather the regions are the junction box in which to engage the different dimensions of the development, to support socio-economic, environmental and cultural development initiatives, and which could ensure their sustainability. So this last becomes the key to interpret the ability to develop a local system and its distinguishing features with a long term strategy.
The sustainability of the development of a territorial system with rural vocation is due to the adequacy of its social, economic, environmental facilities and to its ability to dynamic adjustment to the changes of the global scenery.
The pursuit of this objective starts from the awareness that the development of territorial systems with rural vocation requires the presence of drivers of development not only endogenous, related to the presence of a system of excellence of the territories, but also exogenous, the so-called pole of rural excellence, able to get iterative dynamics of sustainable competitive development in the territorial system.
Based on these considerations, this study aims to identify in a territorial system with rural vocation of Campania region (an Italian region), on the one hand, the system of the endogenous excellence and on the other hand, the configuration of an appropriate tool of governance, examining at the same time the role that the modern information and communication technologies can play in the implementation of the pole of rural excellence.
Starting then from an integrated vision of growth factors exogenous and endogenous to the local systems, the study proposes a theoretical framework on sustainability through the definition of the system of excellence (such as driver of endogenous development). In addition, it proposes an overview on the governance of the rurality, with the identification of poles of excellence as exogenous driver of development and appropriate tools to manage the

processes of the territorial systems with rural vocation towards the sustainability of development.
The study, then, aims to make remarks about the possibility of implementing a sustainable rural governance, useful for policy makers and all those who for various reasons and with different responsibilities are managing the fortunes of rural areas of the Campania region.

**Territorial systems with rural vocation and sustainability**

In the recent years, the global trends have exposed the local systems with rural vocation to challenges that have begun to expect from their development processes, a declension in terms of sustainability. In particular, the complexity of these challenges has demanded strategic responses that were at the same time, articulated, based on an integrated vision, shared by all the stakeholders of the region and above all sustainable.
In fact, sustainability is a key theme in the current debate at theoretical and applied level.
The numerousness of interpretations and, often, their ambiguity, make it one of the most important objectives but difficult to achieve (Papini *et al.*, 1997). The main difficulty in dealing with the issue of sustainability is linked to its multidisciplinary feature, and thus to the complexity of its meaning.
This difficulty is often related to interpretations that do not take into account its multidimensional feature and, especially, the interrelationship and interdependence of its various dimensions (economic, social, environmental and cultural).
In the remarks and experiences on sustainability, so it has been stating the need to deepen and integrate the cultural or territorial issues as drivers of sustainable development processes (Sbordone, 2001).
The neoclassical economic literature has considered the region as an abstract location within which to observe the functioning of economic mechanisms, but today the regions have become a key factor with to explain the process of creating value in economy.
As acknowledged by several authors (Capello & Hoffmann, 1998; Cesaretti & Bianco, 2006), this innovation can be connected to distinct elements, but convergent.
Firstly, economic development is a complex process, in which the role of "territorial" resources (material and immaterial, and result of not only commercial relations) is central: this kind of remarks were proposed by the neoinstitutionalist geography that has founded the basis for a revision of the categories traditionally associated with the concept of proximity, reflecting on the centrality of relational factors and the relationship between the economy and institutions (Capello, 2004).
The second factor is the growing awareness that, in the formation of competitive advantage, play an important role the intangible factors connected to the non-commercial networks of interdependence between economic agents, society and institutions; these networks are structured in different ways, with original features that depend on local circumstances, but at the same time, they can be stylized on some models of collective action territorialized.
Thirdly, the processes of economic competition are played not so much between companies but between local clusters of which businesses are a constituent part; these areas, unlike nations, compete on the basis of a comparative advantage, but absolute, making it even more central the contribution of regional policies that support the local economy and society.
So, the territory has become the box in which to engage, to support economic development initiatives, socio-economic, environmental and cultural aspects that can ensure its

sustainability. In this way, this last becomes the key to interpret the ability to develop a local system and its distinguishing features with a long term strategy.

According Dematteis and Governa (2005), for example, territorial sustainability can be defined as the autonomous ability to create a regional added value in two ways: transforming in use or exchange value, the potential resources (immovable and specific) of a region (added value of the first kind) and incorporating in the land a new value in the form of increased territorial capital (added value of the second kind); therefore, we have the sustainable self-reproduction of a territorial system (self-sustainability) when the process of development is self-governed and it has as final result a regional added value of the first positive type and a second of non-negative type. Moreover, certain studies, in order to arrive at a definition of territorial sustainability, are based on the concept of quality of life of a region; that is, sustainability would be the capacity of economic, social and environmental systems to assure the community living in a specific region that there will be well-being, not only in the present, but also in the future. With regard to this, the traditional measures relating to the welfare of a certain territory, do not take in account of the multiple connections between the macro-areas constituting the territory itself, as if those parts were completely independent. In this context, of course, issues related to a part of the region are considered isolatedly and not related to the other parties.

However, today, territorial sustainability is an imperative objective of any action or development policy, and this applies more than ever, for systems with rural vocation. This condition, first of all, requires an appropriate vision of the set of variables that make up their *humus*.

The literature has established the strategic vision of the territory in which the analysis of all components, their different combination and their varying stages of rootedness on the same provide a more appropriate theoretical framework in the study of realities with rural vocation (Becattini, 2001).

At the same time, as recognized from Cacace *et al.* (2005) the analysis to support rural development highlights strongly the local characteristics, considering the development as a socio-economic, cultural and environmental phenomenon and result of "local" characteristics.

This, of course, pushes towards the recognition of non-homogeneity of the territory and to the need to read and interpret the diversity of rural areas, aiming to sustainable development trajectories calibrated to the specificities of the regions.

Moreover, the conclusions, in November 2008, of the verification process of the health status of the Common Agricultural Policy (launched in November 2007 with the European Commission Communication "*Preparing for the health check of the CAP reform*") have identified just in the development of territorial systems with rural vocation a fundamental factor for a sustainable future of Europe (Storti & Zumpano, 2009).

As pointed out by Nazzaro (2008), the Health Check took into account the strengthening of rural development, namely the transfer of more resources to the "second pillar" from two points of view: in financial terms, by proposing a more effective modulation (progressive cuts to the first pillar applied to the recipients of more support), and in terms of themes and new skills linked to the new challenges not only of economic type (such as food safety, quality, protection and enhancement of natural and cultural resources, animal welfare) but also of environmental (climate change, renewable energies, water management, protection of biodiversity), social (recovery of rural traditions, rural employment, revitalization of rural

areas) and territorial type (conservation of local resources; conservation of rural landscape and of its visual and cultural value; preservation of the heritage, traditions and cultural identity) (Barbieri & Mahoney, 2009; Morgan *et al.*, 2010; Pittman *et al.*, 2010).
This having be stated, it's clear that, as evidenced by Cesaretti & Regazzi (2007), the maintenance over time and space of the competitive advantages of rural systems requires new approaches and new tools able to re-read the relevant territories by the light of a new *culture of rurality (expression of an increasingly diversified and multi-functional agriculture with strong local roots)* with a multidimensional but integrated view of its specificity (Casini *et al.*, 2002).

**The system of rural excellence**

The sustainability of the development of a territorial system with rural vocation is due to the adequacy of its social, economic, environmental facilities and to its ability to dynamic adjustment to the changes of the global scenery, governed by policies inspired by a constant enhancement of its excellence.
The pursuit of this objective starts from the awareness that the development of territorial systems with rural vocation requires, first of all, the presence of endogenous drivers of development, related to the presence of a system of excellence of the relevant territory.
In this regard, too often, the inability to design in an interdependent way the trajectories of exploitation of the whole heritage of excellences, makes a model of territorial development unsustainable. The solution is instead the ability to seal the "project" of each region with a specific identity perceived internally and externally, as an integrated system of excellences.
The patrimony of a territorial system with rural vocation, in particular, is composed of several factors that altogether determine the system to valorizing for the acquisition of sustainable competitive advantages, and to transmitting to future generations to ensure reproducibility in the time of such advantages. Among the excellences of a territorial system with rural vocation, in particular, there are economic, social, environmental and territorial factors. These last can be tangible or intangible assets (Agarwal *et al.* 2009; Davico, 2004; Cesaretti *et al.* 1994; Coppola *et al.* 1988).

Economic factors, for example, include:

- characterization and typicalness of the local production
- structural characteristics and level of specialization or diversification of farms
- local financial system;
- system of infrastructure and public services;
- system of economic and production valorization (eg. the presence of brand of regional recognition);
- system of research and innovation
- the assets accumulated in infrastructure and facilities, considered as a whole and for the resulting externalities.

Among the environmental factors there are:

- The natural resource of the region;
- The system of environmental services for the region;
- The architectural assets and museum;
- The characterization of the agricultural landscape;
- Geographical position and climatic conditions.

The social factors, then, consist of:

- Presence of the youth component of human capital;
- Security;
- Equal opportunities in the territory;
- Health care system;
- Education and professional training system.

Finally, the territorial or cultural factors, include:

- Relational assets and values, embedded in local human capital, such as local knowledge capital, social capital, cultural diversity, institutional capacity;
- Sense of belonging to rural areas.

Since the development of a regional system with rural vocation draws force from all these components, the sustainability of its process is built on implementing strategies aimed to create synergies between the excellences. In particular, these strategies should consider the excellences as a single integrated system to valorize and maintain over time and space. In addition, this system should be considered as a true public good, worthy of institutional protection because of the characteristics of non-rivalry and non-exclusion of its factors (Stiglitz, 1998).

In this context, the sustainability of development of the territorial systems with rural vocation involves not only the ability to reproduce the territorial capital, but also and above all the self-reproduction of the integrated system of excellence, in order to keep in the time and the space its identity through a continuous change resulting from local innovations. In this regard, as reported by Haartsen & Strijker (2010), it should be underlined that there are various forms of regional identity. For example, the regional identities can be distinguished in six aspects (Groote *et al.*, 2000): social constructions, qualities and symbols of the region, past of this area, context of reference, arenas of power, dynamics (regional identities change over time and within changing context).

**A tool for sustainable governance of rural areas in Campania region**

Territoriality is at the origin of the development of rural systems and of their sustainability; but we have to consider that the definition of the relationship between territory and rural systems is a subject of considerable complexity, because it involves not only the relations that bind tightly these two concepts, but also the characteristics of organizational models present in a given area. This issue, therefore, has many facets to it, which draws attention to several studies, which are located exactly on the border between different disciplines.

Currently, administrative and geographical boundaries are no longer able to define spaces "economically homogeneous", because they are characterized by the presence of internal economies and economic areas with different characteristics and potential development. Approaches from different disciplines have interpreted and supported the growth and consolidation of different organizational forms of enterprises into a space: from the cluster of Porter, to the districts, to the milieu innovateur, to the theory of local production systems, up to the poles, technopoles, science parks and networks (Becattini, 1987; Carbone, 1992; Fabiani, 1991; Iacoponi, 1997; Iacoponi *et al.* 1995).

So, the concept of sustainability and its evolution and the idea of territory are equally complex.

The notion of territory, in fact, has always gave space to different and often controversial interpretations. For example, if the concept of region is interpreted only in a spatial perspective, we can define a space as a relative or absolute reality. In economic theory, classical and neoclassical, the notion of territory is often used in terms of relative space and, therefore, linked to issues of distance and cost. But there are many other definitions of territory that are not restricted to one-dimensional space and that find in it the expression of a particular place of integration of different natural and human components. In this sense, the territory is not a mere physical platform on which are located towns, roads, industries, natural resources or public institutions, but the place in which the social, political, cultural and economic relations are developed and strengthened.

There is, then, who looks to the territory as a system of governance and localized technological externalities (Camagni R. & Capello R., 2002). In this sense, the territory is at the heart of the various forms of geographical clustering of activities justified by the production of specific dynamics especially in innovation and employment and that strengthen the competitiveness of companies that belong to it.

Finally, there are those who, paying attention to the spread of networks and the Internet and overcoming its immateriality, has began to wonder if these new technologies could be considered as a vector of spatial dynamics (Cesaretti *et al.* 2006).

These considerations, in fact corroborate the idea that the interpretation of the territory cannot be restricted to the spatial dimension but rather that it becomes an indissoluble core that is characterized by the spatial, temporal and relational features of its components. Therefore, this having be considered and recognized that regional sustainability is a key element in the analysis of rural development processes, this study wanted to reflect on the possible actions that could support over time and space, the competitiveness of the above systems. In this regard, the strong links between rural systems and territory, makes the rural system as the ideal area on which to experiment innovative and sustainable systems of governance. The concept of governance, in essence, means rules, processes and behaviours that affect the way in which powers are exercised at a level territorially defined, particularly as regards the principles of openness, participation, accountability, effectiveness and coherence.

The tools and multi-objective strategies of the governance, in particular, are the natural completion of the actions of government and they are a factor of substantial advantage to enable the rural systems to respond jointly to the needs of health, wellness, quality of life, struggle climate change, maintaining rural employment and slowing the rural exodus of the countryside and the various other socio-economic, environmental and cultural objectives that development of the rural areas can achieve (Cesaretti & Scarpato 2010).

In such optic, the study aims to make remarks about the possibility of implementing a sustainable rural governance, useful for policy makers and all those who for various reasons and with different responsibilities are managing the fortunes of rural areas of the Campania region.

The Organization for Economic Cooperation and Development (OECD) (2006, p.17) considered that new rural paradigm requires important changes in how policies are conceived and implemented to include a cross-cutting and multi-level governance approach. In particular, as recognized by Natário & Neto (2009, p.125) two principles characterize the new rural paradigm: a focus on places instead of sectors and a focus on investments instead of subsidies.

Starting from these considerations and looking to the French experience regarding the rural territories of excellence, namely the Rural Excellence Poles (REP), the study aims to test the applicability of this initiative in Campania region. Really, the literature on the rural excellence poles is very poor, but the existing studies (Alvergne & De Roo, 2008; Constanta, 2008, Natário & Neto, 2009) underline the strategic role of these tools of governance in the pursuit of sustainability of the development.

Inspiring from the policy of urban competitiveness poles, these are projects that facilitate the development of the rural territories, receiving a partial finance support from the part of the state with the aim to stimulate a network of multidisciplinary investigations which favour permanent paths and results between research and action, to

capitalize the analysis to allow an evaluation with a prospective methodology and to join the investigation works and evaluations to place them at the disposal of investigators and eventually the collaborators and territorial collectivity which are interested (Natário & Neto 2009, p.132). In particular, the principle on which the REP is supported is the fertile and joined development between the university investigation and public territorial action.

According to Alvergne and De Roo (2008) the starting hypotheses of the REP are:

1. From the competitiveness poles to the diffusion of excellence
2. From the productive economy to the residential economy
3. From engineering to the investment
4. An interaction between the rural and the modern

Moreover, according to Natário & Neto (2009, p.134) the rural excellence poles privilege four priority areas in the development of rural territories:

- Excellence for the promotion of natural, cultural and tourism wealth (it is fundamental to develop relevance and attractiveness, the creation of new activities and professional training of lines, always supported by the ICTs);
- Excellence in the valuation and management of biological resources;
- Excellence in the offering of services and acceptance of new populations;
- Excellence in technology, for the agricultural, industrial, crafts and local services productions.

As it regards the Campania region, the Rural Excellence Poles represent a tank of growth and sustainable competitive development that could boost its rural attractiveness and development.

In such optics, the main target of the REP could be that to accompany the repositioning of the rural system of Campania region as a centre of gravitational attraction imagining a strategy to systematize the region and its process of dynamization of territorial systems with rural vocation.

This strategy should be implemented by a project that must meet certain fundamental characteristics: balancing, sustainability and identity.

Firstly, the project must be balanced between its components (infrastructure, knowledge and information) and in its work on business and local systems. Secondly, the project must meet

the requirement of sustainability: an integrated project in which all components (business and local systems) are placed in a network with the common goal of pursuing common trajectories of sustainable competitive development.
Third and final feature concerns the gravitational value of the project so that it can transform the Campania region system in the *gateway to Europe* through the affirmation of its local cultural identity.
As a network of specialized services based on knowledge, the pole of rural excellence becomes a real rural area of excellence capable of repositioning the rural systems of Campania region on the world scene.
The rural reality of Campania region, in fact, is characterized by heterogeneity, the presence of a fragmented production structure in small and micro enterprises. We have to consider, for example, that the average size of farms is 3.5 hectares. From this type of disgregation, together with a management highly focused on the company family, result the problems related to a lack of professionalism and training media is not qualified. In fact, we must not overlook characteristics intrinsic to the business reality of Campania region that is often an obstacle to certain initiatives of development.
However, the Rural Excellence Pole in Campania region requires the fundamental systematization of the territorial system whit rural vocation through the e-Campania, the rural agency of Campania region and the territorial centre of rural knowledge. In particular, knowledge and information economy is the core of the strategy of the REP.

**The strategic role of the Information & Communication Technologies**

Information and knowledge represent notions now in vogue in the debate among economists; some time ago, in fact, they have begun to wonder about the rapid spread of information and communication technology in all the fields of economic activities. The increasingly important role attributed to information, however, depends not only on structural changes that it produces, but also from its location in an evolutionary process that is characterized by the transition to an economy based on services (outsourcing), by the extension of markets and production combinations at the global level (globalization) and by the continuous increase in the general level of the stock of information and knowledge. According to Petit (1998) the diffusion of ICTs in different fields is not the only development that has helped to change the weight of information in our economy. Over time, this spread has interacted more and more with the internationalization of markets and production, as the new division of labour within or between companies. These structural changes are, therefore, a single body with technology changes and help to give them the shape of a revolution in the use of information.
In 1996, the European Commission stated "*in an information society, information and knowledge are the most important commodities*" (Commissione Europea, 1996).
Obviously, the rapid diffusion of ICTs can become a problem to the extent that the structural adjustments needed to gain competitive advantages are not implemented rapidly. In the usual process of "creative destruction" caused by technological progress, this quickness can cause concern that the destruction of jobs, through a gain of competitiveness, does not lead to a short-term job creation by offering new products and services.
So far, the greatest benefits were acquired just from those countries which did not oppose the information age and identified as "information society", from the beginning, are taking advantage of all the benefits associated with it, not only economic but also social and cultural: it is the United States, Japan, Great Britain or Germany, and all those countries with the same

lifestyle. But benefits will arise for those organizations that will respond quickly to local rules that govern the modern economy.
In this context, the Campania region, through the implementation of the rural excellence poles could take the opportunity to initiate an important process of diffusion of ICTs in rural areas in order to bring out the best practices, determine their link with the specialized poles in the rural knowledge and ensure their full integration with the urban areas and production centres.
In practice, the creation of the e-*rural Campania* region will require physical e-infrastructure and, therefore, investment to achieve them. Collaborative platforms, websites and databases will become an indispensable component of the new rural society, while it will be strategic the diffusion and use of information, Internet and telecommunications.
The reorganization of territorial systems in rural vocation of the Campania region, using modern information and communication technologies, will facilitate the government and the management of complexity of the region, weaving a network of rural areas, excellent production of food and wine and cultural heritage.
Among the ICT tools for rural systems, for example, there is the creation of a web portal. This last will be an irreplaceable vehicle to give visibility, at a global level, to those systems; at the same time, it will provide a valuable tool for the valorization of the area and of the local handicraft, propagating incentive and promotion activities, and encouraging the exchange of expertise and national and international experience. Innovative instruments also lead to a new packaging, new conservation techniques and quality control, greater visibility of these products in the international arena through the WEB.

## Conclusions

The Information Society and the New Economy, based on virtual networking and Knowledge oriented activities are a reality in different regions of the world. Basically, the question of ICT development in rural areas is addressing a digital divide issue; moreover, lack of advanced telecommunications infrastructure, capable of supporting new ICT applications and services creates a very tangible technical and financial barrier in such territories (OECD, 2001).
Given these problems, however, we have to consider that the many objectives that the international, national and EU policies aim to reach relatively to new model of rural development, have required the emergence of new tools of sustainable governance. Among these, the rural excellence pole was proposed particularly in France as a tool for organizational innovation of the territory based on ICTs, arousing interest in the possibility of implementing the same in several other regions. For the Campania region, with its system of rural landscapes, historical, cultural and agri-food excellences, the rural excellence pole could be an important driving force for sustainable competitive development able to allow the maintenance in time and space of its cultural identity through the diffusion of technological innovation and knowledge.

## References

Agarwal, S., Rahman, S., Errington, A. 2009. Measuring the determinants of relative economic performance of rural areas, Journal of Rural Studies, n.25, pp. 309-321.

Alvergne, C., De Roo, P. 2008. Les pôles d'excellence rurale: regards sur une politique française en direction de l'espace rural (the pole of rural excellence: a look at a French policy towards the countryside). Organisations &Territoires. Volume 17. nº1 – Hiver. 73–78.

Barbieri, C., Mahoney, E. 2009. Why is diversification an attractive farm adjustment strategy? Insights from Texas farmers and ranchers, Journal of Rural Studies, n.25, pp. 58-66.
Becattini, G. 2001. Distretti e ruralità: sfide al riduzionismo economico. Una replica (Districts and rural areas: challenges to economic reductionism. A reply). La Questione agraria, n. 1, pp. 119-127.
Becattini, G. 1987. Il distretto industriale marshalliano: cronaca di un ritrovamento (The Marshallian industrial district: report of a discovery), in Becattini, G (ed), Mercato e forze locali (Market and local forces). Il Mulino, Bologna.
Cacace, D., Falessi, A., Marotta, G. (eds) 2005, I sistemi agroalimentari e rurali in Campania (Agrifood and rural systems in Campania region). FrancoAngeli, Milano.
Camagni R., Capello R. 2002. Apprendimento collettivo e competitività territoriale (Collective learning and regional competitiveness). Franco Angeli, Milano.
Capello, R. 2004. Economia regionale – Localizzazione, crescita regionale e sviluppo locale (Regional economics - Location, regional growth and local development). Il Mulino, Bologna.
Capello R., Hoffmann, A. 1998. Sviluppo urbano e sviluppo rurale tra globalizzazione e sostenibilità (Urban development and rural development between globalization and sustainability). Franco Angeli, Milano.
Carbone, A. 1992. Integrazione produttiva sul territorio e formazione di sistemi agricoli locali (Productive integration in the territory and formation of local agricultural system), La Questione Agraria, n.46, pp.137-163.
Casini, L., Bernetti, J., Menghini, S. 2002. L'agricoltura e lo sviluppo rurale nei nuovi strumenti per il governo del territorio (Agriculture and Rural Development in the new instruments for the governance of the territory). Franco Angeli, Milano.
Cesaretti, G.P., Mariani, A.C., Sodano, V. (eds) 1994. Sistema agro-alimentare e mercati agricoli (Agro-food and agricultural markets). Il Mulino, Bologna.
Cesaretti, G.P., Bianco, M. (eds) 2006. Produzioni agroalimentari e territorio: la ricerca dei caratteri distintivi per l'impostazione dei piani di comunicazione (Agrifood production and territory: the search for distinctive characters for the setting of communication plans). Franco Angeli, Milano.
Cesaretti, G.P., Green, R., Hammoudi, H. 2006. Organizzazione dei sistemi territoriali a vocazione agroalimentare: dalla rete di prossimità geografica alla rete di prossimità virtuale (Organization of territorial systems with agrifood vocation: from the network of geographic proximity to the network of virtual proximity). in Rivista di Economia Agroalimentare – anno X n° 2. Franco Angeli, Milano.
Cesaretti, G.P., Regazzi, D. (eds) 2007. Leve strategiche per lo sviluppo dei sistemi territoriali a vocazione agro-alimentare (Strategic levers for the development of territorial systems with agrifood vocation). Atti del XIV Convegno annuale della Società Italiana di Economia Agro-Alimentare. Franco Angeli, Milano.
Cesaretti, G.P., Scarpato D. 2010 (eds). Politiche integrate per lo sviluppo competitivo sostenibile dell'agroalimentare (Integrated policies for sustainable competitive development of agrifood). FrancoAngeli, Milano.
Commissione Europea. 1996. The Information Society, Office for Official Publications of the European Communities.
Constanta, E. 2008. Modalities to Promote the Rural Development Research Networks – A French Experiece. Agricultural Economics and Rural Development Journal, Volume 5. Issue 1–2. pp. 60–67.
Coppola, A., De Muro, P., Fabiani, G., Favia, F., Martinelli, F. 1988. Un modello interpretativo per l'individuazione dei sistemi agricoli territoriali (An interpretive model for the identification of local agricultural systems). La Questione Agraria, n.30, pp.53-96.
Davico L. 2004, Sviluppo sostenibile. Le dimensioni sociali (Sustainable development. The social dimensions). Carocci, Roma.
Dematteis, G., Governa, F. 2005. Territorialità, sviluppo locale, sostenibilità: il modello SLoT (Territoriality, local development, sustainability: the SLOT model). FrancoAngeli, Milano.

Fabiani, G. (ed) 1991. Letture territoriali dello sviluppo agricolo (Regional readings of agricultural development). Franco Angeli, Milano.

Groote, P., Huigen, P..PP., Haartsen, T. 2000. Claiming rural identities, in Haartsen, T, Groote, P, Huigen, P.P.P (eds), Claiming rural identities. Dynamics, Contexts, Policies, Van Gorcum, Assen, pp. 1 – 7.

Haartsen, T., Strijker, D. 2010. Rural youth culture: Keten in the Netherlands. Journal of Rural Studies, n.26, pp. 163-172.

Iacoponi, L. 1997. Analisi economica della ruralità (Economic analysis of rural), in Agricoltura e Ruralità, I Georgofili. Quaderni, VII, pp.9-109.

Iacoponi, L., Brunori, G., Rovai, M. 1995. Endogenous Development and the Agri-Industrial District, Atti Seminario Assisi, CESAR. On the impact of Endogenous Development in Rural Areas, 25-27 ottobre, pp.25-48.

Morgan, S.L., Marsden, T., Miele, M., Morley, A. 2010. Agricultural multifunctionality and farmers' entrepreneurial skills: A study of Tuscan and Welsh farmers. Journal of Rural Studies, n.26, pp. 116-129.

Natário M.M., Neto P.A. 2009. The new rural paradigm and the public policies in France: rural excellence poles, Agricultural Economics and Rural Development. New Series, Year VI, no. 1, p. 125–144.

Nazzaro, C. 2008. Sviluppo rurale, multifunzionalità e diversificazione in agricoltura – Nuovi percorsi di creazione di valore per le aziende agricole delle aree interne del Mezzogiorno d'Italia (Rural development, multifunctionality and diversification in agriculture - New ways of value creation for the farms of inland areas of Southern Italy). FrancoAngeli, Milano.

OECD. 2001. Information and Communication Technologies and Rural Development. OECD Publishing. Paris.

OECD. 2006. The New Rural Paradigm. Policies and Governance. OECD Rural Policy Reviews. OECD Publishing. Paris.

Papini, R., Pavan, A., Zamagni, S (eds) 1997. Abitare la società globale. Per una globalizzazione sostenibile (Living in the global society. For a sustainable globalization). Edizioni Scientifiche Italiane.

Petit, P. 1998. L'économie de l'information: Les enseignements des théories économiques (The Information Economy: Lessons from economic theory). La Découverte, Paris.

Pittman, J., Wittrock, V., Kulshreshtha, S., Wheaton, E. 2010. Vulnerability to climate change in rural Saskatchewan: Case study of the Rural Municipality of Rudy no. 284. Journal of Rural Studies, XXX, pp. 1-12.

Sbordone, L. 2001. Città e territorio fra sostenibilità e globalizzazione (City and the territory between sustainability and globalization). FrancoAngeli, Milano.

Stiglitz, J.E. 1998. The private uses of public interests: incentives and institutions. Journal of Economic Perspectives, 12(2), pp. 3-22.

Storti, D., Zumpano, C. (eds) 2009. Le politiche comunitarie per lo sviluppo rurale – Il quadro degli interventi in Italia (EU policies for rural development - The framework of operations in Italy). INEA.

# Utilising Grounded Theory research techniques to catalyze knowledge discovery within a Living Lab environment

AAK. Buitendag, JS. Van Der Walt
*Tshwane University of Technology Private Bag X680 Pretoria*
E-mail: BuitendagAAK@tut.ac.za

## Introduction

The role-out and establishment of Living Labs (LL) in South Africa is gaining momentum. One applicable area is in the envisaged establishment of a LL to support emergent, commercial and subsistence farmers. This paper will present a proposed LL framework which incorporates the Grounded Theory research methodology in order to aid in the discovery of new knowledge based on patterns in questions, posed by the Agricultural Community of Practice.
The proposed LL framework relies on Agricultural extension officers placed in various centres as the key enablers of the Grounded Theory process in a collaborative environment.

## Emergent Farmers as Communities of Practice

Emergent farmers as strategic communities of practice are attracting a lot of attention in Southern Africa. The South African government is continuously and extensively involved in the role-out of projects to support the newly established farming communities. One of the projects is that of the creation of Thusong community centres into envisioned Living Labs for rural development. The South African government together with provincial government are funding the establishment of various organisations to support emergent farmers in South Africa. Agriculture with regard to rural development, including land reform, and food production and security has been identified by the National Department of Agriculture Forest and Fisheries as one of the five priority areas to address the major social and economic challenges facing the South African government. (DAFF 2010)
Emergent farmers, who are geographically dispersed and located in remote areas of South Africa, lacks the capacity and resources, to become commercial farmers and respond to the food security needs of the country. The main problems that they face are shown in Figure 1. The main causes of these problems are the transfer of knowledge between knowledge providers and agricultural knowledge users. In addition, the current feedback mechanisms are inadequate to provide the necessary understanding, by the governance role players, of various challenges faced by emerging farmers'.
We belief the best strategy to break the cycle of poor performance, by emergent farmers, is to develop comprehensive programs supporting collaboration and community organising services, to support the individual and community needs of all the role players in agriculture, based on systems thinking. Increasing the effectiveness of anti-poverty programs requires that those designing and implementing programs need not only to develop adequate theories of poverty to guide programs, but also to ensure that community development approaches are as comprehensive as possible (Bradshaw, 2007: 22).
Wang (2005:2) motivated a study on the use of Enterprise Resource Planning systems for agricultural support by explaining that: As is the case in most developing countries, South Africa is characterized by high levels of poverty and food shortage, especially in rural areas

where approximately 70% of poor people reside. The commercial agricultural sector, which utilizes significant agricultural resources, such as human, materials and land, is in the situation of under-using or underutilizing these resources. Employment opportunities in commercial agriculture are largely limited to unskilled workers and a large share of total employment in this sector is usually of a seasonal or temporary nature; Thus, affecting the potential income of the rural population.

Atkinson & Buscher (2005:7) pointed out that one of the four major dilemmas emergent farmers involved with farming on commonage land face is that of a lack of agricultural knowledge as well as the ability to market and sell their products in a globally low price driven arena with cheaper foreign imports.

Figure 1 highlights a summary of the issues and challenges faced by emergent farmers. The summary was obtained through a Grounded theory research process conducted on emergent farmers in South Africa

This research consciously seeks the benefits of building social collaboration models for emergent farmers in SA based on their ethnic, living area, financial and material resources, education and experience, and other limitations and strengths the community currently experience.

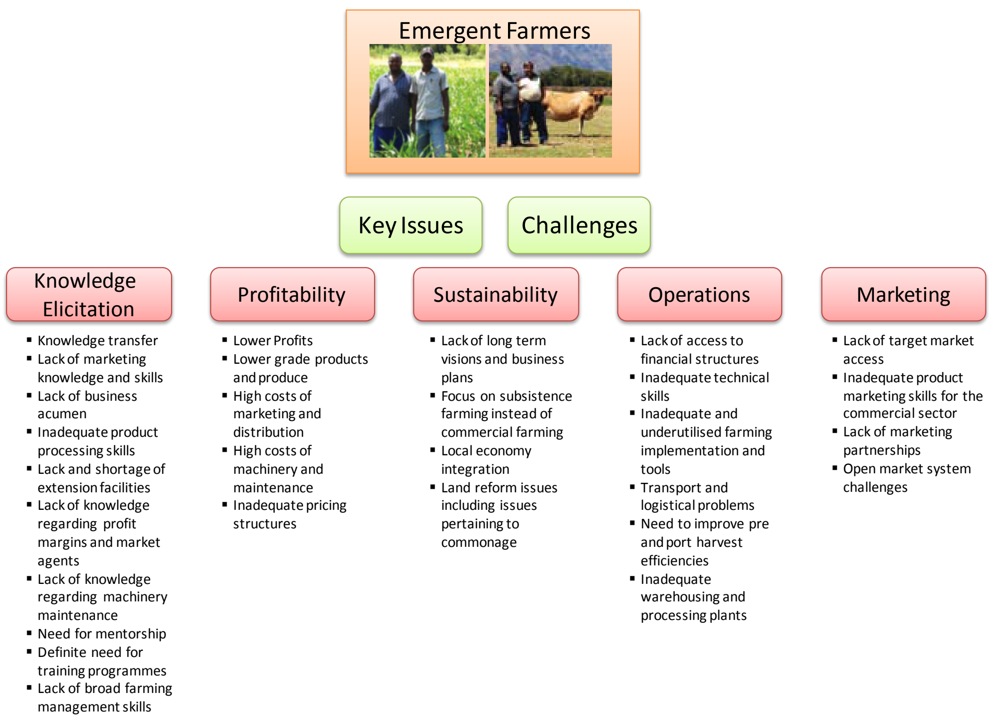



*Figure 1. ssues and challenges faced by emergent farmers.*

In a report on The Joint Initiative on Priority Skills Acquisition summoned by the Office of the Deputy President of South Africa, (JIPSA 2008:21) it was clearly indicated that, there is a shortfall of at least 5000 agricultural extension officers, in order to serve the agricultural sector of SA. The report also highlighted the consistent need for skills and knowledge development of these extension officers. The report correlates to the information represented. In order for a Community of Practice to function successfully, mechanisms need to be in place to ensure successful communication and transfer of knowledge within the group. Community members without the necessary skills need facilitation to aid them to understand disseminate and discuss new knowledge, information and ideas obtained or discovered by the group. Participation in various forms (including learning activities among community members) will ensure that the community performs the core functions of their intent and existence to its fullest capacity.

### Agricultural support structures in South Africa

Various structures and initiatives from both government and the private sector exist in order to support emergent farmers; the figure below highlights some of these structures and organisations and provides an overview of their primary objectives, functions and agricultural involvement.

| Organisation | Primary Objective | Agricultural Involvement |
|---|---|---|
| Thusong Service Centre | **Thusong Service Centres** are one-stop, integrated community development centres, with community participation and services relevant to people´s needs. They aim to empower the poor and disadvantaged through access to information, services and resources from government, non-governmental organisations parastatals, business, etc. enabling them to engage in government programmes for the improvement of their lives. | Provision of agricultural extension officers and services |
| WELCOME TO THE AGRICULTURAL R | The **Agricultural Research Council** are tasked into "conducting of research, development & technology transfer in order to:<br>1) Promote agriculture & industry,<br>2) Contribute to better quality of life,<br>3) Facilitate/ensure natural resource conservation" . | Undertaking programmes or rendering services such as information support and training to extension offices etc. |
| NERPO | The **National Emergent Red Meat Producers' Organisation (NERPO)** was established as a direct result of disparities that exist between the established and emerging agricultural sectors and such imbalances include the following:<br>1) Access to resources and therefore farming opportunities,<br>2) Institutional and infrastructural Support<br>3) Technical, Professional and Life Skills for emergent red meat producers in South Africa. | Undertaking programmes or rendering services such as information support and training to emergent red meat producers in South Africa |
| KAAP AGRI<br>LIMPOPO PROVINCIAL GOVERNMENT DEPARTMENT OF AGRICULTURE<br>agriculture FREE STATE PROVINCE | The **Provincial and Local government Agricultural extension offices**. Various provinces provides agricultural extension services to farmers in their provinces or municipal districts. All major agricultural provinces such as Gauteng, the Western Cape, Free state etc. and municipal structures provides these services to registered commercial and non commercial farmers. | To facilitate and furnish emergent farmers with advice, training, knowledge and information, for capacity building and sustainable local community development. |

*Figure 2. Various Agricultural support structures available in South Africa*

Extension offices are key enablers in the provision of knowledge and alternative methods to persuade clients which include emergent farmers to apply new or bettered practices out of their own free will. This includes actions of facilitation within the principles of help to self help, to prepare clients to adapt and be able to handle future problem situations. The Free State province local department of agriculture (2010) lists and motivates the following services rendered by their department by explaining that:

- Commercial farmers are in need of more advice on management issues and alternative products. Their level of expertise varies and therefore their needs are also divergent. For instance, they need farming management advice, technical advice and financial advice.

- Emerging farmers need training and advice on various issues, such as farming management, technical and financial issues.

Most agricultural departments and support structures provide extension and information services of a general nature to subsistence and landless (potential farmers) as well as urban and peri-urban farmers. The most common activities of extension officers include the dissemination of information and knowledge on farming management, technical issues as well as financial issues.

Currently no collaborative work system exists between the various support structures available to the respective farming communities. In most instances the provincial and municipal district support structures operate in silos, with little or no collaboration with regard to providing for the knowledge requirements of their intended CoP. In many cases extension officers aim to solve problems encountered by the CoP in isolation.. This results in the poor management of resources and sometimes lead to inadequate extension services, which further lead to low utilization of extension services by the CoP.

**Living Labs for knowledge support and innovation**

The implementation of Living Labs in South Africa is gaining momentum, currently 10 living labs exist in various phases of development and operation. LLiSA (Living Labs in Southern Africa 2010)

*Figure 3. LLiSA (Living Labs in Southern Africa 2010)*

Living labs are considered as the driving force for innovation being resource-sharing capabilities coupled with technology advancement demanding extensive infrastructure that is not easy to acquire, especially for small and medium enterprises and those who need high technology to achieve their goals. There is a reduction of technology and business risks, and the large companies have a large pool of ideas to help in their ventures (Lama & Origin, 2006:5 - 10).

Eskelinen (2010) defines a Living Lab with the following formula: Living Lab = User Driven + Open Innovation + Ecosystem

Ståhlbröst (2008) explains that the aim of Living Labs is to facilitate user involvement in innovation processes, suggesting an innovation system that is human-centric/driven, in contrast to technology-centric/driven. In these processes, users are invited to participate in the innovation and development process in their own context in authentic usage situations, facilitating the users to gain deep understanding of how a new product or service will function and correlate to their context based on their own lived experience. Following that line of thoughts, the risk of developing IT systems from a technology driven approach can be reduced in favour of the user-centered approach aiming to consider users' needs and desires in every development phase

Bergvall-Kåreborn *et al.* (2009) as cited by Lepik & Varblane (2010) defined Living Lab as an environment in which people and technology are gathered and in which the everyday context and user needs stimulate and challenge both research and development, since authorities and citizens take active part in the innovation process. The underlying idea is that people's ideas, experiences, and knowledge, as well as their daily needs of support from products, services, or applications, should be the starting point in innovation Lepik & Varblane (2010) highlighted the user-centred approach of the Living Lab concept, promoting innovation.

According to Boronowsky et al (2006) a living lab is more than a digital breeding area; it is a constructed setting of technology, shared by various researchers sharing the same drive, focused on finding the results and helping one another to achieve their goals. Researchers within living labs are restricted to monitoring from the inside what is going on. On the other hand, researchers are part of a living lab and have the capabilities to intervene in order to contribute to a better implementation of technological innovation in social practices, and deal with the unpredictable processes by reflecting on and consequently adjusting their own methodology.

Within South Africa there is a different approach to the modelling of a Living Lab in Southern Africa as opposed to those in Europe due to the differentiation core nature of intent of the Living Lab. According to LLiSA (2010) the main difference is that in Europe they apply different models of innovation as they are more focussed on industry involvement in a LL from the beginning whereas in Southern Africa the focus is currently more on rural community engagement and upliftment and to keep innovation within the community and not to allow the value to go to the industry but to the community as a whole. . According to Herselman, Marais and Pitse-Boshomane (2010) the current established LL within the network in SA each has distinctive uniqueness focussing on the: 1) capacity building of the communities involved, as well as the 2) enhancement of the innovation skills of the individual community

In prior research conducted by the researchers a Living Lab framework for use within the South African context was presented (Buitendag & van der Walt 2007). Figure 4 below depicts a subsequent adaptation to the framework which includes a depiction of the various research activities proposed for the LL, of which the Grounded Theory methodology is also included. The Grounded theory process which is part of inductive reasoning are particularly suited for application within a LL environment due to the fact that observations leads to the identification of patterns which results in the construction of tentative ideas and constructs which lead to a final theory. Extension officers often observe cases, and occurrences of problems and patterns, which lead to the formulation of ideas and constructs which result in theories, and actions towards the theories which are in essence a broad job description of any agricultural extension officer.

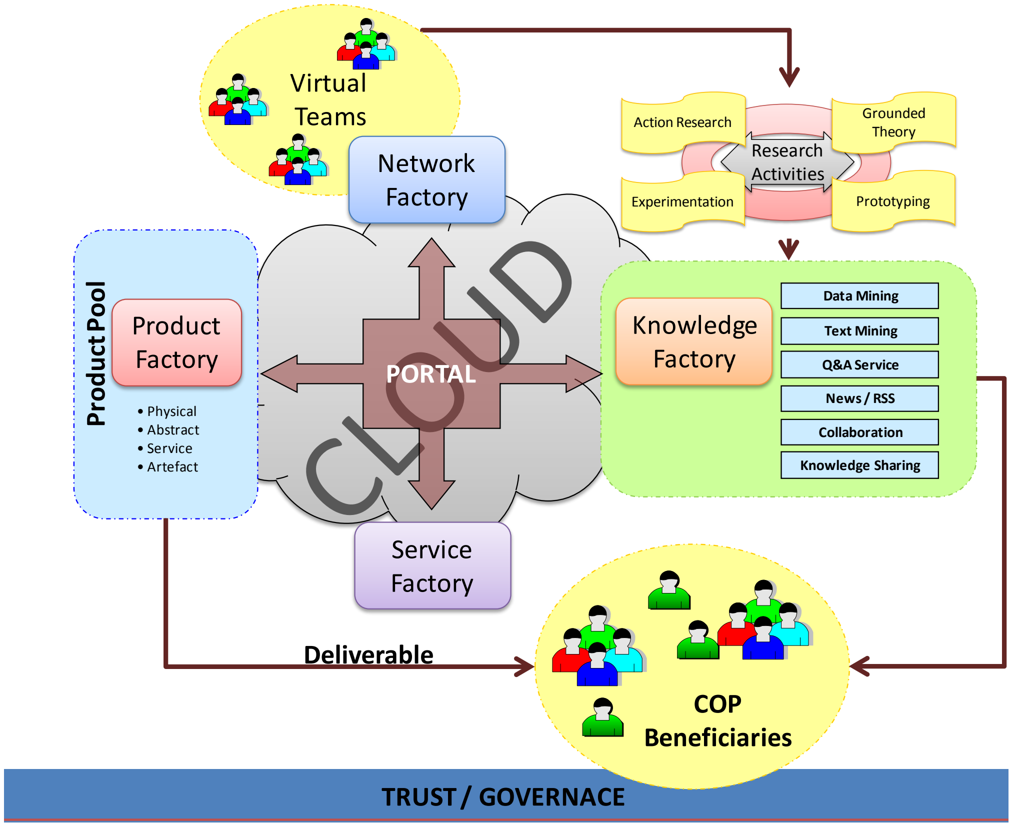



*Figure 4. Living Lab framework Utilising a Portal and the Cloud*

The proposed framework is based on various envisioned LL activities which include amongst others: collaboration, virtualization, learning, value chain optimisation and various applicable research activities, which are considered as one of the primary objectives of the LL.
Figure 4 as developed by the researchers depict a proposed Living Lab Framework as a set of various factories enabled by the utilisation of a Web portal.
In brief each of the factories will render and facilitate the following services and functions:
Product Factory (PF) - represents the processes and activities involved to deliver and create products in various forms. The objective of the PF is to aid in the development of an innovative "product": physical, abstract, a service, artefact. The "product" delivered will normally address and fulfil a specific CoP need or requirement. The researcher sees the creation of artefacts in the form of knowledge objects (KO) as one of the key deliverables of the factory.
Network factory- helps to find people that you the community need. The primary objective of the networking factory is to establish a platform for the engagement of various role-players within the LL. The network factory therefore is to network various communities of practice members and stakeholders into various virtual teams, which include agricultural extension officers from all agricultural disciplines and sectors.

Knowledge Factory - creates a dynamic set of Knowledge Objects implementing a Question and Answer Extrapolation tool (QAET). The QAET utilises the Grounded theory coding process in conjunction with semantic tools to discover patterns and common relationships between questions posed by the CoP. Newman (2010) of Gartner research described how Gartner's proposed pattern based strategy enables an organization to adapt and organization from a "sense and respond" culture to a more proactive "seek model and adapt" environment. Newman (2010) stresses that: The ability to quickly find patterns, understand patterns and change patterns is critical in today's competitive environments. Due to the embedded "sensing" nature of the grounded theory approach the researchers believe that the collaborative efforts of all LL stakeholders in the GT process would enable the CoP to simulate the "seek model and adapt environment" as proposed by Gartner. The primary purpose of the QAET is in the management of user requests, and the formation of knowledge objects. Agricultural extension officers are seen the primary enablers of the process, where requests and queries posed by various CoP members are catalogued, and semantically classified.

The subsequent actions and advice supplied by the extension officer is also stored and semantically classified for future reference. The researchers see the questions and queries posed as well as the subsequent results and advice given as a reusable Knowledge Object. We define a knowledge object (KO) as any artefact that could be implemented by a knowledge seeker in order to learn or expand the user's current knowledge regarding the specific search topic which include prior queries and results. KO's can take on a variety of formats, ranging from digital media to WEB 2.0 and WEB 3.0 mashed objects. All KO's utilized are stored and managed in a Knowledge Object Repository, which is in essence a semantic web cataloguing system.

Services Factory - produces all the web services needed in order for the living lab to function. It included resource planning, Information, business and communication, and knowledge support and analytical services. (van der Walt Et. Al 2009)

The LL in this study has been seen as an incubator and test bed for development of market cooperation's and other community of practice driven collaborations for agricultural community development, based on systems thinking. The tools and product factory, as well as the service factory utilize knowledge objects. The knowledge objects are categorized using standards which are stored and made available for future referencing and use. The knowledge objects are also part of the LL domain which could be available for use by other LL's and communities. In some instances the knowledge objects could be sold to generate income. Innovation in today's day and age it is of vital importance to any organisation. With the current economic climate in a crisis organisations' need to concentrate on increasing their efficiency to the maximum, with the least amount of effort and costs involved. As a possible solution to their problems, cloud computing allows the user to do the most of its computing in the internet cloud, thereby maximising the access to software and the subsequent resources available, that any organisation might need to perform their service.

The Cloud delivers computing and storage resources to its users or customers (Tuncay, 2009). It works as a service on demand, via the Internet, which is dynamically scalable from third party vendors. Cloud computing is a new business model wrapped around new technologies like virtualization, SaaS (Software as a Service) and broadband internet (Tuncay, 2009).

This represents a new paradigm for companies' IT operations where a pay-as-you-go business model will be used. Recent interests offered new applications and elastic scalability with higher computing power. These positive effects have shifted to outsourcing of not only equipment setup, but also the ongoing IT administration of the resources as well (http://www.gridforum.org/, 2009).

**Web 2.0 and Web 3.0 (Semantic Web) technologies as part of a Living Lab**

The knowledge factory incorporates current Web 2.0 and Web 3.0 technologies in order to classify, describe and share knowledge, knowledge objects. Prior research presented at a past WWW ZA conference 2009 c.f. Buitendag & van der Walt (2009) highlighted how Web 2.0 and Web 3.0 technologies could be implemented to support and enable some of the inherent processes of the various factories described. The figure below depicts some of the technologies available and suggested for use.

Many view the Semantic web and Web 3.0 as separate evolutionary successive layered (Cho 2008) entities based on small interpretational issues, which others dispute. For the purpose of this research the Semantic Web and Web 3.0 will be seen as the same entity. Both the semantic web as well as Web 3.0 focuses intelligent web based data applications with an intrinsic expansion to the embedded intelligence of the social web. Many scholars such as Zhou, Ding, Fin (2011) & Hender & Berners-Lee (2010) & Cho 2008 believes that the semantic web will not be a complete replacement of the social web (Web 2.0) but it is rather a natural evolutionary extension to it.

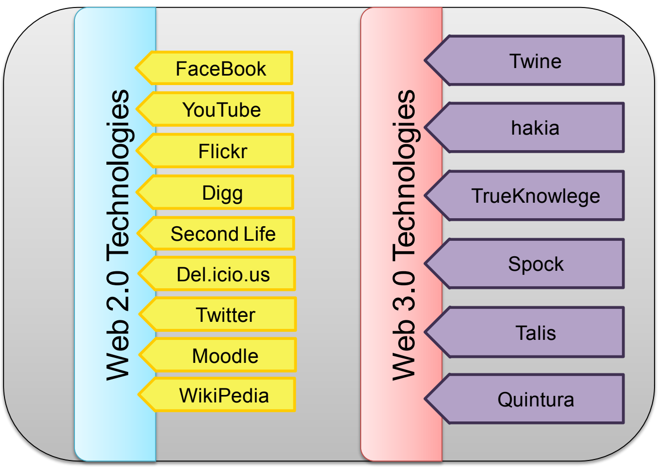



*Figure 5. Web 2.0 and Web 3.0 technologies for use as part of a LL*

McDonald (2010) explains that The Semantic Web provides a common framework that allows data, processes and a company's knowledge to be shared and reused across applications, enterprise(s) and community boundaries. It is an extension of the current Web that allows companies to find, share and combine information more easily, relying on machine-readable information. The Semantic Web is ideal for environments where product collaboration occurs, as well as collaborative knowledge base construction.
Research conducted by Janev & Vranes (2010) relating to the current Semantic web technologies application domains, which also highlights the domain application/coverage

percentages, it was found that the data integration, semantic search, semantic context discovery, and semantic annotation was the top four domains for semantic web applications. The table below constructed by the researchers presents the table as presented by Janev & Vranes (2010) but it also highlights where the domain apply within the LL framework with specific reference to the various factories as presented in Figure 4.

**The Grounded Theory research process as part Living Lab Framework**

The researchers believe that the Grounded Theory (GT) research methodology/approach incorporated as part of the Knowledge Factory would be crucial in the successful functioning of the living lab. Agricultural extension officers as the key enablers of the process often deals with enormous amounts of data both in a qualitative and well as quantitative nature, in order to assist farmers with occurring phenomena and subsequent problems and instances related to the Agricultural sector. Davidson (2002) defines and motivates the use of GT by explaining that: Grounded Theory is most accurately described as a research method in which the theory is developed from the data, rather than the other way around. That makes this is an inductive approach, meaning that it moves from the specific to the more general. The method of study is essentially based on three elements: concepts, categories and propositions, or what was originally called "hypotheses". However, concepts are the key elements of analysis since the theory is developed from the conceptualization of data, rather than the actual data.

Muller (2010) at IBM research motivates the use of GT by stating that: The GT process is good for explorative research, which lead to the disciplined development of new ideas, and developing a theory and structure in domains where there is no a priori guidance, while working with both qualitative and quantitative data.

Dick (2005) explained that: Grounded theory begins with a research situation. Within that situation, your task as researcher is to understand what is happening there, and how the players manage their roles, which evidently lead to new knowledge or support current theories on a certain topic. Charmaz (2008) highlights one major advantage of the GT process by stating that: "The process of GT fosters building explicit "What" and "How" questions into the GT process."

The researchers are of the opinion that: There is a definite need to better understand current the challenges and the environment emergent farmers face. With a better understanding of this challenges and environments, an improved design for the Product Factory of the LL could be presented. The researchers also trust that this approach best satisfies and describes the intended purpose of Agricultural extension officers.

The envisioned GT research process revolves around various activities Figure 6 as constructed by the researchers are based on the GT process as illustrated by Muller (2010), Dick (2005) and Pandit (1996). The figure highlights the overlapping nature of the various stages of the research as well as the iterative coding, memoing and sorting processes, and the finalisation in the "publishing" of the emergent core concept.
The functioning of the GT process is supported by the GT Research Portal interface which will allow various extension officers and other experts to collaborate. The Portal interface utilises the cloud for semantic knowledge discovery.

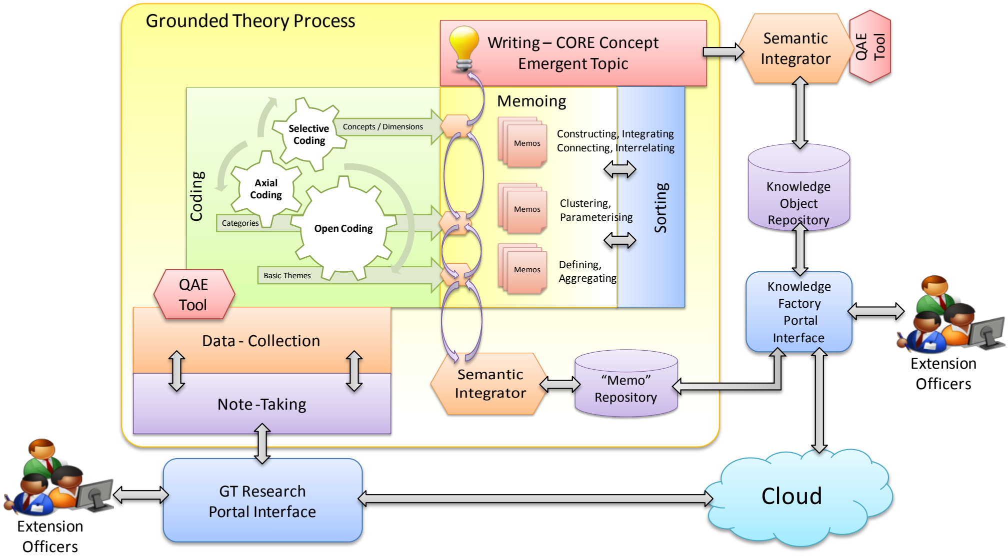



*Figure 6. Grounded Theory research process as part of the Living Lab framework*

The GT knowledge support framework as represented in figure 5, allows a knowledge seeker e.g. emergent farmer or to interact with extension officers, directly or via a portal interface. The knowledge seeker will post a request for knowledge in the form of a question, either directly or indirectly with the aid of an extension officer. A posted question or request are dissected and broken down into common sentence units, such as verbs, adjectives and nouns. The text mining service utilizes the sentence parts and performs an initial matching activity with prior questions that were stored within the Questions and Answer Repository. Similarities and AI matching methods are applied and matching result-sets from the Question and Answer repository are returned. The returned result-set and original question are then further analysed by utilizing Natural Language Processing tools and services. The Ontology wrapping service, implements service ontology for a Question and Answer Web service based on OWL-S.

Extension officers from various domains constantly take notes and collect data as part of their day to day operations. Capturing the notes and data via the portal interface will broaden the data and "memo" repository which in turn allows for the better coding of concepts, and phenomena observed. The QAE tool semantically classifies the content utilising services based on OWL-S, and RDF's. The data collection and note taking phases of the GT process has been described by Scott (2009) in stating that the essence of grounded theory phases resides in the mixed mode approach where both qualitative and quantitative data may be used in constant comparison, often guiding to researcher to better focus the research questions towards a more substantial theory. Data collection and note taking may me done in various forms which include amongst others: observations, interviews, questionnaires, recordings, photographs, group conversations, minutes of meetings, etc. Within the current South African context

Extension officers would be capable of performing such a function optimally, as they are considered to be the first link between farming communities and other entities.
Coding, Memoing and Sorting phases - The coding process is considered by many (Haig 1995, Scott 2009, Esteves et.al 2002) as the main phase of the GT process.

Straus and Corbin (1990) as cited by Omar et. al. (2010:276) explained that in building a grounded theory, the term coding is used to look between relationships between categories. "Analysis or coding of qualitative data represents the operations by which data are broken down conceptualized and put back together in new ways. It is the central process by which the theory is built from the data" Straus and Corbin (1990)
Omar *et. al.* further explained the coding process involves the processing of the data and the inductive analysis thereof using a three stage model of:

1. Open coding forms part of the analysis which involves the identification, naming and categorisation of the phenomena to be studied from the text or data presented. Esteves et. al. (2002) explains that open coding involves the labelling of concepts that represents discrete happenings and other instances within the phenomena
2. Axial coding is described by Hayhoe (2010) as the process where the connections between the concepts and the actors involved are connected. Esteves et. al. (2002) states that axial coding involves the procedures by which the data are put back together in new ways after the open coding process, this process gives the researcher to opportunity to make new connections between the current categories of the data collected. Borgatti (2010) also stressed that the process of axial coding involves relating codes (categories and properties) to each other using a combination of inductive and deductive thinking.
3. Selective coding is the process of selecting the core category, systematically relating it to other categories, validating the relationships and identifying the categories that need further refinement and development. (Esteves et. al. (2002))

The process of memoing is commonly referred to the process of generating notes and small memo's relating to the researcher's hypothesis or the ideas or insights gained about a category or property and their interrelated connections. Borgatti (2010) defines a memo as short documents the researcher constructs as the research progresses through the analysis of a corpus of data.

Dick (2005) highlights the fact that the action of memoing is a practice that continues in parallel with the data collection, note-taking and coding procedures. He further explains that the core proposition of GT is the fact that the researcher assumes that the theory is concealed in the data for the researcher to discover. On differentiating between coding and memoing Dick states that "coding makes visible some of the theories components, while memoing aims at adding relationships which links the categories of the theory to each other" Sorting is the process, of arranging the memos and finding the theoretical codes which best organises your substantive codes. The sorting process also involves the process of connecting and integrating the relevant literature to your theory through selective coding. (Scott, 2010 & Dick, 2005) Rhine (2009) explains that the sorting process does not refer to the sorting of data but rather the process of conceptual sorting of the memos into an outline to better highlight the emergent theory, which shows the relationships between the various concepts involved. The process often stimulates the creation of more memo's and sometimes necessitates the need to perform more data collection. The coding and memoing and sorting processes are improved by the utilisation of the OAE tool and semantic integrator.

The semantic extrapolation process generates tags which are compared with existing metadata, using semantic pattern clustering in the Semantic Knowledge repository, which matches existing classes, relations, axioms, functions and instances of prior searches and results. The Knowledge Object Repository (KOR) contains metadata descriptions of KO's applicable to the current living lab domain, and the Semantic Knowledge repository references semantic knowledge from external domains. Horrocks 2008, explained that the goal of semantic web is to allow the vast range of web-accessible information and services to be more effectively exploited by both humans and automated tools, which is facilitated by utilising, Resource Description Frameworks (RDF's) and Web Ontology Language (OWL). RDF and OWL have been developed as standard formats for the sharing and integration of data and knowledge, the latter in the form of rich conceptual schemas, called ontologies.

The semantic web utilizes ontologies as the primary enabling technology. Fensel & Siorpaes (2010) explained that ontologies are the backbone of the Semantic web which provides the knowledge that is required for semantic applications of all kinds. Ontologies allow the creation web content by using annotations which are machine-readable and machine-understandable.
The web service integrator incorporates web services bus incorporating WSDL and OWL's for the retrieval and discovery of possible data sources, which is not part of the current SKR from external web content, as well as external domain knowledge bases. The results retrieved from external sources, as part of the knowledge seeking process is then evaluated as part of the research processes by various knowledge officers, the subsequent new knowledge or discoveries are tagged, semantically described and stored as part of the KOR for future use.

In order to address the knowledge related issues, emergent farmers face within the South African context, the LL framework as presented relies on the collaborative creation and generation of Knowledge Objects to fully or partially full fill their knowledge requirements. As explained in the first section of this research KO's may take on various forms. Using the collaborative efforts of various agricultural extension officers existing web based documents in various formats could be tagged, and annotated with relation to specific cases utilising standard ontologies and metadata wrappers. These wrappers and annotations are managed by the QAET as part of the Knowledge Factory and stored in the KOR. In order to facilitate the representation of web papers in various forms Garcia-Crespo Et. Al. (2010) developed the POAP (Paper-Of-A-Paper) ontology. The POAP ontology represents the network of concepts and associated external resources derived from the tagging activity of related web documents. The POAP ontology is modelled on the Friend-Of-A-Friend FOAF ontology which facilitates interaction and interrelations based on people's connections and friends.

Other applicable ontologies and a short description thereof which could be utilized as part of the Knowledge Factory Tools are provided in the table below.

| Ontology name | Brief description & web reference | Knowledge Factory framework application area and description |
|---|---|---|
| Meaning- Of-A-Tag Ontology (MOAT) | MOAT is a MOAT, a lightweight Semantic Web framework that provides a collaborative way to let Web 2.0 content producers give meanings to their tags i.e. (URIs of Semantic Web resources ) in a machine-readable way. (Passant & Laublet 2008) & (http://moat-project.org/ontology) | This ontology will be particularly applicable to tag current knowledge objects as references and to provide descriptive classifications thereof. Other documents of a similar nature trying to represent the same core content are located at various locations. E.g. video clips of training material stored on YouTube etc. |
| Friend-Of-A-Friend Ontology (FOAF) | The FOAF ontology is used to express metadata about persons: their names, addresses, depictions of them. FOAF defines an open, decentralized technology for connecting social Web sites, and the people they describe and is heavily implemented in social websites such as Facebook and LinkedIn. (Möller, Predoiu, Bachlechner 2004) & (http://www.foaf-project.org/) | Friends and network colleagues could also be considered as tactical sources of knowledge. Linking experts with the same expertise will greatly improve on the turnaround time of knowledge queries posted. |
| Social-Semantic-Cloud-of-Tags Ontology(SCOT) | SCOT (Social Semantic Cloud of Tags) is an ontology for sharing and reusing tag data and representing social relations among individuals. It aims to describe the structure and the semantics of tagged data as well as offer interoperability of data among different sources. (Kim Et.al. 2010:527) & (http://www.semantic-web.at/1.36.resource.245.scot-let-x27-s-share-tags.htm) | Tagged descriptions of data will considered as a key deliverable as the Memoing process as part of the GT methodology rely on the creation of conjunctures between notes to aid in the discovery of patterns between the observations and the data. |
| Semantically Interlinked-Online-Communities Ontology (SIOC) | SIOC is an ontology for describing discussion forums and posts on topic threads in online community sites, it aims to aims to enable the integration of online community information. SIOC provides a Semantic Web ontology for representing rich data from the Social Web such as forums, blogs, wiki's in RDF. (Bojars Et. al. 2008) & http://sioc-project.org/ | Blogs and wiki's will form a integral part in the knowledge repository of a living lab. The KOR as described in the framework collates to the integration of online community information. |
| Simple-Knowledge-Organisation-System Ontology (SKOS) | SKOS is an area of work developing specifications and standards to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading systems and taxonomies within the framework of the Semantic Web. (http://www.w3.org/2004/02/skos/intro) | In order to manage the data and tags and references stored as part of the KOR the SKOS will provide classification services relating to the various types of knowledge objects referenced |

*Table 1. Applicable ontologies as part of the GT knowledge support framework*

The factory integrator allows for the integration of the various LL factories and manages the flow of data and processes.

The emergent theory and subsequent knowledge generated stored in a Knowledge object repository for future reference via the Knowledge Portal Interface and Cloud. Using services based on OWL-S (Language for describing web services) will create and establish a dynamic problem solving environment using various resources.

## Conclusion

A Living Lab can support knowledge transfer (important know how) and improve farming practices - rooted in innovation, education, research and technology use. Innovation can improve food security through wealth creation and business growth. A LL is supported by a set of tools to create an innovation factory. The LL in this study has being seen as an incubator and test bed for development of market cooperation's and other community of practice driven collaborations for agricultural community development, based on collaborative thinking.

We believe that knowledge support services enabled by knowledge discovery will become a key deliverable in the development of any information driven portal. Extension Officers working in a collaborative environment utilising various tools such as described as part of the Living Lab framework would surely contribute to the improving extension services to emergent and commercial farmers. The grounded theory approach is well suited for this purpose. From a South African perspective such services can play a critical role in limiting and overcoming obstacles such as information poverty and knowledge deprivation. The objectives, use and advantages of knowledge support services are not limited only to the agricultural domain but apply to any knowledge/information driven environment. We believe that semantics based web service technologies comply with satisfying the requirements and improving the interoperability of distributed service component integration.

## References


Atkinson, D., Buscher, B., 2005. Emergent farmers- local knowledge and the implication for land reform: A profile of commonage users in Philippolis, Free State [Online] Available at: http://www.aridareas.co.za/pdf%20documents/philippolis.pdf [Accessed 2010/02/11]

Bojars, U., Cyganiak, R., Passant, A., Breslin, J., 2008. Weaving SIOC into the Web of Linked Data. [Online] Available at: http://events.linkeddata.org/ldow2008/papers/01-bojars-passant-weaving-sioc.pdf [Accessed: 2011/04/19

Borgatti, S., 2010. Introduction to Grounded Theory. [Online]. Available at: http://www.analytitech.com/mb870/introtoGT.htm [Accessed: 2010/07/21].

Boronowsky, M., Herzog, O., Knackfub, B. & Lawo, M. 2006. Wearable computing – an Approach for Living Labs. [Online]. Available at: www.wearitatwork.com. [Accessed: 2010/09/10].

Bradshaw, T. K. (2007). Theories of poverty and Anti-Poverty Programs in Community Development. Journal of the Community Development Society, 18, no 1,Springer 2007.

Buitendag A.A.K., Van Der Walt J.S., 2007. A Question and Answer Knowledge sharing system to support emergent farmers in Southern Africa. EFITA/WCCA 2007 Joint Conference. Glassgow, Scotland. Electronic CD proceedings.

Buitendag A.A.K., van der Walt J.S. (2011) Presenting a framework for the utilisation of the Grounded Theory research methodology to aid in knowledge discovery as part of an Agricultural Living Lab. EFITA/WCCA 2011 Joint Conference. Prague, Czech Republic. 11-14th July

Charmaz, K. 2008. What is grounded Theory. [Online]. Available at: http://eprints.ncrm.ac.uk/208/1/What_is_Grounded_Theory.ppt.4-24-08.SS.ppt. [Accessed: 2009/05/21].

DAFF SA. 2010. South African National Department of Agriculture, Forestry and Fisheries. Strategic plan for the DAFF. [Online] Available at: http://www.nda.agric.za/doaDev/topMenu/StartPlan2010/DAFF_StratPlan2010.pdf. [Accessed 2010/10/02]

Davidson, A.L. 2002. Grounded Theory – Defined [Online] Available at: http://www.essortment.com/all/groundedtheory_rmnf.htm [Accessed 2009/01/09]

Dick, B. 2005. Grounded theory: a thumbnail sketch. [Online] Available at: http://www.scu.edu.au/schools/gcm/ar/arp/grounded.html [Accessed 2009/01/09]

Eskelinen, J., 2010. HELSINKI Living Lab (Volume 2) Convergence of Users, Developers, Utilizers and Enablers. [Online] Available at: http://www.forumvirium.fi/en/project-areas/innovation-communities/helsinki-living-lab [Accessed 2011/02/10]

Esteves, J., Ramos, I., Carvalho, J. ,2002. Use of Grounded theory in information systems Area: An exploratory Analysis. In: Proceedings of European Conference on Research Methodologies on Business Management. ERCM 2002, Reading: UK 129-136

Fensel, D., Siorpaes, K., 2010. Semantic Web: Ontologies and the Semantic Web. [Online] Available at: http://www.sti-innsbruck.at/fileadmin/documents/semweb10-11/SW-Lecture10-handouts.pdf [Accessed 2011/04/19]

Garcia-Crespo,A., Gomez-Berbis,J,M., Colomo-Palacios,R., Garcia-Sanchez, F, 2010. Digital libraries and Web 3.0. The CallimachusDL approach, Computers in Human Behavior, In. Social and Humanistic Computing for the Knowledge Society 27(4): 1424-1430,

Haig, B.D., 1995. [Online] Available at: Grounded Theory as Scientific Method. http://jan.ucc.nau.edu/~pms/cj355/readings/Haig%20Grounded%20Theory%20as%20Scientific%20Method.pdf [Accessed 2009/09/15]

Hayhoe, S., 2010. Overview of Grounded Theory. [Online] Available at: blindnessandarts.academia.edu/.../BrusselsGroundedTheory.ppt [Accessed 2010/09/15]

Hendler, J.,Berners-Lee, T., 2010. From the Semantic Web to social machines: A research challenge for AI on the World Wide Web, Artificial Intelligence, 174(2):156-161

Herselman, M.E., Marais, M.A., Pitse-Boshomane, M.M. 2010. Applying living lab methodology to enhance skills in innovation. Proceedings of the eSkills Summit 2010. Cape Town, 26-28 July 2010, pp 7

Horrocks, I., 2008. Ontologies and the Semantic Web. Communications of the ACM, 51(12):58-67.

Janev, V., Vraneš, S. 2010. Applicability assessment of Semantic Web technologies. Information Processing and Management. [Online] Available at: http:/doi:10.1016/j.ipm.2010.11.002 [Accessed 2011/04/20]

JIPSA 2008. The Joint Initiative on Priority Skills Acquisition – Report. [Online] Available at: www.info.gov.za/view/DownloadFileAction?id=98945 [Accessed 2010/10/21].

Lama, N., origin, A. 2006. Innovation ECOSYSTEMS: Services Engineering & Living Labs a dream to drive innovation? [Online]. Available at: www.atosorigin.com [Accessed 2010/10/02].

LivingLabs Europe.com – 2009. Living Labs User group for open innovation [Online] Available at: ec.europa.eu/info.../livinglabs/docs/brochure_jan09_en.pdfasp [Accessed 2009/01/09].

Lepik, A., Varblane, U., 2010. How to speak the same language with European Innovation-policy in terms of living labs? In: Discussions on Estonian economic policy [Online] Available at: http://www.mattimar.ee/publikatsioonid/majanduspoliitika/2010/2010.pdf# [Accessed 2009/02/02].

LLiSA 2010. Living Labs South Africa. [Online] Available at: http://llisa.meraka.org.za/index.php/Living_Labs_in_Southern_Africa [Online]. Accessed [2010/03/12]

McDonald, K.M., 2010. The Semantic Web: Out of School and Into the Business World. [Online] Available at: http://www.armg-usa.com/SemanticWeb.html Accessed [2010/04/24]

Möller, K., Predoiu, L., Bachlechner, D., 2004. Portal Ontology: Semantic Web Portal Project. DERI Research Report. [Online] Available at: http://sw-portal.deri.at/papers/deliverables/D1_v1.2PortalOntology.pdf [Accessed 2010/10/18].

Muller, M. 2010. Grounded Theory Methods. Presentation at HCIC 2010, Winter Park, CO, USA, February 2010. [Online] Available at: www.slideshare.net/traincroft/grounded-theory-method-hcic-2010-muller [Accessed 2010/09/15]

NERPO, 2011. National Emergent Red Meat Producers Organisation. [Online]. Available at: http://www.nerpo.org.za/ Accessed: 2010/10/20

Newman, D., 2010. The Adaptive Organization: How Boyd's Decision Cycle and Pattern based strategy Drive Rapid Change. Gartner – Research [ID G0020671]

Omar, A.H.H., Dayang, H.T.A.H.H., Norma, A., Md, R.I. 2010. Grounded Theory: A Short Cut to Highlight a Researchers' Intellectuality. In. Journal of Social Sciences 6 (2)

Pandit, H.R., 1996. The Creation of Theory: A Recent Application of the Grounded Theory Method [Online] Available at: http://www.nova.edu/ssss/QR/QR2-4/pandit.html [Accessed 2009/01/15]

Passant, A., Laublet. P., 2008. Meaning of a Tag: A collaborative approach to bridge the gap between tagging and Linked Data. In Proceedings of Linked Data on the Web. [Online]. Available at: http://events.linkeddata.org/ldow2008/papers/22-passant-laublet-meaning-of-a-tag.pdf [Accessed 2010/09/10]

Rhine, J. 2009. What is Grounded Theory – Grounded Theory Institute. [Online:] Available at: http://www.groundedtheory.com/what-is-gt.aspx [Accessed 2010/10/21]

Ståhlbröst, A. 2008. Forming future IT: The living lab way of user involvement. Unpublished PhD Thesis, Luleå University of Technology Department of Business Administration and Social Sciences, Division of Informatics, Finland.

Thompson, W.J.J., Buitendag A.A.K., van der Walt J.S. (2011)

The Agile Farmer: "A beef farmer case study, in Southern Africa" EFITA/WCCA 2011 Joint Conference. Prague, Czech Republic. 11-14th July

Tuncay, E. 2009. Effective use of cloud computing in educational institutions aYasar University, Department of Computer Engineering, Selcuk Yasar Kampusu, Agacl Yol, No:35-37, Bornova 35500, Izmir, Turkey Received October 8, 2009; revised December 17, 2009; accepted January 5, 2010

Van Der Walt, J.S., Buitendag, A.A.K., Jansen Van Vuuren, J.C., Zaaiman, J.J. 2009 Living Lab as a Collaborative Innovation Environment. In. Journal of Issues in Informing Science and Information Technology (IISIT) Volume 6, pages 421-436. USA

van der Walt J.S., Buitendag A.A.K., Thompson, W.J.J. (2011)

A systems thinking model for emergent farmers performance improvement in South Africa. EFITA/WCCA 2011 Joint Conference. Prague, Czech Republic. 11-14th July

Wang, K., 2005.. Architectural Framework For Building A Virtual Knowledge Support System M.Tech IT. Dissertation, Tshwane University of Technology.

# Exchanges of Environmental Information: the next challenge?

[1]G. Waksman, S. Masselin-Silvin, [2]F. Brun, [3]J. Oberti, [4]J.F. Samie, [5]G. Cheruy, B. Prépin, [6]A. Wissocq, [7]J.C. Champomier, [8]B. Dreux
*[1]ACTA Informatique, [2]Acta, [3]Institut de l'Élevage, [4]Ctifl,[5]AEE, [6]Arvalis-Institut du Végétal, [7]Cemagref, [8]Anela*
E-mail: Guy.Waksman@acta.asso.fr

**Introduction: The objectives of our research project over 2 years (2009-2010)**

The research project GI-E-EA (Gestion des informations "environnement" des exploitations agricoles) is to demonstrate the value of "environment" information and the electronic exchange of "environment" information between farmers and their partners, advisory organizations, economic actors and in some cases governmental administration. We estimate that the volume of exchange of these "environment" data could overtake in the coming years the volume of exchanges of economic data (invoices, bank statements) and of data related to animal traceability (identification and tracking of cattle mainly).
The GI-E-EA project is an extension of works performed in France by the Assemblée permanente des chambres d'agriculture (APCA) on an inventory of concepts used in the technical management of farms (GIEA project) and Agro EDI Europe on exchange of electronic information between farmers' Farm Management Information System (FMIS) and their partners, technical partners, commercial or institutional (Resagri project).
Specifically, with the GI-E-EA project, we are interested in "environment" information with our project partners:

- Arvalis – Institut du Végétal (large crops)
- Institut de l'Élevage (cattle, sheep, goats ...)
- Centre technique des fruits et légumes - Ctifl (fruits and vegetables)
- Acta, association of agricultural technical institutes, and specialist in crop protection
- Cemagref (research team interested in the repositories and their consistency)
- Agro EDI Europe (AEE), EDI specialist in Agriculture, at the initiative of the UN-CEFACT working group (TBG18), and creator of the e-Daplos standard that describes parcels and cultural operations
- Anela, association of French publishers of agricultural software, Isagri being the most important

The objectives of this project GI-E-EA were:

- to perform an inventory of most relevant and shared "environment" information, to enable the environmental "traceability" of cultural operations as well as operations carried out in livestock holdings
- to create the corresponding data models data in three areas: inputs, water and energy
- to develop prototypes showing the interest of "environment" data exchange.

**Working methods**

The first working method was to collect and review current regulations on cross compliance and to describe information required by these regulations. In large crops productions, this approach was not really successful, but in animal productions, the Institut de l'Élevage has achieved significant results, with a comprehensive table of information used in the livestock regulations (e.g. identification and health of livestock, management of cattle manures, etc.).
The survey of environmental regulations applicable in cattle productions provides farmers and technicians with a simple and comprehensive table of environmental regulations and farmers' obligations that apply on 1 January 2010.
The second working method was to gather information on tools dedicated to traceability, decision support, simulation and diagnosis, etc. and to identify "environment" information used by these tools (See GI-E-EA form example bellow).
Within the framework of different projects and domains, ACTA performed in 2009 a survey and an analysis of decision support systems and models used by farmers, advisors and searchers, and ACTA is proposing to distinguish two major types of projects involving models for agricultural extension (See modelia project : wwx.modelia.org).

<u>To help production management (i.e. to pilot farm operations)</u>

These tools provide farmers (i.e. the decision makers) with personalized information and advices related to a specific context.
The efficient dissemination of results of this kind of tools may enable to reach a wide audience of farmers and advisors e.g. through the Internet.
Either the farmers perform by himself simulations when directly using these tools, input his own data and getting specific results, or use risk maps and advice bulletin produced by advisory organisations.
Examples are tools for predicting risk of diseases and pests, irrigation management, forecasting yield and harvest date (to help planning of operations), environmental diagnostics, computing of techno-economic results. In crop productions, we can note the importance of tools for crop protection, with over 120 different models listed in France.

<u>To evaluate production systems</u>

These models are used by searchers to prepare studies and establish benchmarks. Their dissemination is limited since the targeted users are searchers (mainly from Inra) and specialists collaborating with Inra teams in our technical Institutes and sometimes advisory organizations.
The examples of the se tools are many: a study of genotype environment interactions comparison of production systems (integrated pest management versus classical management), environmental impact assessment of different farming systems, etc. These models are often vectors facilitating dialogue between partners involved in common research topics.

Were particularly studied:

- A large set of decision support tools for inputs (fertilisers, chemicals, water) computation, greenhouse gases production and energy balance, etc. by Arvalis - Institut du végétal,
- All the tools agro-meteorological risk assessment in plant fruits and vegetables productions by Ctifl,
- the Dexel software programme created by the I-Cone ICT company for Institut de l'Élevage, that is widely used throughout France and performs a diagnostic on manure production at farm level, by Institut de l'Élevage
- the Dia'Terre software programme under development by Agence de l'Environnement et de la Maîtrise de l'Énergie (Ademe), that perfoms an energy balance and production of greenhouse gases diagnosis at farm level,
- the Diaphyt software that performs an evaluation practices of farmers in crop protection by Acta

A third method has been implemented by Agro EDI Europe that has successfully brought together stakeholders (software publishers, R & D organizations, cooperatives, chemical companies and administration services) to work in groups to:

- improve the "agricultural plot" message named e-Daplos that describes the agricultural plots and the cultural operations performed on these plots,
- develop the "agro-obs" message that compiles biological observations in the fields.

From this experience, we believe that this working method which is also adopted by the ISOagriNET in Germany as well as by the TBG-18 of the UN / CEFACT is very successful and could be expanded successfully at European level e.g. through the AgriXchange platform.

********

## GI-E-EA form example

| | | | |
|---|---|---|---|
| **Numéro :** | 2 | **Nom du message :** | **I-PHY légumes** |

**Population agricole concernée :** Producteurs, techniciens de développement

**Production (s) agricole (s) concernée (s)** Maraîchage

**Type de message:** Diagnostic environnemental

(OAD Phyto ; OAD Irrigation ; OAD Fertilisation ; OAD Énergie ; Diagnostic environnemental ; Diagnostic Énergie ; BdD Surveillance du territoire, etc.)

| **Contenu du message** |
|---|
| Informations de la parcelle et des pratiques phytosanitaires jouant un rôle dans l'impact environnemental des pratiques phytosanitaires sur l'environnement. |

| | | | |
|---|---|---|---|
| **De :** | Agriculteur / technicien | **Vers :** | Base de données |
| **Fréquence Saisonnalité :** | 1 fois / campagne | **Volumétrie :** | 15 envois / an |
| **Événement déclencheur :** | Saisie en ligne | | |

**Description du message :**

| Objet | Champs / données | Type : numérique, texte | Unité | Noms de tables de référence (à compléter – tableau ci-après) |
|---|---|---|---|---|
| Parcelle | Nom | AN..100 | | |
| Parcelle | Sol drainé | B | | |
| Parcelle | pH sol | N..50 | | |
| Parcelle | Taux MO | N..50 | % | |
| Parcelle | Distance entre bord cultivé & eau surface | N..50 | m | |
| Parcelle | Sol filtrant | B | | |
| Parcelle | Profondeur sol | N..50 | cm | |

| Objet | Champs / données | Type : numérique, texte | Unité | Noms de tables de référence (à compléter – tableau ci-après) |
|---|---|---|---|---|
| Parcelle | Texture sol | AN..50<br>L | | Table des sols : Sableux (S) ; Sablo-argileux (SA) ; Sablo-limoneux (SL) ; Limoneux (L) ; Limono-sableux (LS) ; Limono-argileux (LA) ; Argileux (A) ; Argilo-sableux (AS) ; Argilo-limoneux (AL) ; Argile lourde (AA) |
| Parcelle | Sol battant | B | | |
| Parcelle | Sol hydromorphe | B | | |
| Parcelle | Pente | N..50 | % | |
| Occupation culturale | Espèce | N..10 | | Espèce |
| Occupation culturale | Date semis / plantation | D | | |
| Occupation culturale | Aménagement parcellaire bas pente (largeur bande enherbée, haie) | N..50 | m | |
| Traitement phytosanitaire | Produit phytosanitaire | N..19 | | Intrant |
| Traitement phytosanitaire | Dose | N..7 | | |
| Traitement phytosanitaire | Unité | N..19 | | Unité |
| Traitement phytosanitaire | Date | D | | |
| Traitement phytosanitaire | Position application | AN..250<br>L<br>Enfoui ; Sur sol nu ; Sur végétation | | |
| Traitement phytosanitaire | Type traitement | AN..250<br>L | | en plein ; sur le rang |
| Traitement phytosanitaire | Gestion résidus culture | AN..250<br>L | | pas de résidus ; résidus enfouis ; résidus en place |
| Traitement phytosanitaire | Type buse | AN..250<br>L | | Fente classique ; Basse pression ; Pastille de calibrage ; Miroir ; Injection d'air ; Pas de buse |

**Description de la table de référence :**

| Objet | Champs / données | Type : numérique, texte | Unité | Noms de tables de référence (à compléter en annexe) |
|---|---|---|---|---|
| Espèce | Identifiant | N..10 | | |
| Espèce | Nom | AN..50 | | |

| Objet | Champs / données | Type : numérique, texte | Unité | Noms de tables de référence (à compléter en annexe) |
|---|---|---|---|---|
| Intrant | Identifiant | N..19 | | |
| Intrant | Nom | AN..100 | | |

| Objet | Champs / données | Type : numérique, texte | Unité | Noms de tables de référence (à compléter en annexe) |
|---|---|---|---|---|
| Unité | Identifiant | N..19 | | |
| Unité | Nom | AN..50 | | |
| Unité | Abréviation | AN..50 | | |

"Environment" information: a definition of environment information already present (i.e. input) or not in the FMIS

Our works allowed us to define the "environment" information as all information describing how the environment impacts the agricultural productions and / or describing how farming operations impacts the environment.

This information forms a consistent whole and are linked together: if one can predict the development or non-development of a plant disease from weather data and biological observations, it becomes possible to spread less chemicals and thus reduce the impact of treatment on surface water or groundwater quality.

We also observed the reality of what we called "agri-environmental traceability" as consisting of all data (plant and animal) to evaluate and / or pilot production systems taking into account their environmental impacts.

When we examined the information necessary to run models for predicting disease risk or to run software computing farm energy balance, agricultural software publishers told us that this information is rarely captured by operators when the FMIS software allows, and, even worse, the FMIS software available on the market does not yet take into account all this information, particularly those describing agricultural equipments.

In the example of GI-E-EA form given above, we see that the type of spray nozzle is included in the I-Phy vegetables model, and we know that such information is not managed in the FMIS.

Another example: software programmes enabling traceability of crop productions, at least in France, do rarely take irrigation into account. But since the new e-Daplos takes into account the amount of water distributed, then one can imagine that:

- publishers will complete their field management software, used to guarantee the traceability of field operations, insert the description of irrigation operations, and expand the export of "irrigation" e-Daplos.
- it will be possible to route the "irrigation" e-Daplos file to a software assessing water stress and computing irrigation water requirements such as Irribet for sugar beets (which is operating as web service), developed by Itb, the Institute in charge of sugar beets.

**3 examples of exchange of "environment" information between FMIS and DSS**

A. Example of exchange of "environment" information between FMIS and DSS in fruits and vegetables

In some cases, information from FMIS software and exported as Daplos or better e-Daplos may providing information necessary to operate a decision support system, for example, assessing epidemiological risks of certain diseases or pests in fruits and vegetables.
Ctifl (Institute of fruits and vegetables) has built two prototypes that show the interest of the exchange of environmental data.
The producer of leeks extracts form his FMIS the information necessary to the model "leek Trips," which are then transmitted trough the Internet to the Ctifl server as a daplos file. The server processes this information and sends "its" diagnosis. To use the "leek rot" model, the producer must provide complementary field observations in the form of a message agro-obs.
If we imagine that these models are used once a week with this organization that allows only having to input a minimum of information, we see that the amount of information exchanged will quickly become important. In both examples, the messaging exchanges generate more traffic than would traceability which can be limited to one message per year at the end of the season.

B. Examples of exchanges of "environment" information between FMIS and DSS in large crops: observation of cultural practices and health observatory

For several years, Arvalis-Institut du végétal offers a verification service that enables producers to check whether they comply with the quality specifications required by Arvalis / IRTAC charts.
E.g. cereals co-operatives transmit to the Arvalis service files describing plots and cultural practices (format Daplos created within Europe Agro EDI), and in return receive a diagnosis at filed level about 70% of the control points of the charts.
The improvements with e-Daplos by AEE including introduction of new information on irrigation, improved soil types repository, facilitated the development of an observatory of cultural practices within Arvalis-Institut du Végétal. This observatory is a database that will be

interfaced with many decision support systems already developed by Arvalis, and primarily with those focused on water quality at farm level and watershed level.
So here we find again the GI-E-EA project issue of the interface between the registration tools for practice in order to ensure traceability, and tools for diagnosing these practices in terms of impact on environment.
Arvalis-Institut du Végétal associated with other agricultural technical institutes (Cetiom Itb, Itl) also developed Vigicultures, a management tool for field observations on the presence of diseases and pests in largecrops.
The contents of the Vigicultures database will be partly transferred into a database managed by the Administration thanks to the "agro-obs" message created by AEE in collaboration with all stakeholders.
Again it will be many exchanges of data "environment."

<u>C. Example of exchanges of "environment" information between FMIS and DSS in cattle breeding</u>

Dexel is a software programme that has been widely used to assess the quantities of effluents French farms and to build a diagnosis. The first question asked by Institut de l'Élevage has been whether Dexel could be "fed" from information already present in the FMIS breeders. But the information required by Dexel are too many and specific. So we abandoned this idea. Dexel is interfaced with "Clé de sol", a fertilisation software management that enables to build plans for spreading manure. This interface is still "closed", that is to say, it works only with this single software. The current work focuses on creating a message gathering information stored in the Dexel software that are necessary to run any fertilisation software programme.
This message could then be proposed for standardization to AEE, contained, for all animal species, some or all of the following information:

- Types of animals;
- Types of barns;
- Average and authorized numbers of animals;
- Kg of N, P and K produced per animal per year;
- Duration of presence on the farm, barns and meadows;
- Destinations of the nitrogen produced in buildings with details of each storage structure and characterization of organic material stored;
- Nb of kg of nutrient products, and consideration of treatment options;
- Consistency between the quantities produced, the volumes of available organic matter and contents;
- Balance of the quantities available, taking account imports - exports to third parties or industry.

**The identified needs of actors (ICT companies, developers, searchers)**

As indicated above, in addition to Dexel, we studied two software programmes that are focused and complex:

- Dia'terre Software of "Agence de l'Environnement et de la Maîtrise de l'Énergie" that allows to compute an energy balance and to evaluate the production of greenhouse gases at farm level,
- Diaphyt Software of Acta that enables to evaluate practices of farmers in crop protection.

The objectives of these targeted tools are: for Diaphyt, to improve security and especially the practices of users when performing pesticide treatments, and for Dia'terre, to identify opportunities for energy savings and reduce greenhouse gas emissions.
We found that among the huge quantity of information needed to run these tools, very few already are in the FMIS most commonly used. Users therefore have or will enter - after being collected - a large amount of data.
Two approaches may be envisaged:

- trying to raise awareness of designers of such programs about FMIS, so they use as much as possible data that may have already been input in the FMIS
- and inversely, to share with developers and distributors of FMIS, the information needs of software such as Diaphyt or Dia'terre.

We observed that very often are missing in the FMIS descriptions of materials (for example: spraying equipment) or other data describing the context of the operation... In some cases, these data are already present, but not based on the reference tables. We hope that co-operation will rapidly develop since software tools like Dia'terre or Diaphyt, which are very demanding in data entry, will never be really used by many farmers unless they fit with specific regulations, what had been the case of Dexel.

A. The need for reference tables accepted by all stakeholders

In the course of our work has appeared again the need to establish reference tables that are accepted by all actors, or in other words that meet the needs of all stakeholders. In this regard, AEE and its partners have done a great job to prepare lists of crops, insects, diseases, weeds, as well as to compile lists of soil types.
But challenges remain. For example, the nomenclature of manure types, with all effluent mixtures that can be observed in farms has been the subject of special study.
The comparison of repositories of manure types used in various information systems including fertilization operations / manure plan, aims to achieve a common model nationally.
Several reference tables exist in the various information systems related to livestock manures. Some are considered valid at the national level, others describe more local variations.
The work under progress has for objective to propose a common data model.

B. The need for management tools and reference data models

The difficulties encountered during exchanges of different models produced by different Software Engineering Workshop (SEW) led Cemagref to work on interoperability between SEW.
SEW generally produce exchange XMI files (XML Metadata Interchange) but the different versions of used UML (Unified Modeling Language) and the different versions of XMI import products often make it difficult.
The aim of recent work is to study the compatibility between XMI files to improve interoperability between SEW.
These studies showed that some SEW do not produce files compliant XMI associated.
To do this the application developed using pivot language as the language of the Web Ontology Web Ontology Language (OWL). Thus the application MOST (Model Ontology Stylesheet Transformation) is able to transform such a file on a standard XMI 1.0 file in XMI 2.1.

The MOST application includes three processes:

- Fixed XMI file to be compliant XMI source
- Processing of source files XMI file OWL
- Processing files OWL XMI file target.

These improvements to the interoperability of information systems agri-environment is a preliminary step in building an agri-environmental ontology.

**Conclusion**

In 2006, within the Resagri project, AEE has obtained very interesting results:

- mapping information exchanges "around farms" and
- testing an eb-XML platform that is now used by a number of agricultural co-operatives.

The great advantage of an eb-xml platform of exchanges of information is that the sender of information is certain that this information will be transmitted in its integrity to the recipient who is certain that the information is really coming from the sender said.
This securing of trade is not essential as far as the "environment" data are concerned, but if an eb-xml platform is used for other information exchanges, this platform will probably also used for exchanging environmental data.
These other exchanges, what are they today? The mapping established by EEA and its project Resagri partners showed that farmers exchange information with their technical and economical stakeholders when it is mandatory: farm book-keeping and identification and traceability of animals mainly. The exchanges of traceability information in crop productions are extremely low because in this area, there is no obligation.
In this context, which importance can take the exchange of "environment" information? The answer to this question will depend on the coordinated efforts of software and R & D

organizations, government offices, cooperatives, and all actors who have developed software decision support (DSS).
Whatever be the qualities of DSS software, they will not be used if they have no real interfaces and therefore opportunities for exchanging information with existing FMIS. The publishers of FMIS generally have a key role not only in the development of FMIS, but also in the distribution of FMIS and user training. If tomorrow these publishers find ways to expand export data in standard formats to transmit the data to servers of DSS producers, and to retrieve results to be included in the FMIS, then the exchange of information "environment" will increase quickly and may exceed the volume of accounting data or animal traceability data.
A difficult question about electronic data exchange is the need for standards: can we settle for standard "local" adopted by our organizations, or is it necessary to propose and negotiate standards to the relevant international organisations?
The development of "local" standards can certainly meet our needs and costs less working time and less meetings than the development of an internationally recognized standard.
But the development of international standards has real advantages:

- Methodology perfectly developed,
- Discussion at international level for further work,
- Security of investments made

Of course we are probably not advanced enough to provide a general standard for the exchange of "environment" information, but after the e-Daplos UN / CEFACT standard, AEE can usefully propose as a standard message agro-obs**.**
Exchanges of "environment" information do not exist today but can quickly become significant but this requires:

- the development of interfaces and therefore communication messages by teams like AEE (FR) or Agroconnect (NL) between FMIS and all the tools for modeling, simulation, decision support that have been developed over the years 1980
- taking into account in the FMIS, and thus by their developers (computer companies agricultural machinery manufacturers, etc.) information required to run these tools for modeling, simulation, decision support
- taking into account in models, simulation and decision support tools, and therefore by their developers (research centers, universities, chemical company, administrations), information already present in the FMIS, information which we must try to valorise... before asking users to enter new information, any additional input being an additional obstacle in the use of this kind of tools

We see that the development of necessary exchange of "environment" information requires a lot of cooperation and consultation, but the way to an Agriculture that has to become both more productive and more environmentally friendly, using more efficiently inputs (y including water), definitely requires progresses in this domain of exchange of "environment" information.

## References


Bosc (J P), Samie (J F), Vibert (J), Lepaumier (B), 2005, Atractiv - application de l'outil d'aide à la décision - L'exemple de producteurs de légumes - Infos CTIFL

Bresson (J), Garcin (A), 2008, Sensibilité des pèches à Xanthomonas Arboricola pv. Pruni - conditions climatiques, Infos CTIFL

Brun T., Poyet P., Bopp M., Vigier F. (2005) « Towards an agricultural ontology in France: Contributions of the farm Information Management Project (GIEA) », EFITA/WCCA 2005 Joint Conference, Vila Real, Portugal, July 25-28, p. 1296-1302

Garcin (A), Giraud (M), 2010, Maladies bactériennes des fruits à noyau et à coque, Infos CTIFL

Institut de l'Elevage. Dexel Méthode et référentiel - Collection : Synthèse - Dépôt légal : 2e trimestre 2006 Tous droits réservés à l'Institut de l'Élevage - Réf.190633002 -ISBN 2-84148-211-1

Institut de l'Elevage. Logiciels pour les éleveurs et les techniciens

Voir : http://www.inst-elevage.asso.fr/html1/spip.php?page=un_espace&id_espace=938

Poissonnier (J), 2007, Outil de prévision des risques - le modèle Inoki Alternaria de la carotte, Infos CTIFL

Serot (S), 2007, Intégration des outils d'aide à la décision risques phytosanitaires et prévisions météorologiques, Infos CTIFL

Roussey C., Soulignac V., Champomier J-C., Abt V., Chanet J-P., 2010, Ontologies in Agriculture.

See: http://cemadoc.cemagref.fr/cemoa/PUB00029288]

Waksman G., Brun F., Donnat E., Coffion R. and Wallach D. (2010), Decision Support System in French Agriculture: The need for information exchanges, IAALD congress, May 26-29, 2010, Montpellier, France

Waksman G., Suau A., Coffion R. and Holl Cl. (2009), CAP subsidy claims: trying to encourage French farmers to abandon paper forms and to use the Internet, EFITA 2009, July 6-8, 2009, Wageningen, The Netherlands

Waksman G., Escriou H. and G. Gentilleau (2007), Situation of ICT in French agriculture, EFITA 2007, July 2-5, 2005, Glasgow (UK)

Waksman G. and Masselin-Silvin S. (2005), Traceability and cross compliance: towards common data descriptions, EFITA 2005, July 27-29, 2005, Villa Real, Portugal.

Waksman G. and Coffion R. (2005), Promotion of the use of ICT by the agro-food sector, EFITA 2005, July 27-29, 2005, Villa Real, Portugal.

# NASYP: Use of environmental modeling for preventing the impacts of major accidents with dangerous substances

Daniel Kahuda[1], Marek Vach[3], Otakar Čerba[2], Jan Bojko[2], Jan Ježek, Přemysl Vohnout[2], Vítězslav Dvořák[3]
[1] *VODNI ZDROJE, a.s., hydrogeological departement,* [2] *Czech Center for Science and Society,* [3] *Czech University of Life Sciences in Prague, Faculty of Environmental Sciences, Department of Water Resources and Environmental Modeling*
E-mail: jan.bojko@gmail.com

**Introduction**

Use of modelling tools integrated into database applications are valuable addition for state administrative authorities and stakeholders involved in IPPC and accident prevention agenda. A prototype tool (NASYP) has been developed in the form of a uniform geoportal component to evaluate the risks in harmony with the implementation of relevant regulation. The tool enables simple editing and analysing to simplify following actualization of information, classification, etc. The results contribute to the Law No. 59/2006 of the Collection of Laws on prevention of major accidents in the Czech Republic (implementation of the directive of the European Parliament and Council No. 2003/105/ES from December 16th, 2003 in the Czech Republic) and the related operating regulations. The aim of modelling is to provide sufficient processing and analysing capacity for rapid decision making during early stage of accident situations. The tools cannot substitute detailed analysis necessary for assessing the impacts of contamination occurred, but it enhances the operability of first stage measures preventing the impacts of major accidents. The modelling focuses on two transport pathways – surface water and air – supposed to be critical for early accident situation measures. To keep the modelling tool both up to dated and simple to use there're precalculated data combined with actual data collected from Czech Hydro Meteorological Institute observations and observing from the accident site.

**Modelling the contaminant transport in surface streams**

In the project framework there was a simple linear tool developed to simulate pollution transport scenarios for surface streams in the pilot region (Liberec Region), where most of input data were processed with use of GIS software. In the proximity of major and minor surface water streams there're sites chosen having a potential contaminant leak (in most cases it's going on industrial plants containing dangerous substances accordingly to regulation nr.2003/105/ES on the control of major-accident hazards involving dangerous substances). There're buffering zones in the radius of 1km around these sites determined where's potential accidents are likely to occur having impact on surface water source. The model calculates 2 basic values. The first is the transport time necessary for the contaminant advection from point A to point B accordingly to actual flow rate in the open channel of the water stream. Second value is the concentration of contaminant on specific location. Under the necessary input data we recognize hydrological data - flow rate measurements at reference profiles - data on individual sections of surface streams (length, average flow rate, catchment area) and the volumes of released contaminant.

For these model calculations the data are prepared using GIS software. In the vicinity of the watercourses there're selected areas with potential leaks polluting substances. Concerning leakages to surface water clusters with a radius of 1 km are used to identify possible pathways of the contaminants. The model calculates two basic variables. The first variable is the time for a contaminant to reach from point A to point B concerning the current flow. The second variable is the concentration of pollutant at a particular time. Among the required data for calculations (the hydrological model) we can identify the flow rates at specific sites, data on individual sections of the streams and data on the quantity of discharged pollutant.

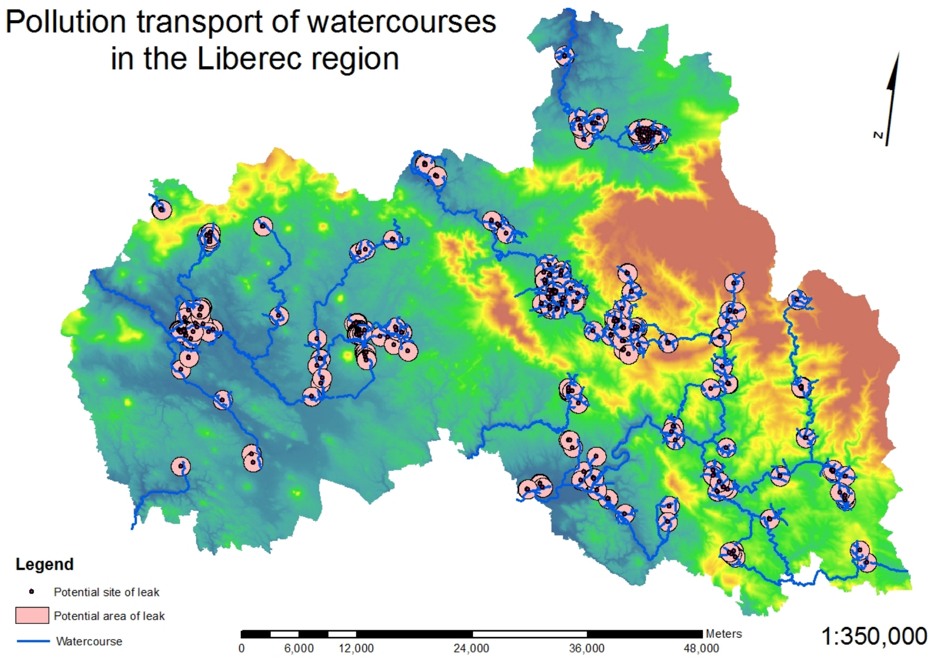



*Figure 1. Streams and potential threats for surface water contamination – Liberecky kraj, Czech republic*

For time measurements for contamination transport from point A to B, the method of sections is applied. With the use of interpolation and data modification, we know the average flow rate of individual sections of the watercourse and their lengths. This calculation is performed for all sections related to the possibly contaminated site and monitored at the end point. The resulting times are summed to create the final time. The necessary quantities for calculating the transport time of contamination are average flow velocity and length of individual segments. The average velocity set in individual sections is calculated using the methodology published in reference "Protecting agricultural land from erosion" (Janecek, 2002). This methodology estimates the velocity based on the average watercourse gradient and information on hydrotechnical adaptation of the watercourse. The longitudinal gradients are determined based on the Liberec Region Digital Elevation model (DEM). The hydrotechnical adaptation differs whether the flow channel is paved or unpaved having high importance

because of different flow velocities and flow rates consequently (with a lower friction the flow velocities are higher in paved channels) and is established on the basic aerial photo-map of the Liberec Region. Length is determined by measuring the individual sections of the liner layer of water flows. Pollution time from the upper end of the section to the lower end = length / speed. This calculation is performed for all sections between the contaminated site and the referenced end point. Summing the resulting time we obtain the resulting final time.
Concentration measurements of the pollutants are determined through flow rate interpolation in individual sections. Furthermore, we know the amount of spill and the time at which the contaminant escaped. First step is to find out the mass of water flowed during the contaminant discharge time (V = Q * S). Secondly the volume is increased by volume of the contaminant itself. The sum of share volume and contaminant concentration is resulting for the referred time. For the next sections we consider as if the contaminant flowed at the top of this section, which consequently changes the water flow on underlying sections. The measured concentration value is used for each section separately.
When measuring time transporting contamination from point A to B, we use the method of sections. Interpolation is performed using different sizes of catchment areas at a site with known value of the flow rate (gauging profiles by Czech Hydro Meteorological Institute - CHMI) and on the edges of rivers, which empty into the profile of follow after it. Specifically, the size of the basin on the edge of the stream is divided by basin size in the specific profile and the result multiplied with the flow rate measured at the reference profile. Flow rates for individual edges are the hourly intervals depending on the changing flow in reference profiles. Furthermore, we know the quantity limitation and the lasting time at which the contaminant escaped. First we find out the volume of water flowed over a period of contaminant discharge (flow rate = volume * time). Then add up the volume with the volume of contaminant. By sum we divide the contaminant volume and obtain the concentration relevant to specific section. For further sections we consider, as if the contaminant flowed at the top of this section. It will change the water flow rate, which will be on low-lying sections above. The concentration is measured at each section separately. The concentration of pollutants is then evaluated accordingly to applicable standards and regulations of the Czech Republic dealing with pollution in surface waters.
Part of the project was also the integration of the above mentioned model to the emergency portal. We have developed a process for calculation of the time and water contamination. The process was created according to the WPS (Web Processing Service) standard 1.0.0, therefore can be easily accessed by another service or user. The necessary input for calculation are coordinates of sites A and B, which we send to processing in the GML (Geography Markup Language) format. However the process can easily be enhanced for reading some OGR formats. The output in the XML (Extensible Markup Language) format from the process represents path of the pollution, transport time when a contaminant reach point B and also every segment possess information about contamination.
The process is based on several GRASS modules, which are mostly used for searching the path of the pollution. The PyWPS (http://pywps.wald.intevation.org/) is used as an essential part of the solution for its simplicity, robustness and possibility to cooperate with other spatial or statistical software. Therefore we could use GRASS for spatial analysis and python for calculations, moreover through the PyWPS we can access the database of sites.

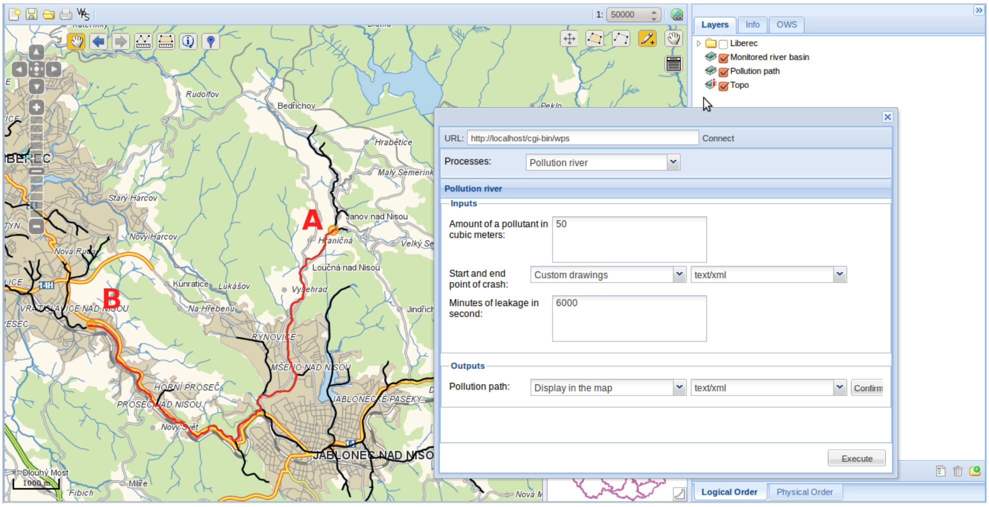



*Figure 2. Example of the HSLayers WPS client (black line: monitored river basin; red line: polluted river segments)*

The simplicity of this model and its implementation allows a rapid use in emergency situations. Due to the large scale of the model area just the principle of advective diffusion is applied. Available data do not allow more detailed modelling.

**Modelling the air pollution transport**

The aim of model calculations is to carry out the theoretical analysis of potential impacts of accidental pollutant releases from sources in the Liberec Region on tropospheric air quality in their neighborhood. Modelled outputs - field ground concentration of pollutants emitted only apply to extraordinary emergency situations of potential losses from evaluated sources, which are the standard operation of the technology equipment effectively eliminated.

Sources of emergency leaks can be in most of the cases considered ground. Technological equipment of which can result in disruption of emergency leaks, integrity, etc. escape pollutant are usually located at the level meters, more than tens of meters above the surface. The smoke plume is then formed in contact with the surface in the flow field of determined shape much of the surface around the source.

Calculation of the three-dimensional flow fields in the boundary layer of the atmosphere in the study area is performed by the computational fluid dynamics (CFD) code as a numerical solution of the system of Reynolds averaged Navier-Stokes (RANS) and continuity equations. The Reynolds stresses tensors is calculated following the Boussinesq hypothesis. The isotropic coefficient of turbulent viscosity μt is obtained from the turbulent kinetic energy and its dissipation rate using the standard k - ε model of turbulence closure (Jones and Laudner 1972). Numerical steady state solution of the system of equations was performed by the standard method of control volumes in a three dimensional non-orthogonal boundary fitted grid using a collocated variable arrangement in hexahedral cells. The lower boundary of the computational domain is delimited by the detailed model of the surface in the study area in the altitude map. The domain of 11000 m by 11000 m in the horizontal and 1000 m in the vertical was gridded with a 220 × 220 regular grid in the horizontal plane and a 27 node grid in the vertical direction. The vertical step in the hexahedral grid was designed as boundary layer i.e. was increasing in the direction towards the upper boundary of the computational domain by using a factor 1.2 for the distance increment. First vertical grid node was adjusted 1 m above ground surface.

The calculation of the flow field in a neutrally stratified PBL was performed for eight basic wind directions set through the upstream boundary conditions as lateral inlet boundaries of the computational domain. The inlet wind profile was set through the logarithmic law.

At the inlet profiles for k – ε turbulence modeling is set estimation approached by Richards and Hoxey (1993). At the lower ground surface boundary of the computational domain, both components of the turbulent energy were standard determined by the wall functions.

The second order upwind approach was used as the spatial discretization scheme. Temporal discretization was achieved by implicit scheme. The continuity was resolved through a pressure correction of the SIMPLEC type (Van Doormal and Raithy 1984). To get steady state solution, the iterative process was converged after about 230 iterations.

The modeled three-dimensional vector flow field generated input data for the related Lagrangian simulation of transport and diffusion dispersion of passive, gravitationally non-falling (or neutrally buoyant) model particles assumedly emitted from a point source. Local transport motion of each particle is determined by the velocity flow vector active in the partway of this motion. This vector is obtained through linear interpolation of the input flow field. The motion of the particles simulating the dispersion through turbulent diffusion is based on solution of the Langevin equations. The components of the fluctuation vector are given by Equation:

$$u_i(t+\Delta t) = au_i(t) + b\sigma_{u_i}\xi + \delta_{i3}(1-a)T_{Lx_i}\frac{\partial}{\partial x_i}\left(\sigma_{u_i^2}\right)$$

The new position of the passive, gravitationally non-falling particles is caused by the joint action of the two vectors.

$$x_i(t+\Delta t) = x_i(t) + (u_i + u_i)\Delta t$$

The amplitude of the modelled turbulent fluctuations including random parameter ξ varies in the range delimited by b = (1 – a2)1/2; a = exp(-Δt/TL); TL is the Lagrangian integral time, which, in agreement with the approach of Zannetti (1990), was set at 200 s in the horizontal plane and 20 s in the vertical plane. Variances of wind speed components σu were estimated from the modeled turbulent kinetic energy: σu = 0.91k1/2, σw = 0.52k1/2. The relevant constants are taken from the work of Panofsky *et al.* (1977) and correspond to neutral stratification. The time step for the motion of passive particles Δt was selected as 5 seconds. The emission flow of the source was set at 2000 fictive particles per Δt.

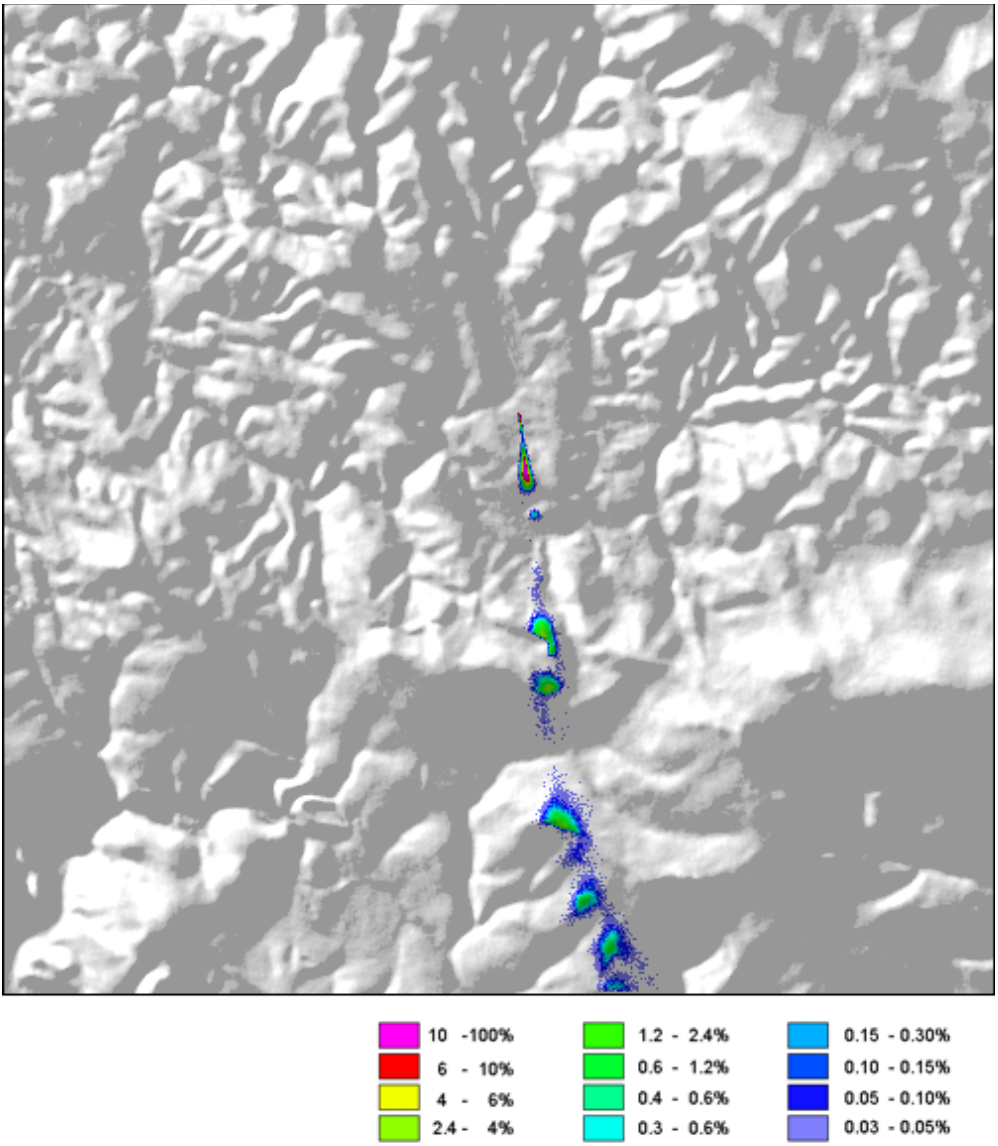



*Figure 3. Example of modelled concentration field relative gaseous contaminants (fictitiously emitted from a source nr.528) in the ground "breathing" layer for the north wind typed on the border area. Air pollution concentrations are expressed as a percentage of the maximum value achieved in the immediate vicinity of the source.*

**Acknowledgement**

The paper was prepared on the basis of outputs from the *NASYP* project, whose solution was achieved with financial support from the state resources provided by the Ministry of Education, Youth and Sports of the Czech Republic for support of project of the program "EUREKA" with registration number OE09015.

## References

www.geoportal4everybody.eu

Janeček, M. a kol.: Ochrana zemědělské půdy před erozí, metodika, Výzkumný ústav meliorací a ochrany půdy v. v. i., Praha, 2007

Jones, W. P., & Launder, B. E. (1972). The prediction of laminarization with a two-equation model of turbulence. Int. J. Heat Mass Transfer, 15, 301-314.

Panofsky, P. A., Tennekes, H., Lenschow, D. H., & Wyngaard, J. C. (1977). The characteristic of turbulent velocity components in the surface layer under convective condition. Boundary Layer Meteorology, 11, 355-361.

Richards, R. J., & Hoxey, R.P. (1993). Appropriate boundary conditions for computational wind engineering models using the k–ε turbulence model. Journal of Wind Engineering and Industrial Aerodynamics 46/47, 145–153.

Van Doormal, J. P., & Raithy, G. D. (1984). Enhancements of the SIMPLE Method for Predicting Incompressible Fluid Flows. Numer. Heat Transfer, 7, 147-163.

Zannetti, P., Air Pollution Modeling – Theories, Computational Methods and Available Software (Wit Pr/Computational Mechanics,1990).

# Enabling the use and exploitation of environmental data for decision making: Bridging the gap between needs and requirements

[1]Ana Saez, [2]Karel Charvat, [1]Francesco Molinari, [3]Mariano Navarro, [4]Lisa Maurer
[1]*Next Step International s.r.l.,* [2]*Help Service Remote Sensing,* [3]*Empresa de Transformación Agraria S.A, Technical University Graz*
E-mail: a.saez@nextstepint.eu

## Introduction

The HABITATS' wide range of 'real-case' (s) challenges and 'tangible' solutions are addressed to supports an IST development process based on an enhanced understanding of past and current challenges. The sound progress shared in this chapter is the result of the integrated efforts of the HABITATS' consortium: Empresa de Transformación Agraria, Spain; Next Step srl, Italy; National Microelectronics Application Center, Ireland; Help Service Remote Sensing, Czech Republic; Tecnologías y Servicios Agrarios, Spain; Technische Universität Graz, Austria; Ente Parco delle Madonie, Italy; Istituto Superiore per la ricerca è la protezione ambientale, Italy; Ústav pro hospodářskou úpravu lesů Brandýsnad Labem, Czech Republic and Latvijas Universitãtes Matemãtikas un informãtikas institũts, Latvia. This consortium is formed by research institutes, national agencies, SMEs, and local authorities from five countries in Europe. The project's consortium is formed by a highly qualified team built around research institutes, universities and SMEs working in distinct but complementary areas of research, application to busnissess, national agencies, all of which provide the skills and experience critical for addressing future challenges in IST and ICT.and the the successful implementation of this ambitious project. HABITATS consortium constitutes a sound pillar for advancing on the linkages between knowledge sharing and innovation in ICT and the application of ICT to address real needs of local, national and regional stakeholders. Therefore aspects of how ICT could promote sustainable development and rjural development are also part of the HABITATS project as one of the greatest challenge of our time: meeting the needs of the present without adversely affecting future generations.

## The HABITATS project

The HABITATS project (Social Validation of INSPIRE Annex III Data Structures in EU HABITATS) focuses on the adoption of INSPIRE standards through a participatory process to design and validate Data, Metadata and Services Specifications with real citizens and businesses. The access/availability/transformation/interoperability/adoption of environmental data in the European context under the EU INSPIRE Directive (http://www.inspire-geoportal.eu/ http://inspire.jrc.ec.europa.eu/) is at the core of the HABITATS project.

The INSPIRE Directive considers that spatial information is needed for the implementation of Community policies which must integrate environmental protection in accordance with Article 6 of the Treaty, and establishes the basis for an infrastructure for spatial information in

Europe in order to support EU environmental policies and those activities which may have an impact on the environment, defines 34 spatial data themes related to environmental applications and requires, in order to ensure that infrastructures of the Member States are compatible and usable in a trans-boundary context, that common Implementing Rules are adopted for all Member States, in Reporting. specific areas: Metatada, Data Specifications, Network Services, Data and Service Sharing and Monitoring.

The HABITATS project offers different connecting points to provide a link between the INSPIRE meta-data developments and the local data providers and users. Within the core community, the seven pilots introduced above represent the basis for the social validation process. These pilot communities are predominantly involved through events on a local basis: through workshops, meetings and other dissemination events. In addition to this strengthening of "real-life" communities, these interest groups can participate in online communities as well to discuss experiences on a local, national or international scale.

Against this background, the methodological approach in HABITATS identified as meaningful to relate the social validation process to the critical learning path of current and future development of the project. Thus representing the core community of the HABITATS' project has been a key component in all the wide range of cross-country and cross-pilot community of domain experts, practitioners and policy makers as well as communicating in English. This core community aimed at extending the community on a local and an international scale following different community building cycles to involve additional learning communities Thus one of the important objectives of HABITATS is to demonstrate the benefits of the INSPIRE adoption in practice by involving different user communities, represented by the pilots. First results have shown that the main benefits seem to emerge through the need to interact with other stakeholders and the possibility of providing a better service to specific user groups The different aspects of community involvement are shown below (Table 1.) depicting the different levels of user involvement and habitats-related data within the HABITATS social network:

**The social validation, within the living lab approach, at work: 7 real study cases**

The basis for the efforts of HABITATS [2]consists of the seven pilot areas, dealing with different habitats-related issues[3]:

Management of Natural Resources

- Wild Salmon Conservation (Ireland)
- La Palma Protected Marine Area (Spain)

[2] NEXT STEP srl has developed this methodological framework for the project. This effort is based on the conceptual framework of Complex Systems. The work has been done as a contribution to some of HABITATS' reports.

[3] For more information on the focus of the pilots please see the HABITATS website http://www.inspiredhabitats.eu/ www.inspiredhabitats.eu

Eco-Tourism

- Hiking Trip Planner (Italy)
- Soria Natural Reserve (Spain)

Economic Activities

- Sheep and Goat Herd Management (Italy)
- Economical activity at marine coastal benthic habitats (Latvia)

National Policy

- Czech National Forest Programme (Czech Republic)

These local pilot communities, which are active in their respective languages, together with a cross-country and cross-pilot community of domain experts, practitioners and policy makers, communicating in English, represent the core community of the HABITATS project. This core community is aimed at extending the community on a local and an international scale following different community building cycles, under the Living Lab approach, to involve additional learning communities.

One of the important objectives of HABITATS is to demonstrate the benefits of the INSPIRE adoption in practice by involving different user communities, represented by the pilots. First results have shown that the main benefits seem to emerge through the need to interact with other stakeholders and the possibility of providing a better service to specific user groups.

The different aspects of community involveent are shown below (Figure 1) depicting the different levels of user involvement and habitats-related data within the HABITATS social network:

**Habitats pilots**

According to the HABITATS user-driven approach to standardisation, the full impact of results will be sparked off by the pilot service scenarios and their ability to attract new participants to the communities of adoption. Each of the HABITATS pilots is therefore built on:

a) existing concrete services currently carried out by project partners,
b) potentials of data access through network services and
c) enhancement through usage scenarios developed by user communities, in order to meet the three HABITATS criteria of relevance, openness and responsiveness.

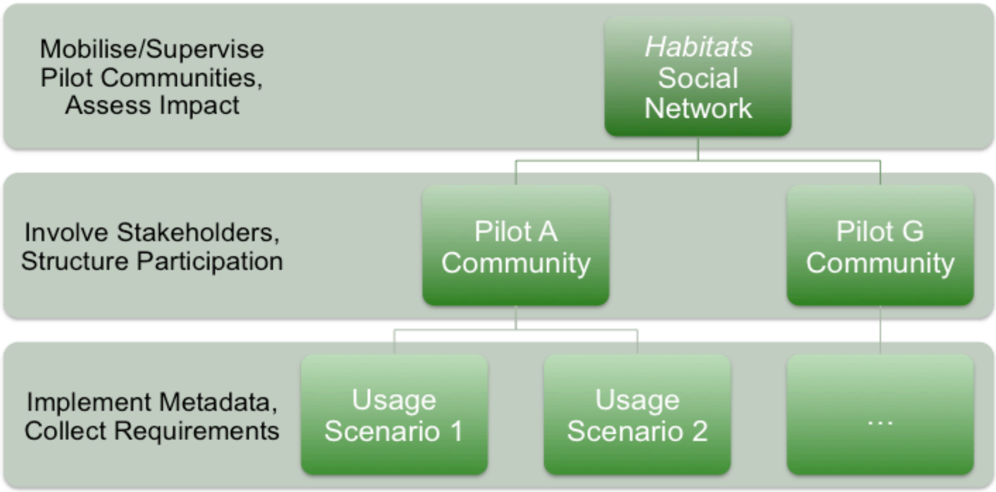



*Figure 1. Prototype for collecting community metadata (HABITATS D2.2.1 2010)*

In addition to these elements the development/identification of use cases within the different usage scenarios represent detailed application cases and serve as specific entry points for community discussions. Based on the community engagement, the social validation process focuses on assessing the social importance of the effects of any practice, standard or any other issue that is central to the project.

In sum, the proposed (multi-perspective) framework of assessment is characterized by user inclusion, as well as their engagement into social validation. It can be seen in a way as instrumental to better understanding and representation of impact, without incurring into limitations of time and logic. This ICT supported framework responds to:

1. The *social significance of stated goals*.
2. The *social appropriateness of followed procedures*.
3. The *social importance of obtained effects*.

The specific usage scenarios, including the state of the art baseline and user requirements coming from them represent the key input for the planned data and meta-data modeling activities and the SDI services that will be developed in the further course of the HABITATS project.

One of the key objectives of Habitats is to evolve from and further develop the taxonomy of social validation 'ideal type' approaches. Putting it differently, the AS-IS (and TO-BE) situation of each HABITATS pilot could be represented as the intersection of three main axes – or impact drivers:

- The level of SDI compliance to INSPIRE prescriptions for the Annex III data themes 16-19 in terms of data and metadata models, standards and application schemes;

- The extent to which new services are developed – or upgraded – with or without proper SDI compliance;
- The degree of end-user satisfaction – however it is defined and measured – regarding those services.

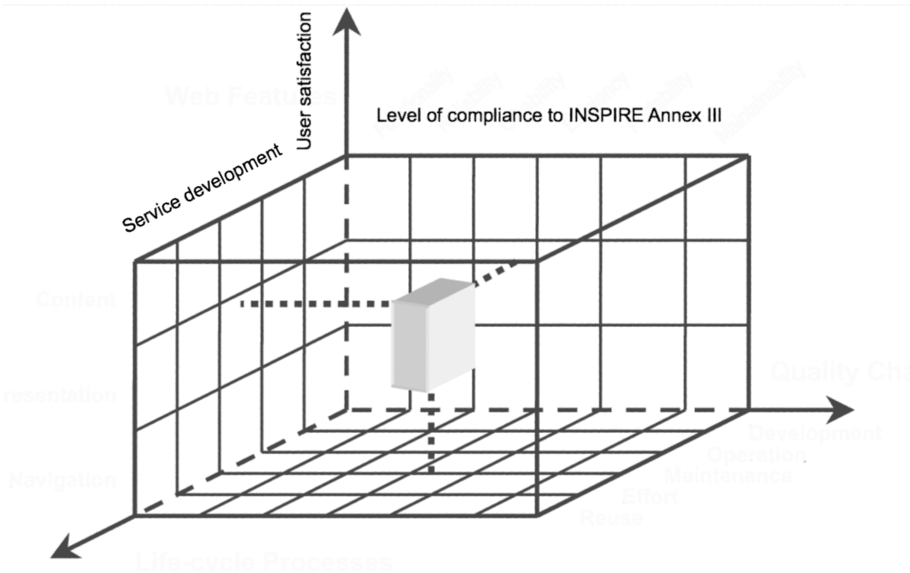



*Figure 2. Drivers of impact in HABITATS pilot*

Generally speaking, we detect a positive correlation in all the pilots, between service development and user satisfaction, while on the other hand, it cannot be taken for granted that the new services provided are also INSPIRE compliant: this can be due to several reasons, two of which seem more prominent than others:

- On the "supply side", the cost of increasing the compliance, in terms of time, resources, etc., from the perspective of the SDI "owner";
- On the "demand side", lack of interest or simply ignorance of the advantages of compliance, from the perspective of the end users.

To exemplify the contrasting effects of these factors, we focus on the hypothetical relation between the number of services offered and the percentage of them that complies with the INSPIRE Directive (reference is made to the Extended Impact Assessment of INSPIRE). As the following picture displays, in principle we may identify four possible areas of consideration (or "quadrants"), each corresponding to the situation of one or more HABITATS pilots

- Low number of services, poor compliance;
- Low number of services, highly compliant;
- High number of services, poor compliance;
- High number of services, highly compliant.

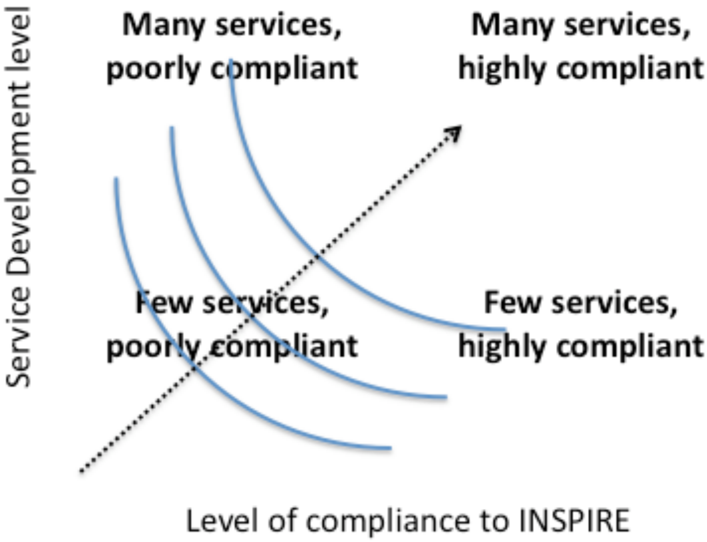



*Figure 3. Service development Vs. INSPIRE compliance*

Each point on the curves that are depicted here represents a (hypothetical) combination that an SDI "owner" may opt to choose in order to spend a given budget assigned to it (in terms of time, resources etc.). One can easily assume that moving along the same curve, from SE to NW, to create several new services may imply some sacrifice in terms of compliance[4], and vice versa, moving from NW to SE, to increase the level of compliance of existing services may exhaust available resources for the creation of new ones[5].

Certainly, jumping from one curve to another (in the direction indicated by the dotted arrow) means to increase the level of user satisfaction, though at the (likely) price of spending more and more resources on both compliance and services[6]. Yet the question is: are we sure that there is no other way to achieve a given target of satisfaction in the user community (supposing it is measurable)? Couldn't it be that, for instance, the SE quadrant compared to the NE, where more services are offered with a similar level of compliance, has the same "value" for prospective users, just because the additional services produced are not of interest to them? Or couldn't it be that for a similar reason, the NW quadrant provides less "utility" than the SE one? Or that for a given number of services, moving from SW to SE does not change much the "wealth" of the current end users, because they do not clearly perceive the

[4] For sure, the amount of resources invested on making an existing service INSPIRE compliant may be lower than those that are needed to add new services to the SDI. Though this may not be taken for granted in all cases, its only implication would be that we are preferably staying on the "flatter" part of each curve, rather than on the "steeper" one.
[5] If, by mere chance, that was not holding true, and the number of services could be increased together with the desired level of compliance, without affecting the budget availability, this would practically mean that we are not "sitting" on the "right" curve, and that we could safely "jump" to the next one, as explained later in the text.
[6] See the previous footnote.

advantages of getting a higher level of compliance? In short: which pathway is to be followed in order to maximise community benefits, while keeping development and/or compliance costs at the lower possible level?

In the HABITATS project, the global answer to these questions derives from a deeper involvement of end-users in data modeling and service co-creation. In the conceptual definition of the social validation "space", the proposed focus of application for behavioural analysis was threefold, namely:

1. The *social significance of stated goals*. Do the specific development objectives correspond to what users really want? Are they fulfilling a need that is shared by the prospective end users? Does the broader community in which the SDI infrastructure is located value the new services as important to them?
2. The *social appropriateness of followed procedures*. Do the ends justify the means? Do users feel that they have a voice in SDI infrastructure improvement? How do they feel they are included in the development, implementation and assessment process? Do users and/or local stakeholders consider the procedures for their involvement acceptable? Do they recommend them in other situations?
3. The *social importance of obtained effects*. Are end users satisfied with all the results, including any unpredicted ones? Do domain experts value the effects and believe that they were indeed caused (or facilitated) by the INSPIRE compliant services developed on the SDI? Does the broader community appreciate the outcomes? Does it value them as something that should be extended to other databases?

The following table 1, matches the taxonomy of social validation "ideal types" with the three main areas of interest for behavioral analysis:

| Social Validation Type | Pilots involved | Focus of evaluation |
|---|---|---|
| Potential benefits of user involvement influence data modeling and standards adoption. | Wild Salmon Monitoring (IE)<br>Economic Activity of Marine Habitats (LV)<br>Forest Management (CZ) | Social significance of stated goals |
| Direct expert/end-user interaction with data modeling process. | Monitoring Protected Marine Area (ES)<br>Environmental Education in a Natural Reserve (ES) | Social appropriateness of followed procedures |
| User-driven co-design of services leading to "demand pull" INSPIRE adoption. | Hiking Trip Planner (IT)<br>Sheep and Goat Herding Management (IT) | Social importance of obtained effects |

*Table 1. Types of Social Validation in HABITATS Scenario-Building Processes*

Given that in most of the pilots, the actors represented as stakeholders are highly heterogeneous, and some of them are possibly very low skilled on ICT and poorly knowledgeable on the contents and implications of the INSPIRE Directive, the specific requirement has emerged in HABITATS for a common framework according to which the added value of standards adoption could be assessed in all usage scenarios.

Based on the evidence collected above, the following conclusions can be provisionally drawn. These are open to the evaluation of HABITATS partners, and of the European Commission, which would rely on a focused and effective management of the second phase of the project activities. It is to be noted that our work in impact assessment will have relevance and will provide insight and content to the efforts made by the Joint Research Center (JRC) and thus HABITATS will contribute to the IA TOOLS developed along the lines of the European Commission's Impact Assessment Guidelines SEC (2005)791.

HABITATS Impact Assessment is tapping unto key knowledge sources such as the INSPIRE Extended impact assessment[7]. In this context the lessons learnt from the work conducted with key stakeholders through the FDS working group is also a guiding principle to HABITATS's work. It is also expected that the INSPIRE FDS working group would enrich its wealth of case studies and expert views with those emerging from the implementation of HABITATS.

The INSPIRE Framework Definition Support (FDS) Working Group issued questionnaires with targeted questions on the impact of INSPIRE to specific user groups including the other INSPIRE Implementing strategy working group, the research communities such as AGILE, the private sector and local/regional administrations. An approach apparently close of that used by HABITATS but that is substance is different as the feedback emerging from stakeholders is based on the early involvement of stakeholders at each validation pilot. In essence the social validation in HABITATS is expected to be complementary to the impact assessment methodologies used by other groups.

Broadly speaking, the INSPIRE Directive and its implications are hardly familiar in any HABITATS pilot community. From this evidence, the global intuition supporting the project is confirmed, namely that it makes no sense to transfer the technical jargon of domain experts to the respective user groups, which are only interested in speaking their own language and seeing their immediate or prospective needs fulfilled by a growingly automatic and interactive SDI implementation in the selected thematic domain. This opinion is mostly shared by the competent authorities and knowledge centres themselves – though not to the extent that the ongoing, "top down" operated, process of transposition of the Directive into national legislation and praxis is somehow affected by it.

In this scenario, some considerations are in order. First of all, as the comparison between the nascent ISPRA pilot, the current Irish and the Latvian or Czech institutional environments is there to easily demonstrate, the process of INSPIRE conformity gets increasingly complicated when the relationships between the institutional setting and decision-making process of the country and the levels of decentralization and governance increase in complexity and/or the number of stakeholder communities involved grows up. In particular, what we have here in

[7] Contribution to the extended impact assessment of the INSPIRE Framework definition support (FDS) working group. Max Craglia. 2003.

HABITATS is a flourishing number of “bottom up” and often very “ad hoc” initiatives, which we have started to call “services” (such as the Augmented Reality environment in Campo del Moro, or the herd monitoring system in Sicily), which aim to respond to very specific needs and (emerging or actual) requirements from the broader population or selected communities out of it. And it is in these ad hoc services, like segregated “islands” from a plethora of more or less available SDI’s, that the level of INSPIRE conformity makes or can make a real difference. We note that while this emergence is a clear commonality of all the HABITATS pilots, it seems comparatively more likely to occur in the larger, rather than the smaller sized, European countries. This probably because in the former, rather than in the latter group, the level of centralised institutional control is weaker, leaving more space to autonomous initiatives from the public authorities and agencies belonging to the policy environment mapped by the INSPIRE directive.

Further to that, the evidence collected from the pilots in practically all EU countries shows a growing concern for the limitation of financial resources available to service deployment, and obviously to the process of INSPIRE compliance. Starting from this evidence, the HABITATS project introduces, like an additional decision-making variable for the allocation of scarce budgets to prospective conformity plans, the level of user satisfaction deriving from the services deployed (see Figure 4 above and the related discussion), or the possibility of prioritizing the time evolution of new and existing SDI’s in association with their perceived impact (and also, the possible economic return from) within the target user communities.
The more linked is the prospective SDI evolution to actual end user needs and requirements, the more crucial become the modalities adopted for the collection of technical feedback from these users, not only during the initial design stage, but all along the development and possibly also the validation of the various service prototypes.

In general, what the HABITATS pilot descriptions strongly support is a distinction between two kinds of end user – and feedback thereof:

- Individual members of the local stakeholder communities (e.g. public authorities, universities, domain specific research centres, business associations etc.). These are normally available on a continual basis to offer their acknowledgement and support for the ongoing deployment of INSPIRE compliant SDI services. Thus, they can be easily associated to all stages of design, development and validation. However, their technical contribution is usually weak, politically driven, or scanty and superficial, with respect to the specific features of the proposed solution. Moreover, they must be “forced” to contribute on a voluntary basis (as the Latvian pilot has stated) and there is no technical solution that can ensure this to happen.
- Actual end users of the system (e.g. tourists, shepherds, anglers, fishermen, etc.). Contrary to the previous group, they are deeply involved into system usage, thus their technical is an expert input which is more precise, and often essential in bugs tracking and usability or functional improvement. In general though, it is quite hard to collect the experience of these people right in the time and place they were using the system. There may be solutions, like the ones proposed in the Campo del Moro pilot, allowing to “jump” into an evaluation system right after the end of a working session and that the development of a systematic registry of the application would both aim to gain more insight on its usability. Yet once more it cannot be taken for granted

that all relevant feedback would be saved in that way. Furthermore, the contribution of this category of end user is mostly concentrated in the post-deployment stage, leaving the previous ones largely unattended[8].

What is probably missing from this scenario is *the possibility of exploiting stakeholder feedback during the development stage*. This lies among the main goals of the PaESI[9] platform deployment. PaESI has also the advantage of not placing any burden to prospective users of a system in terms of INSPIRE related knowledge. All key evaluative questions are proposed in natural language, and the respondents do not immediately perceive their association with the underlying technical development options.
However, the same process can be repeated – though less efficiently – also by offline discussions and surveys, followed by internal briefings and operational meetings between developers and strategic decision makers. This is what is currently happening at the level of most HABITATS pilots.
To conclude, what we are aiming to describe is a process that is staged by necessity, and also highly interactive at each of the proposed stages, as the following picture shows:

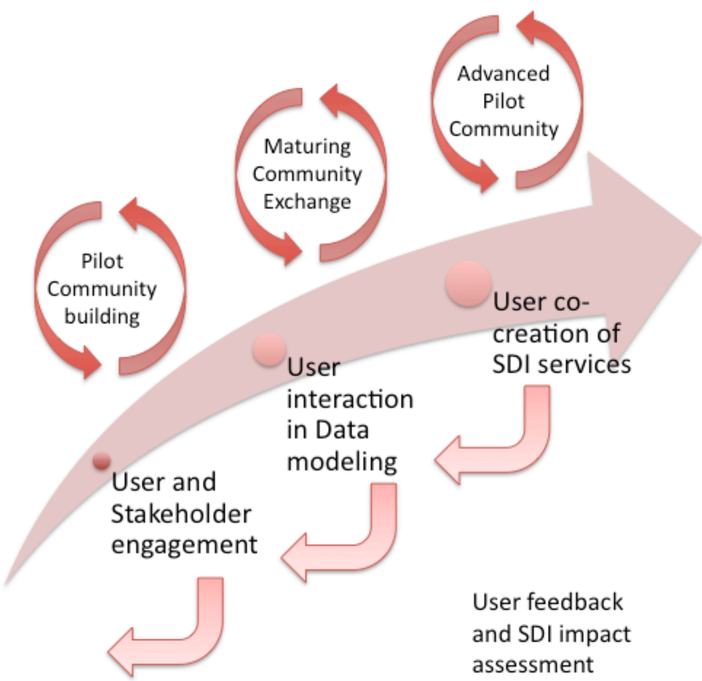



*Figure 4. The social validation 'ideal types' in HABITATS*

---

[8] Of course, the communities of reference for these end users – if existing – may act in support of the design and deployment phases. In which case, they become local stakeholder associations and their inputs can be classified as belonging to the first cluster.
[9] PaESI Plataform. Developed by NEXT STEP srl www.nextstepint.eu NEXT STEP is a partner in the HABITATS consortium

According to the evidence at hand, all the individual pilots may be positioned at a given stage of this process, namely:

- *Validation driven by the prospect of user engagement*

  In this case end-users are not yet directly involved in social validation, but the prospect of user engagement is already influencing institutional behaviour. Pilots involved are: Wild Salmon Monitoring (IE), Economic Activity of Marine Habitats (LV), Forest Management (CZ).
- *Validation through direct user interaction with the data modelling process*

  In this case there is, or will be, direct participation of (expert/non expert) users in data modelling. Pilots involved are: Monitoring Protected Marine Areas (ES), Environmental Education in a Natural Reserve (ES).
- *Validation driven by the co-design of innovative "demand pull" services*

  This is the most user-driven approach, as it actually involves final end-users in the co-design of services that use INSPIRE compliant data sources. Pilots involved are: Hiking Trip Planner (IT), Sheep and Goat Herding Management (IT).

Although the main arrow is there to communicate the idea of a sort of evolutionary growth, there is also the possibility for a given pilot community – once it has been properly established, which means that the first stage in the process is more or less common to all – to "skip" the interim stage and go straight to the third one. This is the case, for example, of the Sicilian pilot, where the first and foremost driver of change is – a little bit paradoxically, given the embryonic level of the underlying SDI and even the poor IT skills and equipment of the main prospective users – the possibility of co-creating a number of added value services directly with the SDI owner. But even at the first stage, like it has been documented in the Irish pilot, for instance, relevant and meaningful feedback as a guideline for action can be fruitfully gathered. This gradually shifts the attention from the plain collection of user feedback to the value that this may have for the SDI owner and the way it can be processed and ultimately taken onboard.

In summary we can also conclude that HABITATS is contributing to a new strategic approach for better policy making and use of the INSPIRE Directive. HABITATS' pilots are proving on real cases at the validation pilot sites that there is a large gap between what is required and what is needed. The bottelnecks identified by HABITATS coincide with those already published in several papers[10] where sound knowledge based on timely, accurate, easily accessed geospatial and environmental information, remains one of the main obstacles in terms of sharing information across European institutions, national agencies, local jurisdictions and stakeholders.

To conclude any SDI should be seen as an evolving concept that sustains (or mediates) various perspectives or stakeholder views. Depending on the user's interest and role within the broader community, its design and implementation (as well as the corresponding SDI

[10] The contribution of INSPIRE to sustainable development. A. Annoni.

assessment process) gets reshaped by a continuous negotiation and re-negotiation actors. In addition, 'space' – or the ultimate object of any SDI – is socially produced as well, which makes the validating role of socio-technical platforms like those proposed in HABITATS even more important.

From this situation, two directions for scaling-up are currently envisaged:

- Strengthening the potential of the so far identified "pilot twinnings", to enable data and service interoperability, particularly driven by the convergence of thematic domains, business interest and technology availability;
- Further elaborating on the evidence gathered from the respective policy environments (hereby including the additional "pilot" from ISPRA), to see whether and to which extent an update of the currently "top-down" transposition and conformity approaches can be facilitated by the emergence of successful "bottom-up" initiatives, like the ones described in the HABITATS pilots.

**Content and Structure of the social validation**

The implementation process of the evaluation paradigm is mainly based on the configuration of the PaESI platform at community level. In particular, the execution of the latter activity has been prioritised in the Sicilian pilot(s), with the purpose of creating a common database that can later be reused by the different pilots. Furthermore, a national-wide survey is being carried out in Italy under the care of the ISPRA partner, which is also preparatory to further localisation of PaESI in a selected number of regional and local administrations.

With an eye on documenting the interim status of social validation activities, the achievements so far are to provide an overviews of the current status of experimentations each HABITATS pilot site. However our main concern goes towards the ignition of a self-reflection process among the consortium members of the HABITATS project, regarding the impact generated by project activities within and outside the identified user communities.
To this end, a list of steps has been taken as such:

<u>Status of Community Building and Engagement</u>

- Conceptual basics for the social validation: creating a critical mass of multi-stakeholder partnerships (different mechanisms within HABITATS: creation of local social networks/participating in other networks…)
- Current status (including local social network groups)
- What do they wish to see and to contribute to in the social networks?
- What do stakeholders appreciate and recommend in terms of dissemination/awareness-raising and other mechanisms for user engagement?

<u>Face-to-face Involvement process: Workshops</u>

- Rationale for HABITATS at the pilot level
- Living Lab Approach (LLA) status of mainstreaming at the pilots
- Advantages associated with the co-design of services targeted to user's requirements

- Potential on-line and off-line services of HABITATS network services
- Discussion on current trends in technology
- What is at stake: technology accessible to regions
- The interlinks amongst the local/regional/national contexts
- Best/good practices in technological advances
- Bottlenecks of the process in technical/operational terms
- How the current policy environment enables or restricts standardization & adoption
- Identification of these policies - at what scale: local/regional/national/EU?
- How can HABITATS support better decision-making?

Impact Assessment

- What are the expected benefits of INSPIRE for your organisation?
- What do you think are the INSPIRE benefits for the community?
- And how about the combined benefits with HABITATS in the pilot?
- Identified gaps in the wide application of the different technological advances in INSPIRE. How have impacts been assessed?
- Innovation in terms of joint accountability of coordination/integration: co-designing evaluation of impact
- In your opinion, how useful is the social validation?
- What should be the criteria for evaluating the impact of HABITATS at the pilot?
- How to quantify and qualify it?
- Resulting benefits from feedback as part of the social validation: expectation of best case scenario
- What are the ‘enabling’ elements of better decision-making?
- What decision and policy framework and level of governance can be addressed by the integration of services/sharing data, etc.
- Expected impact from the overall project?

Additional output :

- Twinning opportunities for the pilots.
    - Potential for twinning/sustainability and up-scaling issues
    - What mechanisms to develop for this?

We have thus examined that the main aim of Social Validation requires the structuring of a wide-spread end-user participation in standards definition and adoption. Some of the mechanisms for user engagement that HABITATS has developed, social networks, will be explained at length in the next section.

**HABITATS Aggregated social networks**

The main aim of this essential building block in HABITATS' project is to structure wide-spread end-user participation in standards definition and adoption through social networks. The developments are following two iterative circles that lead through the different aspects of HABITATS Social Network. The interactive interlinkages of the tasks performed can be described as: different user communities are identified and brought together, while the pilot user communities use the platform for dissemination and awareness raising activities. From there the State of the Art, Scenarios and Requirements are developed in order to achieve the configuration of the validation pilots and support the stakeholder interactions. And finally the impact assessment leads to a framework definition that again supports the further development of the HABITATS Social Network.

The developments of the project had followed two iterative circles that lead through the different aspects of HABITATS Social Network. As such, the interactive interlinkages of the tasks performed can be described as: different user communities are identified and brought together, while the pilot user communities use the platform for dissemination and awareness raising activities. From there the State of the Art, Scenarios and Requirements are developed in order to achieve the configuration of the validation pilots and support the stakeholder interactions. And finally the impact assessment leads to a framework definition that again supports the further development of the HABITATS Social Network.

It is in this context that is worth noting the key characteristics faced with when working on setting up a social network were identified: First, many of the local communities are already active in their own respective networks and second, many stakeholders show some reluctance towards joining online social networking platforms. Based on these assumptions, a basic platform strategy had to be developed which takes these limitations into account while structuring and coordinating an active and effective community.

Therefore the approach of aggregated communities wa s chosen. This means that while a main community platform is used, this platform is not the only relevant platform. Instead, other communities augment the main platform; some of them directly linked to the project, some of them with different connections to the project, some of them communities that are thematically related and others that are just interested in the social validation initiative of the HABITATS project.

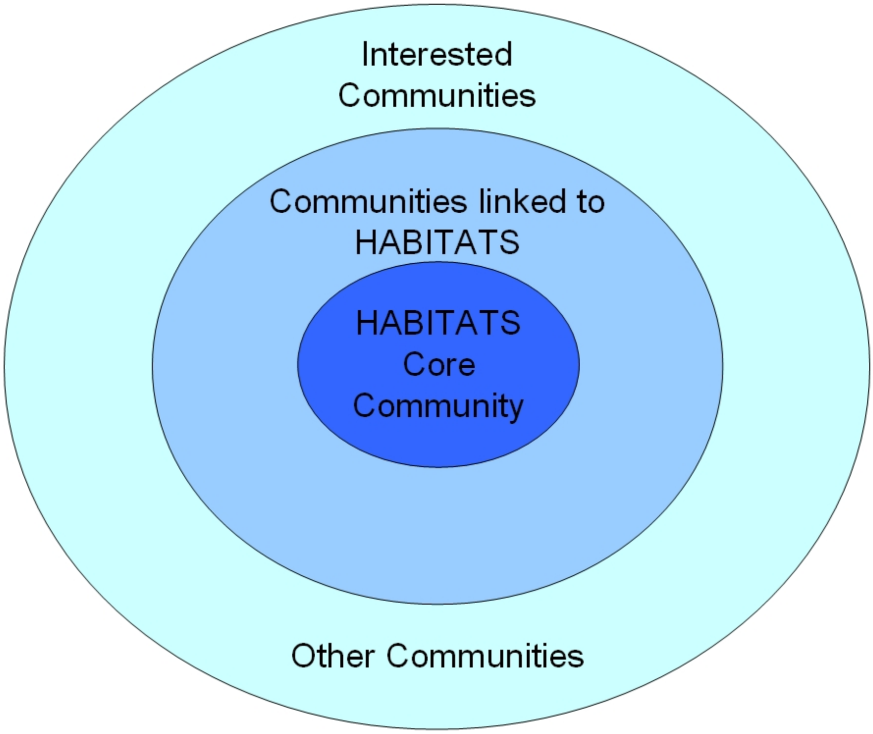



*Figure 5. HABITATS Learning Community Space*

This association of related communities represents the HABITATS Learning Community Space which is "owned and stirred" by the community. In addition to creating a community network of different layers of involvement and expertise, one of the main benefits can be found in the fact that users don't necessarily have to join another group but can still participate remotely through their favourite platform or already existing networks. If stronger interest is sparked, the core community can be joined in order to participate more thoroughly. One of the basic tools for creating a flow of information is using RSS feeds. Other strategies, especially involving community interaction, will be explored in the development process of the community.

In addition to this online community approach, the HABITATS core community also aims at establishing strong ties with local (pilot) communities and stakeholders that only exist in real life and not as only communities. The HABITATS pilot partners therefore play an important role as a link and a facilitator of the information flow between different types of communities.
This information exchange and development process can take place on a number of different levels:

1) INFORMATION: The main HABITATS website[11] (Inspired Habitats 2011) provides information about the project background, the partners and the pilots. It presents the scientific results of the project and informs about news and project meetings. For internal uses it also incorporates a document management system to facilitate information exchange between partners.
2) DISCUSSION: A central communication platform[12] for the HABITATS community has been established on NING (Inspired HABITATS NING 2011). This represents the communication hub of the aggregated community approach as described above. This platform (and the aggregated community around it) offers room for user-driven developments and creative discussions on different topics, e.g. technology, usability, cooperation, requirements, etc.
3) APPLICATION: The HABITATS Reference Laboratory[13] can be seen as the third main building block of the HABITATS social validation community. It allows users and data providers to test existing technologies and thereby generate further research tasks through a user-driven process. The Reference Laboratory will also collect information coming from other projects as input for analysis and discussions within the HABITATS social validation process.

These three elements are the main links that connect the different aspects of the HABITATS project. In addition to the above described community features, they support the management and coordination of and between project partners, and present developments and results of the meta-data modelling activities (work package 3) for further discussion and development while results from the network service architecture, including developed services, are provided (work package 4). The different pilots (work package 5) are given a platform for communication within and between the pilots and through that are the cornerstone for the social validation process. Dissemination (work package 6) is a welcome "side effect" of these processes, as the joint development and learning process goes far beyond simple one-way dissemination activities.

The community development and user engagement process is seen to be happening in three steps. In general the interaction of community members in a co-design process is of high importance as a driver for change and continuous improvement. For that an intensive information flow between different organizational levels (from different types of local community development to discussions in the online community) is aspired.
In general the community development that is strived for in the HABITATS project can be seen in three steps:

**First Phase: Formation of Core Community**
As the first step this information flow is mainly linear, feeding information and news from other groups and communities back to the main core community. Within this network, all project-related groups are even more tightly interwoven by notification tools. This helps to

[11] http://www.inspiredhabitats.eu/
[12] http://inspiredhabitats.ning.com/
[13] http://www.habitats.cz

generate an environment for the main project community in which the network can start to grow and develop, while receiving input and "nutrition" from related initiatives and networks.

**Consecutive Phase: Evolution of Learning Community**
After the establishment of the community framework and the necessary connections, the main objective has to be to breathe life into the interactions between community members while allowing for information flow to the linked and interested communities. This often takes the shape of engaging in a concrete project with shared goals, whose objectives could only be reached through the collaboration of several stakeholders.

Through the sharing of requirements and results and through the testing of applications and technologies, the basic community structure can evolve towards a "learning community", co-designing and developing new ideas and approaches. As the "learning community" begins to assume its own identity, it begins to attract an ever broader scope of stakeholders, including private companies and research institutes interested in investing resources to take advantage of the added value of the cooperative co-design setting.

**Final Phase: Circulation and Collaboration**
The final phase of the community building process can actually be seen as a take-off platform towards a collaborative community that begins to uncouple from the original project basis and follows the requirements of the internally evolved objectives and needs. This phase generally corresponds to the establishment of some form of permanent structure, from a loose Memorandum of Understanding to a legal entity, that can then carry forward other projects and initiatives, building on the "trust capital" constructed in the first experiments.

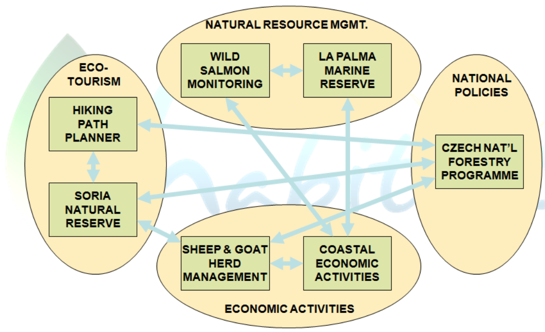
In this framework a number of different user groups have been identified that should be addressed. The four main user groups are: Public Administration, Scientific Community, Industry/Social Stakeholders, Users and General Citizens. The networking of these user groups and their participation in the social validation process is taking place on a number of different levels, where the different social networking platforms are only a small part of the whole. Especially on a local basis, the organised workshops and other face-to-face meetings are of high importance for triggering interactions and discussions. The online communities are especially aiming at creating links across regions, disseminating results and initiating discussions between different user groups.

The HABITATS social network is an approach to bring different interested communities together in order to disseminate information about the project's developments and results. But more than that, the social network offers an (inter)active entry point for all kind of interested people to actively participate in a discussion and validation process of the INSPIRE guidelines. It often seems that information gaps can be found between different actor groups for example between users and developers, system developers and integrators, policy makers and users, as well as between the local and the global view. The wide approach of HABITATS social network tries to present an arena that supports the open flow of information and of the corresponding discussions while reaching a wide range of interested users. This tries to support the validation processes happening in "real life" environments on a regional scale.

CROSS REGIONAL EFFORTS and TECHNOLOGICAL DEVELOPMENT AT THE SERVICE OF BETTER DECISION-MAKING

Much of the success of this is based on cross-regional services and data access. The pilots in the HABITATS project, fall into the following trans-regional categories:

1. Management of natural resources
    - Wild Salmon Monitoring (IE)
    - La Palma Protected Marine Area (ES)
2. Eco-tourism
    - Hiking Trip Planner (IT)
    - Soria Natural Reserve (ES)
3. Economic activities
    - Sheep and Goat Herd Management (IT)
    - Economical activity at marine coastal benthic habitats (LV)
4. National policy
    - Czech National Forest Programme (CZ)




In other to enhance the understanding and to find commonatlities for sharing knowledge acroos the pilots , the following trans-regional groupings had been identified.

1. Potential benefits of user involvement influence data modeling and standards adoption.
    - Wild Salmon Monitoring (IE)
    - Economic Activity of Marine HABITATS (LV)
    - Forest Management (CZ)
2. Direct expert/end-user interaction with a data modeling process.
    - Monitoring Protected Marine Area (ES)
    - Environmental Education in a Natural Reserve (ES)
3. User-driven co-design of services leading to “demand pull” INSPIRE adoption.
    - Hiking Trip Planner (IT)
    - Sheep and Goat Herding Management (IT)

**Habitats Data**

Data in HABITATS architecture is divided according to their origin and their usability. Basic types of data can be found in next table.

| Data type | Description | Examples of data | How data are used | How are data accessible |
|---|---|---|---|---|
| Reference data | Reference data are data, which are coming from other systems, and which are not changed inside of HABITATS | Topographic maps, Pan European data sets, administrative borders, cadastre, orthophotos, | Data are primarily used for visualisation, in some tasks they could be used also as background data for analysis (could define areas of analysis, buffers, etc.) | Data could be stored locally, or their could be accessible using Web Services (Web Mapping Services for visualisation, Web Feature Services or Web Coverage services for data analysis) |
| Satellite imagery | Data coming from Earth observation, which are used for analysis and visualisation | SPOT IMAGE, Quick Bird, MODIS (Aqua and Terra) | Data are primary used for data analysis, partly for visualisation | Primary data will be accessible trough external catalogues using Web coverage services |
| In-situ observation | Data measured by sensors and stored in databases | Meteorological data, quality water measurement | Data are used for analysis (including temporal) and partly for visualisation. Usually visualisation is necessary using graphs etc. | Data from sensors are usually stored in database and then are accessible through services like Sensor Observation Services, or Web Processing Services |
| Terrain measurement | Data collected in the field using specific measurement equipment | GPS measurement, geometrical measurement, photography with positioning etc. | Data are validated and then used for updating of existing data, next analysis or visualisation | Data are accessible from equipment in proprietary formats or using Web services |
| User edited data | Specific data layers, where groups of people could on line edit data and attributes | Point of interests, user collected data about objects, user observation | Data are validated and then used for updating of existing data, next analysis or visualisation | Data are stored in database using system middleware, accessible could be using web services |
| User derived data | Results of analysis of previous types of data | Image classification, extrapolation from sensors | Data are used for reporting and visualisation | Data are stored on server (often temporally) and are accessible using Web services |

Each of the 7 validation pilots relies on trans-regional and trans-European data sharing between pilot settings, within INSPIRE networks present in the project and with collaborating members of the HABITATS user communities.

Pilot actors include stakeholders participating in the management of platforms, data owners and producers and contributors to services and user groups of all pilots. Habitats tabulates the generic use cases from the pilots, as well their user groups and generic roles. From these it identifies that the services based on external data in HABITATS requires the following basic operations:

- Data discovery
- Storage metadata about services on server
- Data visualisation
- Data download

On the base of analysis of user need we identified the following possible services at each of the HABITATS pilots:

1. Management of natural resources
   - Wild Salmon Monitoring (IE)
     1. Online service to allow Fishermen check what fishing they are allowed to do in a particular river.
     2. Or input their rod catch data on their smart phone.
   - La Palma Protected Marine Area (ES)
     1. Online temporal series for each Parameter for general users and expert users.
     2. Eco-tourism
   - Hiking Trip Planner (IT)
     1. Map browsing on Internet and mobile phones.
     2. Define by themselves the composition of maps based on their personal specific interest;
     3. Download waipont, routes, maps from centralized site with official and updated data;
     4. Browse maps with mobile phones or smartphones equipped with gps receivers
     5. Users give feedback on quality of routes.
2. Natural Environment – Augmented Reality (ES)
     1. Visualize the spatial data with 360º images of a outdoor stage via web
     2. Tourist routes – educational.
3. Economic activities
   - Sheep and Goat Herd Management (IT)
     1. services for locating animals;
     2. service to stay in contact with the park administration to assign grazing areas.
   - Economical activity at marine coastal benthic habitats (LV)
     1. Construction Works at the Marine Areas – online information and planning tool.
     2. Fishing Activities – information service on the marine coastal habitats.
4. National policy
   - Czech National Forest Programme (CZ)
     1. Online service on Subsidies in the forest management – important to a lot of people and private growers.
     2. Forest Management (silverculture) procedures to avoid floods – as current tree species do not retain water well.
     3. Mapping of Forest Types – mainstream forest management procedures.

**Habitats architecture**

Based on the pilot scenarios and pilot use cases from the first stage of the HABITATS analysis, we identified common schemes, which can be focused on sharing of data from different pilot localities, but also provides support for pilots with pan-European (and global) data discovery services, with possibilities to provide support for common data.

The HABITATS Networking Architecture has the goal of defining a system able to ensure the interoperability and security of provided data and services. In particular, since an integration with the INSPIRE initiatives is needed, it is based on:

- a methodological approach able to define a system architecture that is scalable and adaptable to the specifications and standards currently being defined;
- the adoption of a Service Oriented Architecture based on WebServices and SOAP technology.

Habitats proposed a view of the system as a set of functional components (or blocks) that communicate through defined interfaces, the system is configured as an architecture made of levels or functional layers:

- data layers – management data and files on storage, eventually guarantee access to external sensors
- Server (engine layer) – define tools, which guarantee basic services on the server side – supplying service
- Client layer – is client side of Web services, which guarantee access of users to services
- Application layer is some form of wrapping elementary client services into application or into such form, which could be used by other Web tools
- Presentation layer contain such web tools, which allow to combine and publish single objects from application level as part of Web presentation

The basic scheme is on next picture:

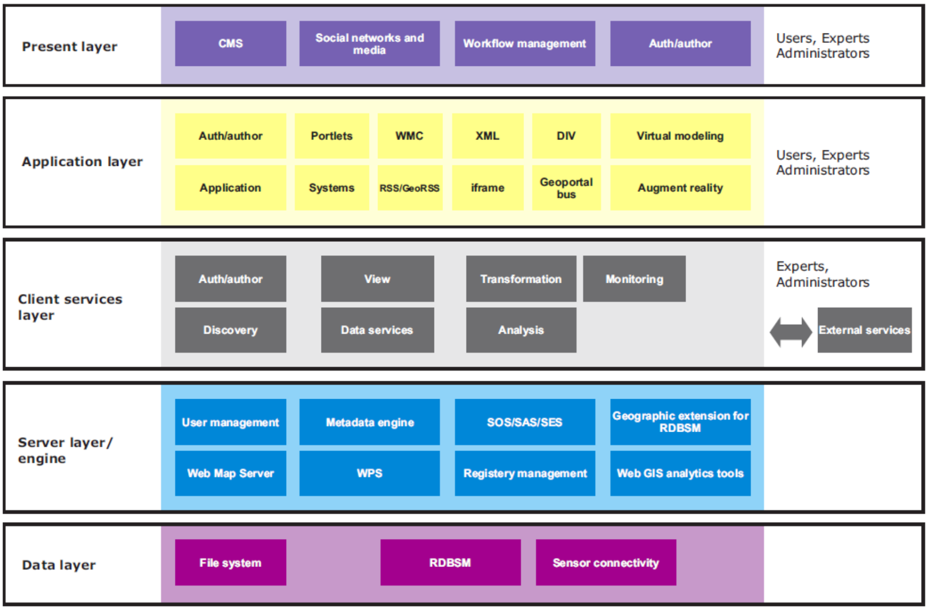



*Figure 6. Habitats Architecture*

These local pilot communities, which are active in their respective languages, together with the HABITATS networking architecture support INSPIRE Network services, where needs represents the fundaments of this concept. INSPIRE networking services are in principle limited only to the management of existing data and metadata. The HABITATS Networking Services also support such functionality with data and metadata management, data and metadata collection, working with non spatial data, etc.

The HABITATS Networking Architecture aims to extend the principles of the INSPIRE networking architecture, because INSPIRE doesn't cover such important aspects such as data management and data collection. So all components of the INSPIRE networking architecture will be included in the HABITATS architecture, but this concept will be extended by other functionalities. From this point of view principles of GEOSS and GMES and also principles of Shared Environmental Information System (SEIS) and Single Information Space in Europe for the Environment (SISE) will influence the HABITATS architecture and its networked services. The HABITATS architecture defines a platform-neutral SDI with a basic set of networking services in compliance with the INSPIRE Directive for sharing environmental data, especially that related to the 4 INSPIRE themes of 16.*Sea-Regions (SR),* 17. *Bio-geographical Regions (BR),* 18. *Habitats and Biotopes* (HB) and 19. *Species Distribution (SD).* This will result in a European Metadata profile for these four data themes, which will be an extension of the INSPIRE profile. Our intention is not only to follow the INSPIRE profile for discovery services, but to also reflect on the extension the profiles for using data; a link to data modeling activities is therefore necessary. This profile is further open to extension by single countries or user groups, but the aim is that it be respected as a minimum set.

HABITATS Service applets re-use existing applications where possible and are themselves designed for re-use. The selection of the specific services to deploy is primarily an user-driven process, as defined in the user scenarios and requirements and as required by the pilot validation platform of task has defined the prototype set of Network Service Applets to be installed in validation pilot platforms, as:

- A series of specific networking service applets deployed and tested for data sharing within the project using the Network Service Architecture (of D4.2.1)
  - Interoperability Services
  - Enabling Services
  - Visualisation of information layers
  - Overlay of information from different sources
  - Spatial and Temporal Analysis
  - "quick" and "light" on-demand applets to meet validation pilot expectations and user needs
  - Usability, simplicity and openness to rapid prototyping mash-ups.
- A set of specific service applets that allow users to identify, access, use and reuse habitats-related data, designed and deployed on-demand to meet user needs,
- Users selected in the T2.3 user scenarios and T5.2 pilot validation platform.
- "Quick Prototyping" service applets respecting the HABITATS service architecture and developed on-demand.

These are based on the outcomes from the earlier tasks and work with the HABITATS RL, and will now lead into the interface tools and toolkit.

The final implementation of the HABITATS Services anticipates that selected concrete services will be deployed for every pilot, and that there will be one central platform (i.e. the Reference Laboratory).

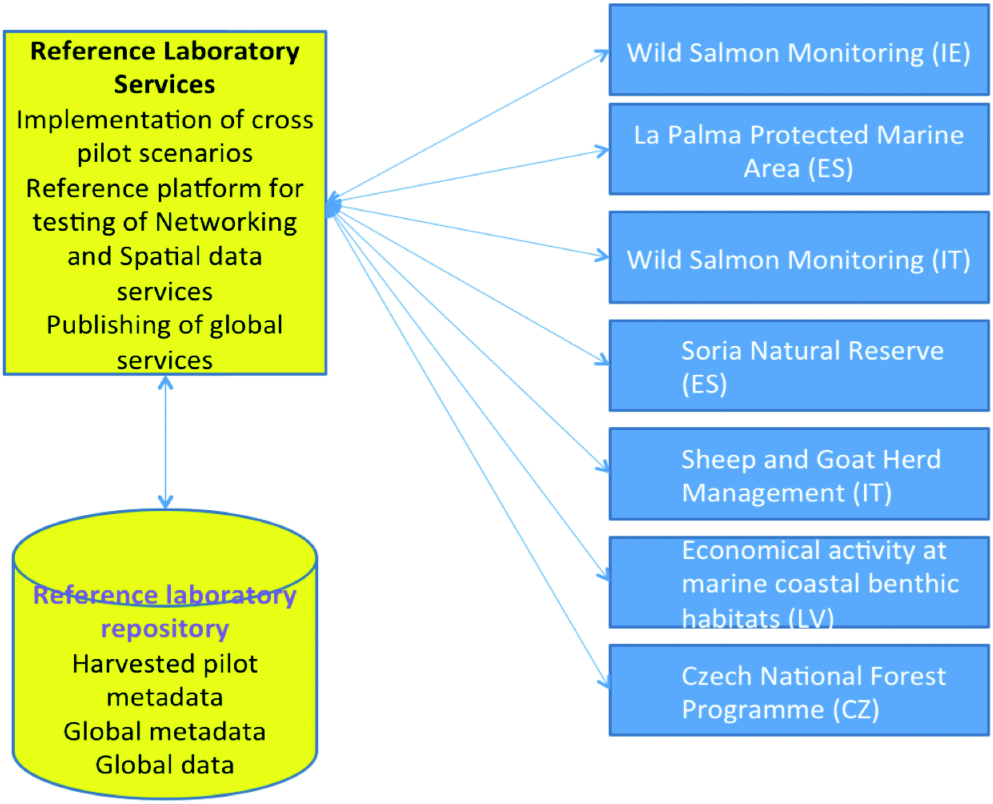



*Figure 7. Connection of Reference Laboratory and Pilots Implementations*

The HABITATS RL provides the Networking Architecture supporting both Networking Services and Spatial Data services, that support the SDI network services to enable trans-European sharing of habitats-related spatial data between public authorities and other stakeholders in the Community, enabling the creation of value added services. The RL is focused on implementation of:

- Cross pilot scenarios based on sharing of data among more pilots
- Validation platform for testing of conformity of implemented pilot services
- Services supporting global discovery, view and downloading
- Repository for common metadata
- Repository for pan European datasets such as Natura 2000, CLC, Urban Atlas and Open Street Maps
- Interlink with social networks

The RL provides the Networking Architecture, that supports the SDI network services to enable trans-European sharing of habitats-related spatial data between public authorities and other stakeholders in the Community enabling the creation of value added services.

The RL enables deployment of specific service applets, including interoperability and enabling services, on-demand from user communities and the pilots for initial implementation and validation. It is being developed further to include an invoking service toolkit integrating the service applets with the goal of facilitating the development of end-user services accessing habitats-related spatial data over time.
We have a platform or common elements on which to grow the different drivers. These elements are intended to facilitate the tasks of development and enable interoperability and services required by INSPIRE.
The first set of HABITATS Networking services have been implemented on the HABITATS Reference Laboratory (RL) geoportal platform. This acts as a client of the seven HABITATS pilots and provides a very rich set of cross-pilot, inter-regional and enabling services and tools that will be validated by users on the basis of concrete implementations in phase 2 of the project. In addition, "quick prototyping" applets for specific applications will be developed in response to user requests at each pilot during phase 2 of the project.

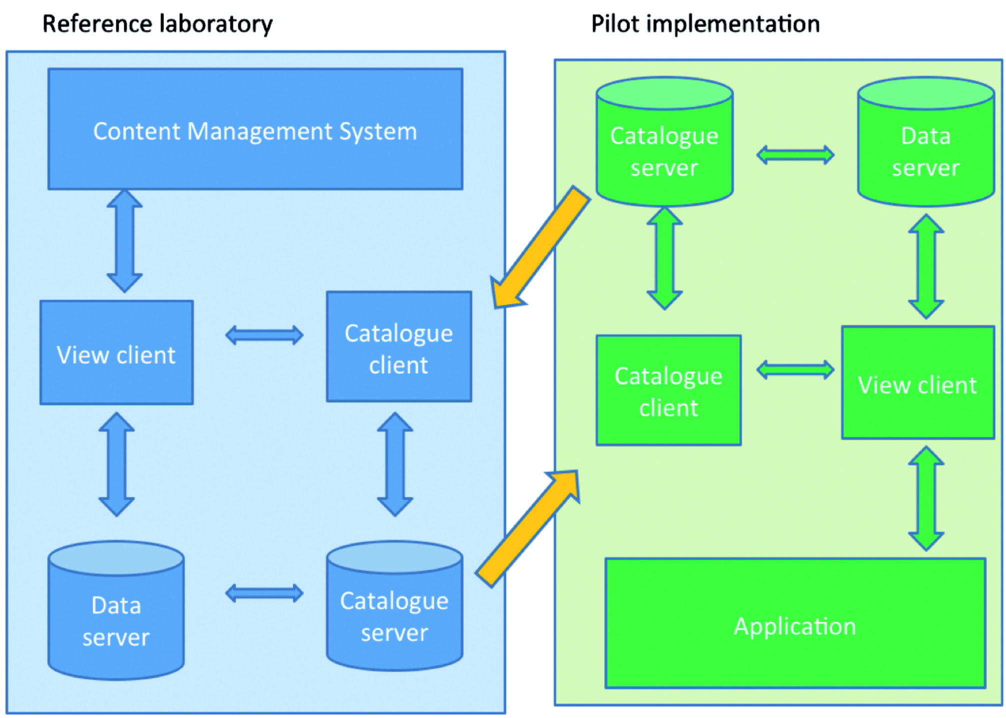



*Figure 8. Connectivity of RL catalogue and pilot catalogue*

Pilot implementation will implement only functionality required by pilot user needs; Reference Laboratory will implement full functionality.
This figure illustrates that the more linked is the prospective SDI evolution to actual end user needs and requirements, the more crucial become the modalities adopted for the collection of

technical feedback from these users, not only during the initial design stage, but all along the development and possibly also the validation of the various service prototypes.
In general, what the HABITATS pilot descriptions strongly support is a distinction between two kinds of end user – and feedback thereof:

- Individual members of the local stakeholder communities (e.g. public authorities, universities, domain specific research centers, business associations etc.). These are normally available on a continual basis to offer their acknowledgement and support for the ongoing deployment of INSPIRE compliant SDI services. Thus, they can be easily associated to all stages of design, development and validation. However, their technical contribution is usually weak, politically driven, or scanty and superficial, with respect to the specific features of the proposed solution. Moreover, they must be "forced" to contribute on a voluntary basis (as the Latvian pilot has stated) and there is no technical solution that can ensure this to happen.
- Actual end users of the system (e.g. tourists, shepherds, anglers, fishermen, etc.). Contrary to the previous group, they are deeply involved into system usage, thus their technical input is more precise, and often essential in bugs tracking and usability or functional improvement. In general though, it is quite hard to collect the experience of these people right in the time and place they were using the system. There may be solutions, like the one proposed in the Campo del Moro pilot, allowing to "jump" into an evaluation system right after the end of a working session, yet once more it cannot be taken for granted that all relevant feedback would be saved in that way. Furthermore, the contribution of this category of end user is mostly concentrated in the post-deployment stage, leaving the previous ones largely unattended[14].

One added-value of HABITATS is precisely the emphasis it has palced in *the possibility of exploiting stakeholder feedback during the development stage of the different ICT developments*. This lies among the main goals of the PaESI platform deployment and its coupling in support to the technological aspects of the project, and in particular with regards to the Reference Laboratory. PaESI has also the advantage of not placing any burden to prospective users of a system in terms of INSPIRE related knowledge. All key evaluative questions are proposed in natural language, and the respondents do not immediately perceive their association with the underlying technical development options.
However, the same process can be repeated – though less efficiently – also by offline discussions and surveys, followed by internal briefings and operational meetings between developers and strategic decision makers.

**Conclusion**

The shaping of future challenges emerges from a clear involvement with key stakeholders who co-design the future technologies based on the intersection among their needs, the ICT requirements in light of availability, usefulness as well as their potential. For this to happen

[14] Of course, the communities of reference for these end users – if existing – may act in support of the design and deployment phases. In which case, they become local stakeholder associations and their inputs can be classified as belonging to the first cluster.

there is need of an enabling environment which can only be generated if based on the identification of common interests and trust on the derived benefits from the process. In this context, HABITATS' strategic vision has been transformed into operational technological developments fit-to-purpose within a bottom-up approach.
Relative little attention has been devoted to the social validation of potentially effective interventions/programmes and projects. Thus it is somehow difficult to identify, to generate and implement evidence-based practices that have resulted in positive outcomes to people and communities. In this context, the methodologies and programming used are often inadequate for the assessment of impact. Thus this chapter is to be seen as a contribution in future policy agendas as it highlights the benefits of ICT in support to rural development, the strategic nature of ICT for sectorial planning as well as the needs for a more responsive for a research agenda to the needs of stakeholders. At large , this chapter underscores different aspects of knowledge in real-case scenarios as a result of the implementation of the Living Lab Approach. The emergence of innovation through mechanisms, approaches and ICT technological developments have been oriented to the formulation of recommendations for policy-makers at the European Commission,region-wide and at the international level.

It is in this context that HABITATS is coupling social content in terms of social validadtion innovative methodologies with highly relevant technological innovatives developments where stirred by *inclusion, participation and technological enhanced learning, lessons learned and knowledge sharing using the methodological approach of Living Lab*.

**References**

HALL, A., R. Sulaiman, N. CLARK, B. YOGANANd: From Measuring Impact to Learning Institutional Lessons: An Innovation Systems Perspective on Improving the Management of International Agricultural Research. Agricultural Systems. 78, 2 (2003), pp 213-241.

RAJALAHTI, R., W. JANSSEN, E. PEHU: Agricultural Innovation Systems: From Diagnosis Toward Operational Practices. Agriculture and Rural Development Discussion Paper 38. World Bank. New York. USA (2008)

PEREZ-TREJO, F. L. DIAS PEREIRA, J. GARCIA GUZMAN, M. NAVARRO DE LA CRUZ. LIVING LABS AND RURAL DEVELOPMENT:TOWARDS A POLICY AGENDA. eJOV Executive The Electronic Journal for Virtual Organizations and Networks. Volume 11, 2009

*SAEZ, A.,J. MARSH.* So*cial Validation of INSPIRE in the Madonie Park, Sicily. E-Challenges 2011. Florence, October 2011.*

SAEZ. A., F. PEREZ-TREJO. Implications for the Individual, Enterprises and Society. New approaches towards innovation and user-driven open innovation methodologies. In the context of development issues. Bled eConference eTrust .June, 2010; Bled, Slovenia.

SAEZ.A. F. MOLINARI.. Millennium Development Goals [MDGs] : inextricably linkages to policies, programmes and projects that bring about socially inclusive and equitable sustainable development. International Conference on Social Spaces for Innovation. Food and Agriculture Organization (FAO). Rome, Italy. June 8, 2011

SAEZ, A., G. URQUIA,, M. NAVARRO, A. DE MIGUEL. A Contribution to advancing the vision of GEOSS and INSPIRE: The social validation of initiatives. GEOSS Conference. Madrid, Spain. January 2012.

# WWW Technologies and Regional Development

Zdeněk Havlíček, Petr Benda, Václav Lohr, Šilerová Edita
*DIT FEM CULS Prague*
E-mail: havlicek@pef.czu.cz

**Introduction**

Web technology continues to develop and bring new opportunities that the young generation takes for granted. Facebook and YouTube are typical examples. Users increasingly create the web contents that they share with the others in a specifically targeted communities and social networks.

The primary reasons for the popularity of these technologies are listed as follows [3]:

- Free or cheap to produce
- Do not require proprietary software to be installed
- Do not require specialist computer skills
- Content can be updated and disseminated easily

Keyword-based tags can be used to classify the submitted content and to conduct searches for retrieval [9].

It is possible to start and operate one‘s own simple web site on different servers, eg Webzdarma. The more complex web sites use the Web Content Management Systems (WCMS). Some are the Open Source Software (OSS), distributed under the GNU / GPL licence, others are usually part of larger software units and adapted to the needs of customers for a fee.

Technological progress in the web area should be seen primarily from the point of view of a typical user. We can notice it, for example, in the advancement of maps applications. The dynamic maps ("Live Maps"), which have recently appeared, allow viewing any territory in different scales. Interactive mapping applications are created with the use of map data, which are stored in a database. Such applications called as RIA (Rich Internet Applications) or web 2.0 applications can be used not only on personal computers but also on other types of devices, such as tablets (iPAD), and Smart Phones.

New web applications must reflect the current state of technology and respond to end-users, so that they can find additional value in these applications. With the increasing power of computer stations the distinction between the client side and the server side becomes blurred. The original architecture model of the client-server technology was based on the usage of powerful server machines with large processing power capacity, operation and disc memory. The requirement of higher performance servers is secured through their virtualization, deployment of grid technology and cloud computing. Most server side scripting languages are developed as platform independent, although some are closely tied to the platform.

This chapter is based on two previous papers [4 and 5] the text is little modified, but the images are different. We would like to emphasize scale of using new ICT technologies for regional development. The chapter is divided in two parts Web Sites Creations and Web Applications Development.

**Web Sites Creations**

Objectives and methods

This part aims to analyse the use of freely available technologies that are used for creating web sites. Two OSS WCMS Systems, WordPress and PrestaShop have been selected from a wide range. With these systems, users without specific IT knowledge have created and published on their websites new presentations or simple e-commerce applications. Practical examples are used to demonstrate the creation of presentations and e-shops, which can contribute to regional development. The paper is based on the IT in agritourism research results that have been previously published [6 and 7].

Results

Different options for creating web sites have been tested in the Master's course "ICT for Managers" in the academic year of 2010/2011. The environment for storing and managing dynamic Web sites has been produced on dedicated server kitlab.pef.czu.cz. The overall design was based on the products of the OSS (Open Source Software) type. In this case the Linux operating system and the Apache Web Server were used. Web sites were created using the WCMS WordPress and PrestaShop. The two softwares are based on the same platform of the PHP scripting language and MySQL database system.

All created web sites are running at the kitlab.pef.czu.cz server. The web sites have been developed for the Czech users and the screen shots presented in the paper have been obtained from the Automatic Translation in Google Chrome.

Distance and full-time students in groups of 3 to 5 have created the web sites. Comprehensive data on the number of projects (created Web sites) are outlined in Table 1.

| Used software | Description of the technology | Type of study | Number of projects / number of students in semesters | Links |
|---|---|---|---|---|
| WCMS WordPress | System for creating and managing (blog) web sites (blogs) | Distance Course ETE77E | Winter Semester 37 / 154 Summer Semester 26 / 115 | Figure 2,3,4 |
| WCMS PrestaShop | System for creating and managing e-shop solutions | Full-time Course ETE41E | Winter Semester 70 / 336 Summer Semester 54/ 231 | Figure 5,6,7 |

*Table 1. Projects overview*

Web sites for micro regions and agricultural farms

Distance students in the “ICT for managers - ETE77E” course used WCMS WordPress and created a total of 63 group projects on the topic of "ICT and Regional Development - the region of XY“ in accordance with the recommended Website Project Assignment - Table 2.

| Task | Type of task | Task description |
|---|---|---|
| Analysis by the questionnaire | Individual | In a selected region assess the situation in the Small and Medium Enterprises (SME) using a questionnaire entitled "ICT and Regional Development“. At least five completed questionnaires is recommended |
| Web site creation | Group | Design a "better" web site for a selected object (eg a farm) Develop a new, modern website using the WCMS WordPress. This is a student project, personal data are not to be published. |
| Documentation | Group | For the region to process documents according to the recommended outline and add to the wiki on kitlab.pef.czu.cz |

*Table 2. Website Project Assignment*

The projects' aim was to create a modern web presentation of an agritourism farm in a selected region. The focus of the presentation had to support local regional development objectives and be consistent with the Regional Development Strategy. Another objective was to increase the attractiveness of a selected region as a tourist destination by using its specific potential, especially the natural and cultural heritage, as well as local products and services.

When developing the projects, emphasis was placed not only on the web site content as such, but also on the use of technologies that are common in professional web applications, such as microformats and mashups:

- Microformats: In the projects the students have used hCalendar (a simple, open, distributed calendaring and events format), hReview (an open standard for distributed reviews) and hCard (a simple, open, distributed format for representing people, companies, organizations, and places).
- Technology mashups using different APIs synchronize the information content from multiple sources (see Figure 1). Students used this particular technology to create map mashups. The Web sites contain maps which show places of interest, sightseeing areas, tourist trails, community news related to GPS data, etc.

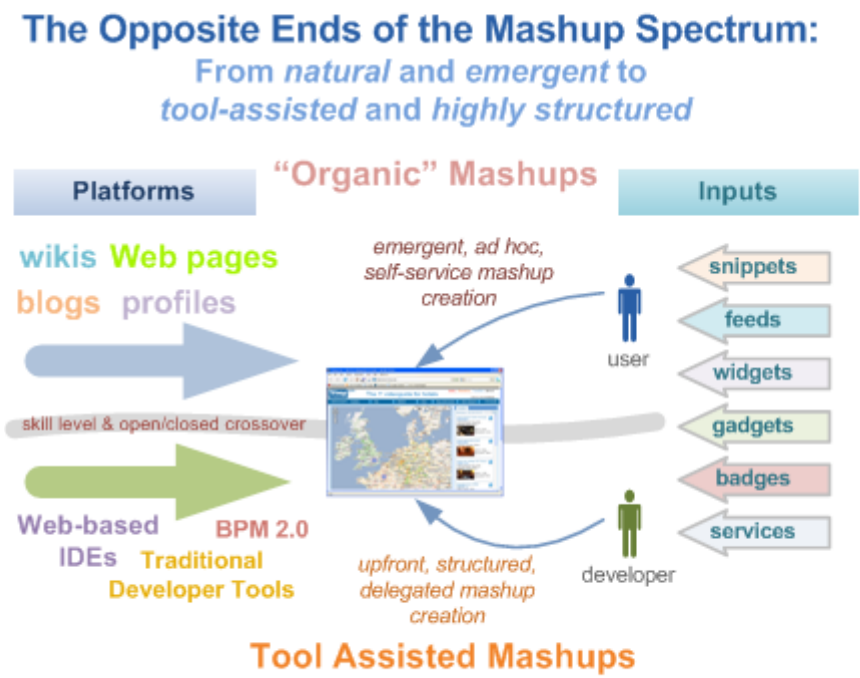



*Figure 1. Mashup schema [8]*

<u>WCMS WordPress</u>

Web Content Management System (WCMS) is prepared software for creating and managing web applications (web sites). Its main capabilities are editable content, document management and user management. There are many different types and forms of WCMS. One of most popular WCMS is WordPress [12]. This software is free under the GNU General Public License. WordPress allows the use of different templates and supports differentiation in user viewpoints. With WordPress is easy to create simple web applications by non-IT specialists.

We have prepared for the students a video support for basic work with Wordpress (installation on server, choice of template etc.). The students can easily create useful web sites with new and editable content. Examples of their work are shown in Figures 2, 3 and 4.

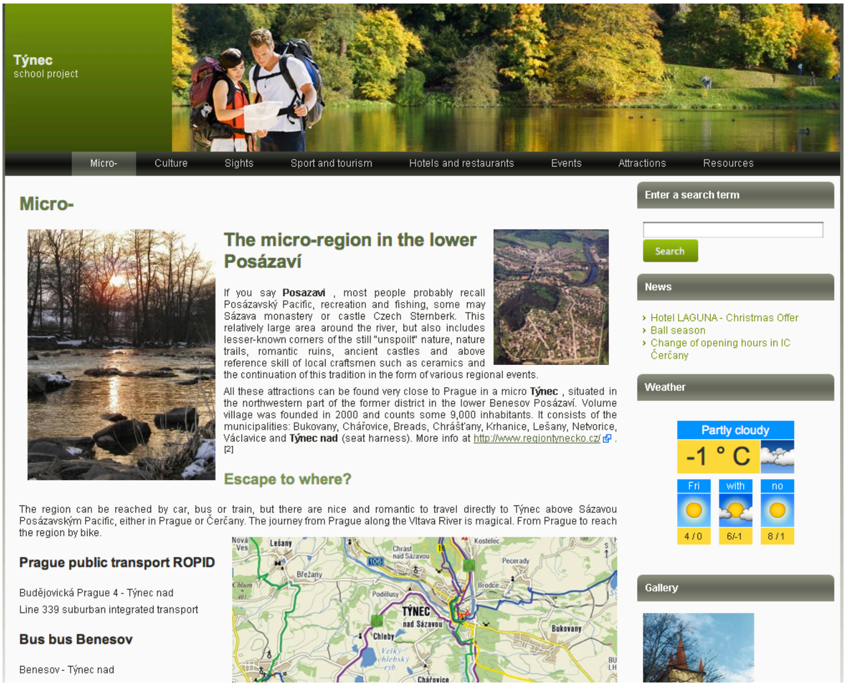



*Figure 2. Micro–region Posazavi*

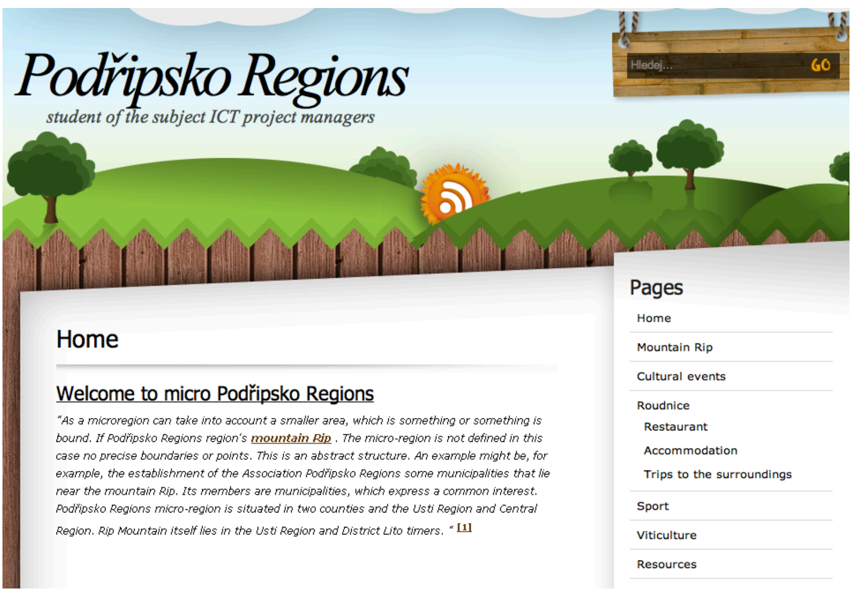



*Figure 3. Micro-region Podripsko*

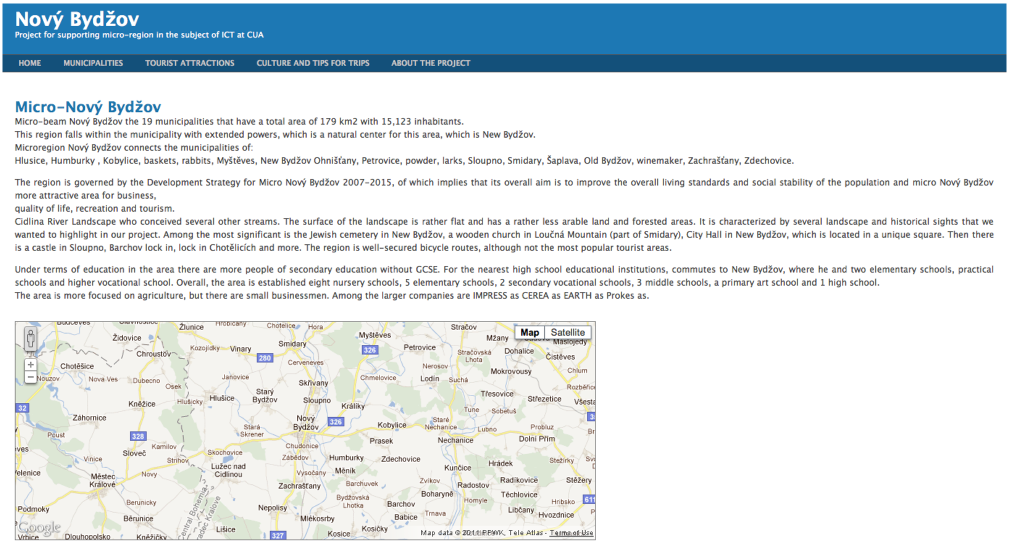



*Figure 4. Micro-region Novy Bydzov*

The projects were developed in close cooperation with the farm owners. Existing Web sites were updated both in terms of overall design and in terms of content and visual form. Some projects will be directly implemented in practice.

<u>e-shops supporting regional development</u>

Full-time students have created e-shop solutions for small and medium enterprises. In this case, the students used microformats and mashups technology as well. The students tried to implement the hProduct microformats so as to improve the position of their Web sites in search engines. Mashup technology was used to synchronize messages, news and articles from Facebook, Twitter, etc. In some projects YahooPipes have also been successfully deployed.

<u>WCMS - e-commerce software PrestaShop</u>

PrestaShop is a free e-commerce system, which is currently very popular. More than 50 000 e-shops in the world are based on this system. The basic kernel is available under Open Software License. More sophisticated modules as payment modules are offered commercially. PrestaShop is useful for creating small to medium-sized e-shops.

PrestaShop is a complex system, with functions for the most requests of the online commerce, including options for discounts on certain products. Furthermore, the latest products are displayed on the first page of the site and the platform offers many other opportunities to promote products.

IT specialists collaborating with the students prepared an initial installation on the kitlab.pef.czu.cz server. A short video on the usage of the system was available to the students in the Moodle Learning Management System. Many students worked with the ftp protocol for the first time when uploading some files to the server.

After the installation the students were able to enter the administration panel to change the site settings and create e-shop contents. The administration panel area is very well organized and easy to use. Anyone has access from it to any part of the site that can be customized. Adding products, currency settings and transport costs are just some of the features offered by PrestaShop to its users.

Three examples are presented in Figures 5, 6 and 7.

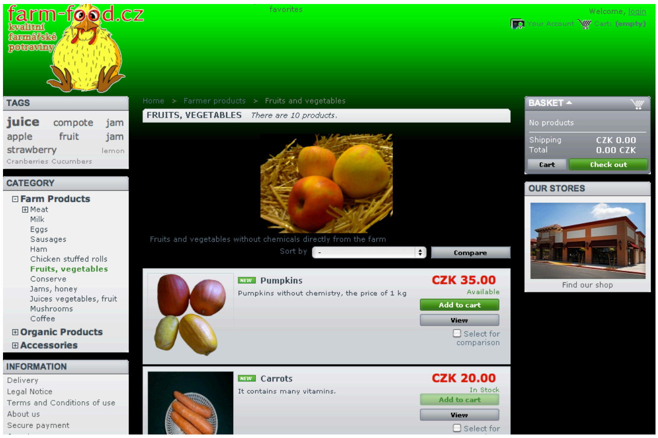



*Figure 5. „Farm Food“ e-shop*


*Figure 6. "Garden Buildings" e-shop*


*Figure 7. "Wood Toys" e-shop*

With WCMS it is relatively easy to create new web sites and manage their contents, including the support for the connection to social networks. Both methods used in the creation of web sites are reviewed in Table 3. The average length of time spent on the web sites created in WordPress was about 40 hours, while it took about 60 hours to create an e-shop solution in PrestaShop. The number of hours is directly dependent on the total number of offered products or services. An e-shop requires the longest time for making product images and descriptions. The created projects have always been designed as separate modules without any direct connection to the enterprise information system.

| WordPress Advantages | WordPress Disadvantages |
| --- | --- |
| • Universal system for the creation of smaller web sites<br>• Wide community supports the overall development of the system<br>• It enables applications in Czech (Czech guides and tutorials)<br>• A large number of templates and graphics expansion modules is available free of charge | • It requires IT support and training for installation on the server |
| PrestaShop advantages | PrestaShop Disadvantages |
| • Universal solutions for e-shop<br>• Modern technology - Drag and drop<br>• Language options, a large number of templates. | • Quite an extensive system, requires training and IT support<br>• Some specific modules (eg, payment by credit card) must be bought. |

*Table 3. WCMS comparison*

**Conclusion**

Rural development should be supported by Web sites with a nice design and current content. These requirements are relatively easy to achieve through the usage of WordPress or PrestaShop WCMS. The solutions used for the web sites described in this study were low cost and could be managed without special IT knowledge. Nevertheless, minimal IT support is desirable.

**Web Applications Development**

Objectives and methods

This part analyses technologies that are used for creating web applications. New approaches, which are known as mashups, are utilized for map applications by incorporating data from different sources. Practical examples show the possibilities for creating applications that can contribute to regional development. Synthetic evaluation of used methods is described in the Conclusions.

Results

In the academic year of 2010-2011 two approaches for web applications development had been selected for the courses of the Internet Technologies (Informatics Study Program). In the first case eleven applications focused on a specific region were created. In the second case nein applications were created on the basis of differences in regional types of food. All created applications are running at the server kitlab.pef.czu.cz. The applications have been developed for the Czech users and the screen shots presented in the paper were obtained using the Automatic Translation in Google Chrome.

<u>Web applications for the Jicin Region</u>

The main objective in the Internet Technologies (ETE32E) course at CULS Prague was to create different applications for the promotion and development of the Jicin Region. The MSc students attending this course visited the Jicin Region and afterwards they chose topics for their groups‘ projects. Eleven projects have been carried out.

Steps in the projects development:

- Data collection
- Design and layout
- Static web site (XHTML and CSS)
- Testing for usability and accessibility
- Implementation of mashups applications using the PHP language and maps
- Creation of the administration interface
- Documentation for the entire project in the Wiki

Two examples of new regional portals are presented in Figures 8 and 9.

The students created the project documentation in wiki technology on the server http://kitlab.pef.czu.cz/wiki/. Example of this wiki documentation is shown in Figure 10.

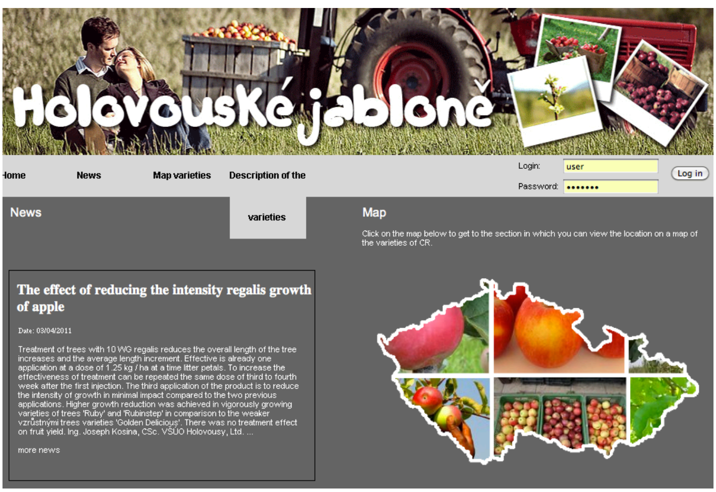



*Figure 8. Portal for Research and Breeding Institute of Pomology Holovousy*


*Figure 9. Portal for regional brewery „Novopacké pivo"*

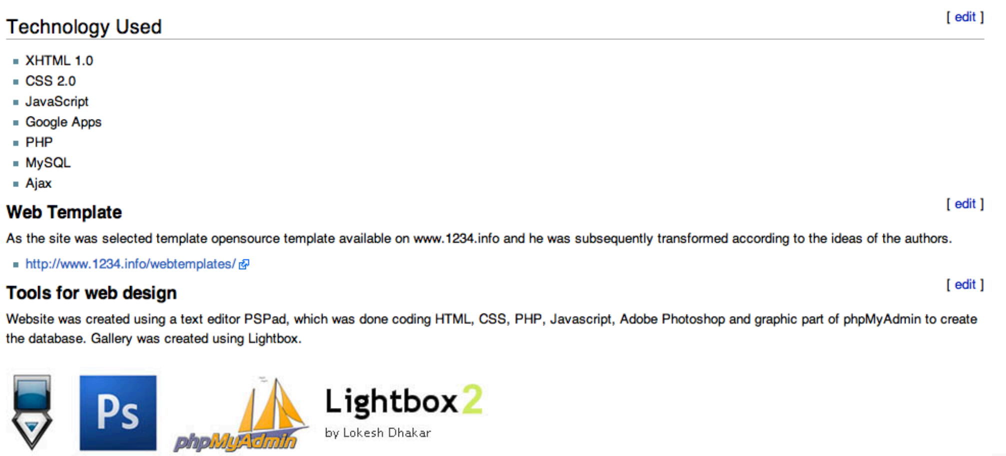



*Figure 10. Example of documentation in wiki*

External people from the Jicin Region evaluated the created web applications. They were very satisfied and, as a result, some projects will be implemented in real conditions.

Web applications - Regional Food

The Czech Ministry of Agriculture launched a campaign for the promotion of regional foods in 2010. Regional foods are specific carriers of cultural values and they are tied to specific areas. They are often the main asset and symbol of the region (e.g. the Trebon Region carp).

Students in the Internet Technologies (ETE89E) distance course created web applications concerning the regional food in 9 different regions. The general project aim was to create a dynamic web site that would inform about the most important regional foods of a given region.

Basic information that is available to site visitors:

- List of regional food from a given region
- Name of regional foods
- Detailed description of the food
- Food photos (if available from the information source)
- Food producers
- Address of the manufacturer
- Summary description of the activities of a producer
- Location of manufacturer of food on a digital map

The manufacturer and retailer will be using maps to locate shops. To promote interest of visitors, the website offers them an opportunity to upload their favourite food recipes and other activities. The website is also designed to add new features such as ratings, comments, etc.

Frameworks supporting the architecture MVC (Model-View-Controller) have been used for the server-side application development. PHP framework Nette was used in the main. Nette framework is a powerful framework for rapid and easy creation of high quality and innovative web applications in PHP 5. It eliminates security risks, supports AJAX, SEO, DRY, KISS, MVC and code reusability. Nette Framework is of Czech origin and has a rapidly growing user community.

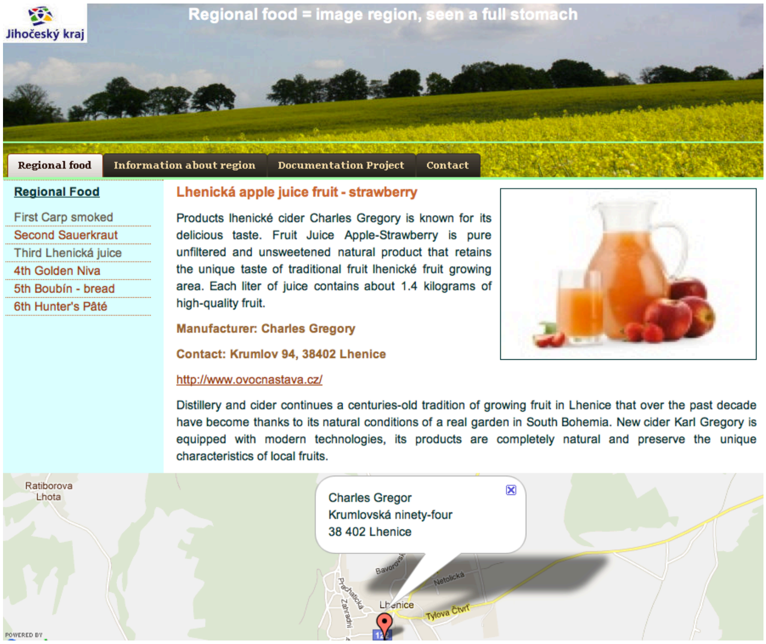



*Figure 11. Regional food – Example for South Bohemia Region*

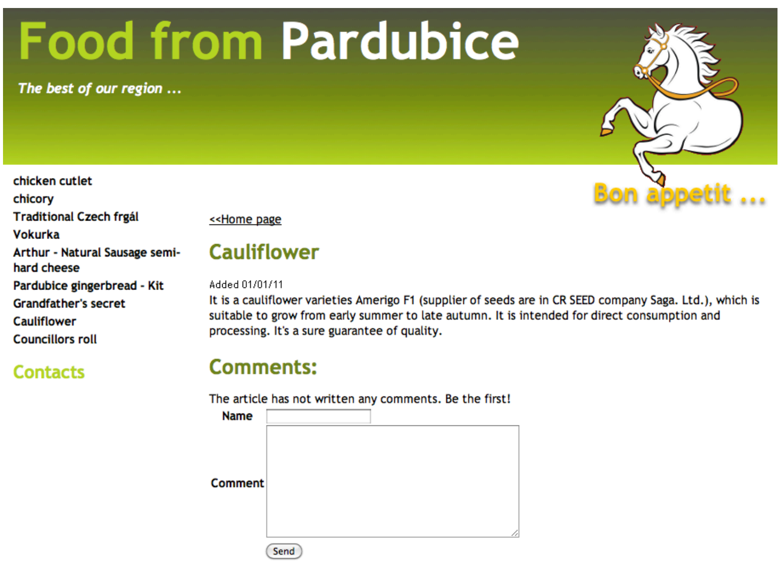



*Figure 12. Regional food – Example for Pardubice*

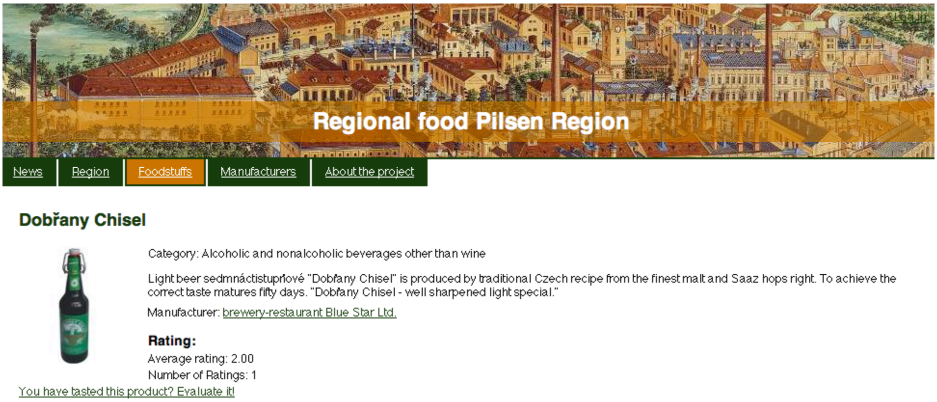



*Figure 13. Regional food – Example for Pilsen Region*


*Figure 14. Regional food - Example from Liberec Region*

Mashups can be customized for the characterisation and description of the regions which have been used in the students website projects. Mashups combine Web 2.0 and Web services technology to facilitate data access and visualization on the Web. Furthermore, the use of mashups provides a cheap and effective platform for displaying real-time data. [2]

Mashup applications therefore bring new possibilities for users. Their main advantage is in an interactive combination of many information sources into one application or its part. The most common usage of mashups is in the news sphere, social networks, and/or in mapping information, which are listed on the examples above.

People collect a huge amount of data about things and activities, all of which are need to be marked with their locations. Mashups success is due to the arrival of Google Maps API. This started new approaches to developing specialised websites and allowed Web developers to mash all kinds of data onto maps. Other softwares, such as APIs from Microsoft (Virtual Earth), Yahoo (Yahoo Maps), and AOL (MapQuest) followed [1, 10].

Another advantage of mashups is in the fact that users can share information, add value to Web applications by using them or aggregate data from different sources creating Web applications using specialized tools (mashup tools) [11].

## Conclusion

Web applications that have been created for the Jicin Region deliver a new "modern" view of the use of Web technologies for regional development. These applications – Portals were well appreciated by the Jicin Region people

Different approaches to creating dynamic web sites have been tested in the "Regional Foods" applications. Nette framework was found to be the best tool for solution of the project.

## Acknowledgements

This paper was elaborated within the framework of the solution VZ MSM 6046070906 "Economics sources of Czech agriculture and their efficient use in the context of multifunctional agri-food systems".

## References

ARDIELLI, Jiří; *et al.* Modern tools for development of interactive web map applications for visualization spatial data on the internet. In Acta Montanistica Slovaca [online]. Special Issue 1/2009. Košice: Berg Fac Technical Univ Košice. 2009 [cit. 2010-08-19]. <http://actamont.tuke.sk/index.html>. ISSN 1335-1788.

DUANE , Merrill. Mashups: The new breed of Web app. In IBM Developer Works [online]. IBM USA : IBM Corporation, 08 Aug 2006 [cit. 2011-04-19]. Dostupné z WWW: <http://www.ibm.com/developerworks/xml/library/x-mashups/index.html>.

Goodfellow, T.; Graham, S.: The blog as a high-impact institutional communication tool. 2007, Sydney, Biennial Conference of the Australian-Library-and-Information-Association, p. 395-400. ISSN: 0264-0473.

Havlíček, Z.; Benda, P.; Lohr, V.; Šilerová, E.: Web Sites and Regional Development. Proceedings ISAF 2001 p. 104 -112. ISBN 978-80-213-2190-8

Havlíček, Z.; Lohr, V.; Benda, P.; Šilerová, E.: WWW Applications and Regional Development. Proceedings EFITA 2011 p. 403 – 409. ISBN 978-80-904830-0-2

Havlíček, Zdeněk; Benda, Petr; Lohr, Václav. Agritourism Web Presentation Evaluation. 2009, Praha: Agrarian Perspectives XVIII, PEF CULS Prague, p.727 -731. ISBN 978-80-213-1965-3.

Havlíček, Zdeněk; Benda, Petr; Lohr, Václav; ICT Support for Agritourism. 2010, Praha: Agrarian Perspectives XIX, PEF CULS Prague, p. 327-334. ISBN 978-80-213-2123-6

Hinchcliffe, D. Mashups. [online] The next major new software development model? [cit. 2011-04-10]. URL <http://www.zdnet.com/blog/hinchcliffe/mashups-the-next-major-new-software-development-model/106>

Hope, G.; Wang, T.G.; Barkataki, S.: Convergence of web 2.0 and semantic web: A semantic tagging and searching system for creating and searching blogs. 2007, Northridge: ICSC 2007: International Conference on Semantic Computing, Proceedings, p. 201-208. ISBN: 978-0-7695-2997-4.

MEDITSKOS, G.; BASSILIADES, N. A combinatory framework of Web 2.0 mashup tools, OWL-S and UDDI. EXPERT SYSTEMS WITH APPLICATIONS. JUN 2011, 6, p. 6657-6668. ISSN 0957-4174.

Poorazizi, M. E.; Alesheikh, A. A. Monitoring real-time environmental information using Web 2.0 and GIServices technology. INTERNATIONAL JOURNAL OF CIVIL ENGINEERING. MAR 2011, 9, p. 63-70. ISSN 1735-0522.

WordPress [online], WordPress.org. 2010 [cit. 2011-04-12]. WordPress - about. URL <http://wordpress.org/about>.

# Section III
# Where we go?

# What are the priorities for future research?

Karel Charvat
*Czech Centre for Science and Society*
E-mail: charvat@ccss.cz

**Introduction**

What will be the RTD priority for next period? What needs to be done in the application domain? What are the future technological solutions?

This chapter is formed by introduction analysis defining future research priorities, which are based on the Cologne Declaration, but also on former recommendations coming from other projects including ami@netfood, Rural Inclusion, COIN IP, Future Farm, NaturNet Plus. This chapter also includes the following list of topics representing examples of research results:

- Transparent agrifood interoperability,
- A business intelligence approach,
- Exploring the dynamics of using ICT for farming purposes,
- Data quality concepts and methods,
- The central role of use cases in enhancing data exchange and interoperability in agriculture,
- Multi-sensor data acquisition, storage, handling and analysis,
- Monitoring in centrifugal irrigation pumps ,
- VlitNod technology
- Wireless sensor networks geovisualisation,
- Server-side solution for sensor data,
- Self adaptive algorithms for analysing sensors data
- Optimally distributing a fleet of robot,
- Using linear programming for tactical planning in agriculture.

In accordance with the Cologne Declaration, we divide the research priorities in two groups:

- Application Research domain for Agriculture, Food, Rural development and Environment,
- ICT technologies for Agriculture, Food, Rural development and Environment.

**Application Research domain for Agriculture, Food, Rural development and Environment**

Currently, we recognise the following priorities for research domain for Agriculture, Food, Rural development and Environment:

- Collaborative environments and trusted sharing of knowledge and supporting innovations in agri-food and rural areas, especially supporting food quality and security - The concept of the trust centres has to represent an integrated approach to guarantee the security aspects for all participants in the future farm. There will be a growing importance of protection of privacy and IPR. Trust of information is one from the priorities for all rural communities. Pan European social networks have to support trust centres and enable such technologies as cloud applications and which will have to guarantee knowledge security. [1]
- ICT applications for the complete traceability of production, products and services throughout a networked value chain including logistics - to develop world-class network management solutions that facilitate communication and co-operation between networks of SMEs and large enterprises in the agri-food and rural development domains. These solutions will enable the management of food supply chains/networks, virtual and extended enterprises through collaboration and knowledge exchange. [2]
- New generation of applications supporting better and more effective management of agriculture production and decision making in agriculture - Future farm knowledge management systems have to support not only direct profitability of farms or environment protection, but also activities of individuals and groups allowing effective collaboration among groups in agri-food industry, consumers and wider communities, especially in rural domain. Having these considerations in mind, the proposed vision lays the foundation for meeting ambitious but achievable operational objectives that will definitively contribute to fulfil identified needs in the long run. [1]
- **ICT applications supporting the management of natural resources** - With better understanding of the environmental relations, the necessary valuation of ecological performances will become possible. Pilot projects and best practice samples will be the key to demonstrate for a wide auditorium the benefits of environmental caretaking. New model of payment of the different groups of beneficiaries have to be worked out (local, regional, national, continental and worldwide) as well as the best practice between today's "government owned" environment or "private owned with social responsibilities" has to be worked out [3]
- **ICT application supporting rural development and local businesses** - Rural businesses are usually small or medium size businesses according to the number of people they employ. Therefore, knowledge management and internal processes are different from large companies. Future knowledge systems have to be based on each community's own concepts of value, cultural heritage and a local vision of a preferred future. The objective is to develop human-centred reference models of sustainable rural life-styles that overcome social divisions and exclusion and include unique rural

features and create new rural businesses and social infrastructures and attractive computer-based education.

- **ICT application for education and awareness –** Agriculture will require highly educated staff. There will be large shift from manual work to knowledge management. It will be necessary to provide effective knowledge transfer to as many people as possible, through a range of services and to meet the diverse knowledge and information needs of our customers and stakeholders and incorporating management practices and technologies on the home farm, supervised project work and discussion groups: linkages with higher level education institutions.

- **ICT applications reducing administrative burdens in rural areas** - Future ICT applications have to reduce the administrative burden of enterprises and citizens in rural areas by reducing the information elicitation process of businesses when they want to use a particular instance of some public service, or making more effective use of the resources. It has to include, adapt and deploy a web infrastructure combining semantic services with a collaborative training and networking approach, in the rural setting. Furthermore it should include e-Government services that regional public authorities already offer and support them by a rigorous and reusable service process analysis and modelling, and then deploy a semantic service that facilitates the disambiguation of the small businesses needs and requirements when trying to use the particular services. At the same time, the semantic service is complemented by a number of other web-based services that support the creation of communities of learning and practice in rural settings, thus facilitating the communication between the rural businesses with the regional public authorities**.** [4]

This list cannot be complete, but we would like use this initial list as a basis for further discussion.

**ICT technologies for Agriculture, Food, Rural development and Environment**

The development of knowledge-based systems for the farming sector has to be supported by ICT focused on:

- **Future Internet and Internet of things including sensor technology, cloud computing and machine to machine communication** - Future Internet architectures must reflect the needs and specificity of rural communities. It has to be resilient, trustworthy and energy-efficient and designed to support open access and increasing heterogeneity of end-points. Networks should sustain a large number of devices, many orders of magnitude higher than the current Internet, handle large irregular information flows and be compatible with ultra high capacity, end-to-end connectivity.

- **Service Oriented Architecture** as a key element of architectures for future knowledge management systems. Service Oriented Architectures have to provide methods for systems development and integration where systems group functionality around business processes and package these as interoperable services.

- **Methods of knowledge management** including aspects of interoperability. It is important to support the development of machine-readable legislation, guidelines and standards to integrate management information systems with policy tools. Major priorities for future knowledge systems will be the integration and orchestration among services based on semantic integration of collaborative activities, including semantic compatibility for information and services, as well as ontologies for collaboration.
- **Management and accessibility of geospatial information** as a key information source for any decision. Geospatial information includes not only digital maps, ortophotos or satellite imagery, but also location services and sensors. Importance of geospatial information will become more and more important and its amount will rapidly grow with new sensors technologies. Effective methods of management, accessibility and analysis will be important.
- **Open Source development** - Open Source development can bring a lot of advantages not only for research, but also for commercial community. Open source can help to users (private and public) and for all IT sector. The future will give more opportunity for open solutions and also for smaller flexible companies, which will be able to adopt their behaviour and react on new situation. Other important issue is large international cooperation of small companies and not only inside of one country.
- **New modelling methods –** new modelling methods will become in future more important. Their importance will be not only in application like precission farming, but also iin long time planning and decision making. It will require growing computing capacity to introduce methods including linear programming or theory of game.
- **The power of social networks and social media** - The future development of technology has to be based on a broader utilisation of social networks and social media.

## References


Cologne Declaration, http://www.wirelessinfo.cz/?menuID=1&articleID=5&action=article&presenter=ArticleDetail

Dr. Bjarki A Brynjarsson, Dr. Dragan Stokic, Harald Sundmaeker, Jesús de Juan. DRAFT Strategic Research Agenda. Collaborative Work Environments – ICTs supporting Agri-food businesses and Rural development

Charvat K. at all D1.1.2, Analysis of external drivers, Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards.

http://www.rural-inclusion.eu/?q=en/node/9

K. Charvat junior, P Gnip, Karel Charvat COIN – Module for Tactical planning of agriculture production. EFITA/WCCA 2011, Praha

Karel Charvát Future of Open Source systems, Geoinformatics FCE CTU 2009

Karel CHARVAT, Zbynek KRIVANEK Dual Licensing Policy – New Way for Software Development, IST Africa 2010

Armantas Ostreika / KTC; Michael Printzos / PRO; Guillermo Schwartz / LOGCMG, Thorsten Reitz/FHG-IGD, Karel Charvat/HSRS, Humboldt, A2.4-D1 Software distribution strategies and business models, 2007

# Transparent Agrifood Interoperability, Based on a Simplified Dynamic Simulation Model

Monika Varga, Bela Csukas, Sandor Balogh
*Kaposvár University, Department of Information Technology*
E-mail: Varga.Monika@ke.hu

## Introduction

Agrifood sector has a strategic importance, both quantitatively (almost 7 billion people) and qualitatively (great impact on public health). Nowadays, an increasing part of the conscious consumer society claims for more reliable, healthy and safe food products. Advanced countries cannot prescind the fact of starvation. That is why, besides the safety and the quality of food production, another important question is the appropriate amount of food. To assist these efforts, many initiatives have been elaborated (ETPFL, 2005; IFPRI, 2010) for the forthcoming years.

Furthermore, in the last decades, agrifood sector have been confronted with several food scandals and with the request of the sector spanning traceability in the complex agrifood networks. Problem solving claims for more conscious social control over manipulations and value chains. The above reasons placed this topic to the focus of R&D interest many years ago.

According to the established solutions, we can distinguish three levels of the agrifood traceability systems:

- internal: inside the actors, with many field specific issues;
- external: between the actors for interoperability of the neighbors; and
- sector spanning: for the whole network of the sectors.

a.) Regarding internal systems, there are numerous ERP softwares, which offer complex solutions for process and production control in food chains, related to the procurement, inventory, production, requirements planning, sales, traceability, quality management and laboratory information system. Advantages of these systems are the integrated process management and the process based mapping of the materials in every step of the production, inside the actor's system. The problem is that the SMEs or private businesses are often not able to integrate such complex and high-priced systems. It is emphasized that very few SMEs have their own solution (Bunte *et al.*, 2009).

In a recent paper Fritz and Schiefer (2009) gave a comprehensive analysis about the decision complexity in the agrifood tracing and tracking from economical point of views. Authors summarized the results in a cost-benefit decision table.

There are many isolated case studies in the up-to-date literature to develop simple and feasible methods for the internal traceability. For example, Bertolini and his colleagues presented an application of an industrial engineering tool, named Failure Mode Effect and Critical Analysis (FMECA) to manage the production processes of the food chain (Bertolini, Bevilacqua and Massini, 2006).

An MILP based batch dispersion model was published by Dupuy and colleagues for the example of a French sausage manufacturing company (Dupuy, Botta-Genoulaz and Guinet, 2005).
Skoglund and Dejmek (2007) discussed the problem of continuous liquid food processing with the application of fuzzy batches. They introduce the so-called fuzzy traceability concept. They concluded that traceability systems in continuous production need further development.
In a relatively early work, Minegishi and Thiel (2000) analyzed how system dynamics could contribute to improve the knowledge of the complex logistic behavior of integrated food industrial processes, based on causal diagram.
b.) In external systems we can find various solutions, too. For example, Jansen-Vullers and colleagues proposed a gozinto graph model based approach to design an information system for tracing the flow of goods (Jansen-Vullers, vanDorp and Beulens, 2003).
To eliminate the problem of the island-like solutions, the most important question was the standardized communication between the various stakeholders with the unambiguous identifyability of the actors.
First of all, XML/EDI became a leading standard in the format of messages, both in the internal and external system of the connected actors (Bryan, 1998).
For identification the GS1 system (the former EAN.UCC standard) became the most frequently used one (GS1, 2011). The objective of the GS1 System is to make efficient the communication between the connected partners, by establishing a precise, but flexible method for unique identification of products and packages in both human and machine readable formats.
c.) The most unsolved field is the trans-sectorial system, because of the complexity of the agrifood networks. Different kinds of information from different actors of different chains make difficult the information flow "from the farm to the table of the consumer".
In 2010 Wolfert and his colleagues described explicitly, that a sector spanning system has not yet been developed. They made an effort to outline a really sector spanning ICT method (Wolfert, Verdouw, Verloop and Beulens, 2010).
To explain the motivation of our work, let's see the problem to be solved from another aspect. We can distinguish three control levels of agrifood processes:

- process control,
- production control and
- process network interoperability control.

The lowest level process control operates the detailed technological processes of a single actor in the network. In agrifood context it means e.g. the GIS supported precision farming. Probably, it is the most elaborated level, because the rapid development of IT makes possible the use of advanced identification (e.g. RFID, optical identifiers) and standardized analytical (TraceFood, 2008) methods.
The medium level production control exists mainly in the ERP systems. According to a recent survey about the Hungarian meat production chains (Füzesi, 2010), the most frequent ERP softwares, applied by the food chain actors are CSB-System and Microsoft Navision.
The most unsolved level is the process network control, in spite of the standardization efforts. Although the necessary data are available from the previous levels, there is not

a comprehensive, widely usable method for the systematic elaboration of these data. This is not surprising, because the large networks, the globalized trade, the heterogeneous actors, and the individual interests make almost impossible the transparency in the agrifood networks.

To sum up the available methods, we can find two extremely different concepts both in the practice and in the literature. First is the solution with standardized IDs, according to the prescribed "one-step-up, one-step-down" principle. However, in this way we lose the most important thing, the transparency of the whole network. The other approach plans huge central databases, where all data of the various actors are stored. The widespread character of the agrifood network, the heterogeneity of the stakeholders, and the large amount of data make this approach irreal. We can state, that we have to search for an appropriate intermediate solution.

Schiefer (2008) highlights that because of the complexity of the food chain, enterprises cannot solve the problem alone, but it requires concerted action. According to his opinion, feasible and workable solutions require agreements between the different actors and sectors, and any new solution should be based on the bottom-up organized, cooperative system of the actors.
Wolfert, Verdouw, Verloop and Beulens (2010) in a recent paper suggest the development of sector-specific, SOA based open models and standards for information integration.

**Objective**

Main objective of the present work is to develop a new, dynamic process model database, containing the essential building elements of the macro level processes. The continuously upgraded, model databases shall be used for the problem specific, ad hoc generation of the dynamic models in the solution of various tracking and tracing problems. The skeleton of the simplified dynamic model is the mass balance, with the known or estimated stoichiometries. These mass balances can be associated with the actually studied case specific intensive variables (concentrations, prices, etc.).

Actually, we try to find the general building elements for the unified modeling of the quite different agrifood processes. With the knowledge of these model elements, we implement a methodology for the unified generation and execution of these macro level models. Our final objective is to develop scalable methods for problem solving by the application of these multiscale process models.

The solution outlines an organizational architecture, where the so-called Agrifood Interoperability Centers help the cooperation of the actors, as well as the communication between the actors and the various authorities.

Example systems for the development of unified agrifood process models
To determine the common building elements of the various actors we tried to choose consciously different stakeholders, with quite different activities.

**Arable farming**

First example was an existing farm with arable farming activity in almost 80 parcels.
The actual state of the system can be described by the storages of the input raw materials, and output products, as well as by the parcels.
There are input storages for the seeds, fertilizers, pesticides, etc., and output storages for the various products (crops, straw, hay, etc.). At the beginning of the crop year, farmers allocate the planned cultures to the identified parcels. It is a keynote activity, because unambiguous identification of the parcels is necessary for the tracing and tracking. However, it causes special difficulty, because the parcels may be planned yearly, as well as, according to the (Hungarian) regulations, the name and area of the parcels are recognized officially only in early May of the next year.

Parcels can be modeled as special "containers" for the input of the various raw materials and for the output of the respective products.

The most important transitions are the cultivation processes, associated with the parcels. The input and output quantities, as well as the optionally used estimated stoichiometries are described by the respective transition elements of the cultivation process.The macro level models have to describe only those operations that effect on the overall material balance with the environment. Considering the continuous actualization of the model, instead of the single, gross mapping of cultivation we use individual transitions for the elementary steps of the cultivation process. In the reality they are usually event-driven processes, depending on many environmental factors like wheather. However, in the stepwise supply of data they appear as time-driven elementary processes, characterized by a specific date or period. The overall stoichiometry of the cultivation is usually not known, however can be estimated, as needed. Without the knowledge of the stoichiometries, cultivation means an ensemble of sinks and sources. In problem specific investigations the experts can estimate the actually needed stoichiometries (e.g. for the percantage of a harmful components forwarded into a product, considering also a time dependent decomposition). It is to be noted that in the future it would be very useful to describe the approximate stoichiometries completely. This would make possible for example to simulate the contribution of the given cultivalition to the carbon balance. Of course, this extension needs to embed additional knowledge into the model database.

The re-allocation and renaming of the parcels can be described by special „stoichiometric" rules. They can solve the declaration of N new parcels from M old ones. The tracing and tracking algorithm have to take into account these changes dynamically.

Additional transitions can model the purchases and sales, while the model supports also the handling of the respective prices, that is useful in the possible value chain analysis.

**Game management**

In the investigated deer farm a given, individually identified part of animals is reserved for high quality meat production.

Accordingly, the most important state characteristics refer to the groups of animals, while in the dedicated group(s) the deers have individual IDs. Also the age of these animals is registered and incremented continuously. The remaining part of animals is contained in various groups, where the group members are characterized by the same, approximate properties.
Other important class of state properties declares the various nutrition recipes that depend on the age of the animal and on the actual season.
The storage of the inlet materials (herd's grass from the separated fields, consumed fodders and food industrial by-products, etc.) and of the outlet products (slaughter animals, surplus arable farming products) are defined in the respective state elements, too.

The nutrition model of the animals can be determined by a set of transitions. Each of them prescribes the actual animal age and season depending stoichiometry, according to the knowledge of the experts. These stoichiometries describe the consumption of the various naturally growing and bought fodders, as well as the related weight increase of the animals. The elementary models of the nutrition can be identified and validated according to the systematic weighting of the animals. Also the accidental health problems and the necessary medication of the animals are recorded, generating special transitions that extend the description of the individuals with this information. The continuous decomposition of the medicines can also be taken into consideration.

The extensions of the set with the newborn deers, and with animals arriving from other farms are taken into account by additional model elements. Also the selection for slaughtering, the death and the transfer of the animals from the game preserve to the hunting area are taken into account with special events.

The set of the transitions should be supplied with the commercial transactions (purchases and sales).

**Slaughterhouse**

The state elements of the slaughterhouse are described by the stored slaughter animals, as well as by storages for the carcasses, refrigerated carcasses, chopped products and packed products. All of these storages must be characterized by the set of individually identified items that makes possible to identify the animal, as well as the age of the product since the chopping. The parts of the "bulk" products cannot be identified individually; rather the whole amount is characterized by the respective herd of deer.

The chopping and cutting transitions can be described by well-defined stoichiometries, determining the ratio of the various products and by-products.

Additional transition elements determine the purchase of the animals, as well as the sale of the carcasses, or of the cut and packed products.

Common features of the building elements in the different actors
In Table 1 we summarized the main characteristics of the three example systems.

| Actor's profile | Actor 1. Arable farming | Actor 2. Game management | Actor 3. Slaughterhouse |
|---|---|---|---|
| **Data sources** | obligatory data registers (e.g. the formal parcel identifying system) | obligatory data registers | individually dedicated software, supporting data registration |
| **Frequency of data acquisition** | 1-2 weekly, in accordance with the agricultural processes | various (e.g. daily for feeding, or monthly for rare events) | daily |
| **Transition elements (activities)** | agricultural and commercial processes with mass flows | registration, feeding, growth, selection for slaughtering and commercial processes | slaughtering, chopping, cooling, storage, commercial processes |
| **State elements (components)** | parcels, input and output storages | individually registered animals, fodders, medicines, feeding strategies | individually registered animals, warm and cooled carcasses, cutted, packed and bulk meat goods |
| **Difficulties, specialties** | redistribution or rename of the parcels | feeding strategies depend on seasons and age, individual identification of animals | identification of packages, produced from different carcasses |

*Table 1. Overview of the investigated examples*

It is to be noted that all cases can be described by conservational models, i.e. by (finally conservation based) extensive quantities and balance processes. The conservational models are supplied with the necessary signs and rules, described in the language of informational processes.
In every case the most important issue is to outline the structure and the initial conditions of the model on the basis of the previous period's data. Having fixed the initial state, the real-time, stepwise data supply can be described by the respective new transportations and transformations. The frequency of the upgrade depends on the characteristics of the investigated system (weekly or bi-weekly upgraded day scale for arable farming and game management, but daily detailed hour scale for slaughterhouse).

## Computer implementation of the unified agrifood process models

The computer implementation is based on the principles of Direct Computer Mapping of the process models (Csukas, Varga and Lakner, 1996; Csukas and Balogh, 1998; Csukas, Balogh and Bankuti, 2005; Varga, Csukas and Balogh, 2010). The developed software is written in GNU-Prolog (Diaz, 1999-2010). The temporarily used interface is solved by the appropriate extension (Varga, 2009) of the open source code Graphviz (2009).
According to our conception, the primary structure of the various process models is a net, consisting of two kinds of nodes (states and transitions), as well as of three kinds of edges (increasing, decreasing and signaling).

The GNU Prolog representation of these elementary building blocks provides us a toolkit, to build quite different process models for the various agrifood actors, easily. First we introduce the basic building blocks (prototypes) of our tool, followed by the basic principles of the model building from this prototypes. The meta-prototypes of the building elements, shown in Fig. 1, are the followings:

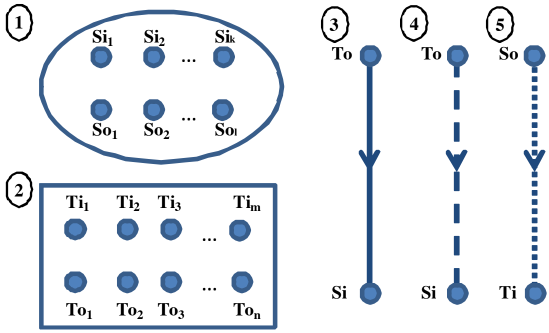



*Figure 1. Meta-prototypes of the building elements*

1. State elements, characterizing the actual state of the process (ellipses in the graphical representations);
2. Transition elements, describing the transportations, transformations and rules, corresponding to the time-driven or event-driven changes of the actual state (rectangles in the graphical representation);
3-5. Connection elements, designating the directed transport of the respective measures or signs between the state and transition elements (different lines correspond to the different changes in the graphical representation).
The state and transition elements contain lists (i.e. arbitrary number) of input (Si or Ti) and output (So or To) slots. The identifier and the type of slots must match to the sending (input) and receiving (output) end of the connections, respectively.
The simplified description of the state and transition elements is the same, as follows:

- element(Name,Coord,ProgramName,ParameterList,InputList,
- OutputList,Timing).

Both state and transition elements are characterized by the following major attributes:

Name: identifying name;
Coord: coordinates, determining the scale and place of the given element in the geometrical and parameter space;
ProgramName: identifying name of the program;
ParameterList: parameter slots, prepared for the storage of the local data, associated with the given elements;
InputList: input slots, prepared for receiving data from the containers of the designated connections;

OutputList: output slots, prepared for sending data to the containers of the designated connections;
Timing: instructions about the temporal behavior of the given element (see later).

The slots of ParameterList, InputList and Outpulist are described by the following properties:

SlotName: determines the local identifier of the given slot;
SlotType: gives instructions to the interpretation of the value, associated with the slot;
SlotValue: contains the list of data, e.g. in the form of

d(DataName,DataValues,Dimension)

functors, where

- DataName: identifies the individual data set;
- DataValues: is the list of data (numbers or atoms);
- Dimension: determines the measurement unit or n/a.

The local functionalities of the state and transition elements are described by the program code, identified by the respective ProgramName. Usually many elements use the same program, declared by the prototype of the given subset of elements. In the local execution the elements receive input, execute program and send output.
The programs, referred by ProgramName calculate the values of OutputList from the data of InputList and ParameterList according to the

program(ProgramName,InputList,ParameterList,Outputlist) :- ProgramCode.

clause, where ProgramCode binds the free variables of OutputList with the knowledge of the bound variables of InputList and ParameterList.
In the general case, the state and transition elements may contain both conservational and informational slots. Conservational input slots can receive data only from the increasing (solid lines) and decreasing (dashed lines) connections, coming from conservational output slots. Informational input slots can receive data only from the signaling connections (dotted lines), coming from informational output slots. In contrary, conservational output slots can send data only via increasing and decreasing connections to the conservational input slots, as well as informational output slots can send data only via signaling connections to the informational input slots. There may also be pure conservational and informational state and transition elements, as special cases.
The syntactically identical state and transition elements can be distinguished structurally and functionally. The structural difference means that, in the sense of the General Net Theory (Brauer, 1980), only the state → transition and transition → state connections are allowed. The functional difference, in the sense of the State Space Model (Kalman, Falb and Arbib, 1969), is rather semantic than syntactic. Regardless to the fact that both kinds of elements are associated with programs, at a given point of time the actual state of the process is determined by the state elements. In contrary, the dynamic behavior of the process is determined by the

transition elements. The functioning of the state elements is limited to the collection, interpretation and distribution of the static characteristics, while the transportations and transformations are modeled by the transitions.
The syntax of the

```
connection(SendOperator,SendElement,SendCoord,SendSlot,
        ReceiveOperator,ReceiveElement,ReceiveCoord,
        ReceiveSlot,DataType,DataSet,Timing).
```

is general for all increasing, decreasing and signaling connections. All of them carry data in the container of DataSet from a sending slot to a receiving slot and they are characterized by the following major attributes:

SendOperator: determines the action to be done at sending slot (e.g. read, etc.);
SendElement: identifies the sending element;
SendCoord: refers to the (scale and place) coordinates of sending;
SendSlot: defines the sending slot of the SendElement at SendCoord;
ReceiveOperator: determines the action to be done at receiving slot (e.g. write, decrease, increase, remove, extend, etc.);
ReceiveElement: identifies the receiving element;
ReceiveCoord: refers to the (scale and place) coordinates of receiving;
ReceiveSlot: defines the receiving slot of the ReceiveElement at ReceiveCoord;
DataType: gives instructions to the interpretation of the DataSet;
DataSet: contains the list of data in the form of the above mentioned functors d();
Timing: contains instructions about the temporal behavior of the given connection.

Increasing and decreasing connections transport DataSet from transition to state elements. Signaling connections can transport DataSet both from state to transition elements and *vice versa*. The special reading connections of the conservational substructure transport intensive parameters from the output slots of state elements to the input slots of transition elements.
The temporal behavior of the elements and connections is declared by the associated list of Timing, containing the

t(From,To,[When1,When2,…,WhenM],Step)

functors, where

From: is a possible starting time;
To: is a possible ending time;
When1, When2,…,WhenM: are prescribed discrete times of the execution;
Step: is the individual time step of the repeated execution.

In Figs. 2 and 3 simplified examples for a transition and a state element of an agrifood process are illustrated.

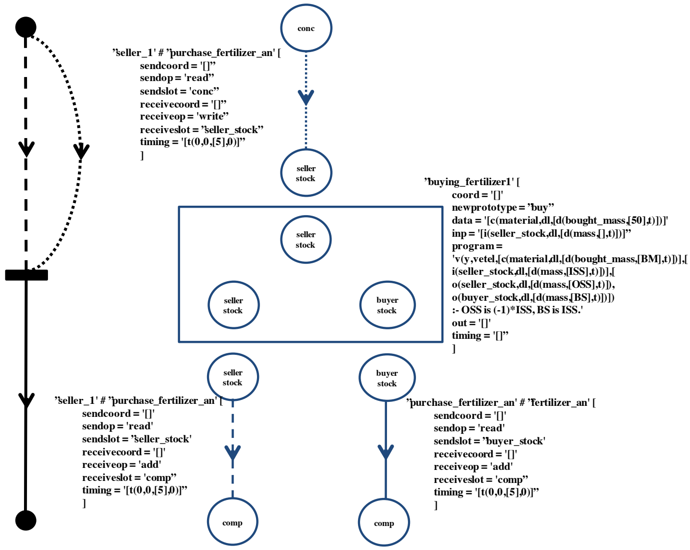



*Figure 2. Simplified example for a transition elements with its connections*

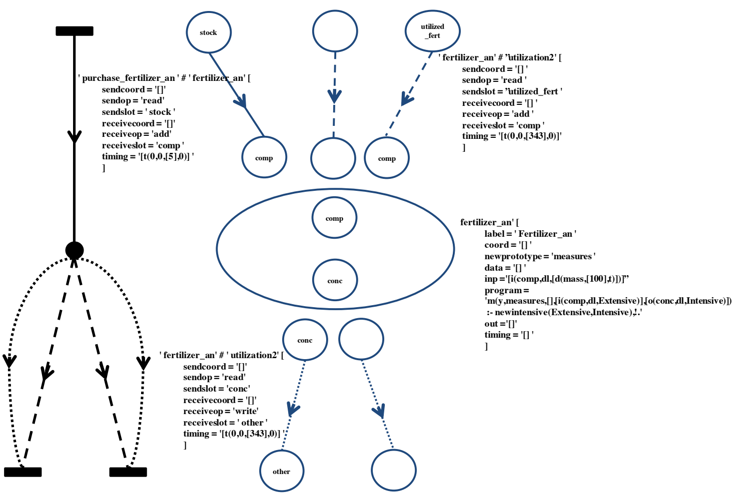



*Figure 3. Simplified example for a state element with its connections*

The multi-scale modeling is supported also by the arbitrary number of integer coordinates given in the lists of Coord. Say, Coord = [3,7] refers to the seventh element in scale II, embedded in the third element of scale I. The connections can be interpreted both within a scale and between the scales.

The model is extendable, because the number of data, containing and transporting functors is not prescribed, rather both the slots and the containers of the connections are characterized by the lists of respective functors d(.). The only convention is that the optionally single first element (first functor of conservational data) refers to the mass, while the following ones refer to the intensions or extensions, carried with the given mass batch or mass flow. The conservational state elements receive extensive changes and send intensive properties, while the conservational transition elements receive intensive characteristics and send extensive changes, *vice versa*. The increasing and decreasing of the measures, as well as the extending and removing of functors can be initiated by the 'increase' and 'decrease', as well as 'extend' and 'remove' operators, respectively. The distribution of the investigated new components is calculated by the exactly known or estimated stoichiometries.

**Execution of process models**

The schematic architecture of the computer implementation is illustrated in Fig. 4. The prototypes of the underlying process structures, as well as the brief programs, determining the calculation of the transitions and state functionalities of the given application are added by the expert through the editable expert interface. The structure and the parameters of the actual problem are described by the user, via the user friendly graphical and template generating interface.

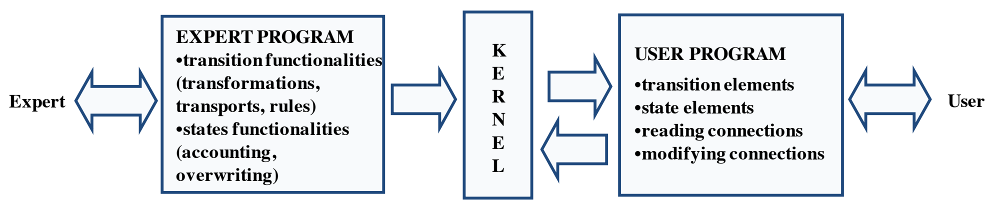



*Figure 4. Computer implementation of process models*

The simplified execution of the models consists of four cyclically repeated consecutive steps, as follows:

1. transition elements read the content of the associated state elements through the reading (signaling) connections;
3. brief programs, associated with transition elements calculate the changes;
4. state elements are modified according to the changes carried via modifying connections;
5. brief programs, associated with state elements calculate the new state.

The methodology makes possible the reverse dynamic simulation of conservation based stoichiometric processes, i.e. we can start from any simulated final state and can simulate the process backwards in time. The transitions are calculated causally right, while the increases are replaced for decreases, and *vice versa*. Also the extending and removing of the add-on intensive characteristic should replace for each other. Consequently, with the knowledge of the stoichiometries, the inverse simulation supports the quantitative tracing of the various problem specific components backward in space and time.

The recent implementation of the methodology has been written in GNU-Prolog, while the case specific unification of the higher level structures, contained in dynamic partitions, supports the generalized method development.
Temporarily an extended GraphViz input interface (Varga, 2009) and a CSV file based Microsoft Excel output interface are used for testing of the methodology and problem solving. The ongoing new implementation is a platform independent and partly open source tool, with a Qt based, interactive GUI (Qt, 2010). The interface involves a graphical modeling and design environment that allows both the user and the field expert an easy access to the input/output data, while the expert can modify and extend also the field-specific program prototypes.

**Examples for the application of the unified agrifood process modeling methodology**

In the following we illustrate the adaptation of our modeling methodology for the investigated agrifood processes. We show only a few important puzzle elements of the model building with some characteristic results from the three examples.

**Description, interpretation and execution of the model**

First step of model building is the determination of the actor's major activities with the accompanying material flows. Next we outline the structure graphically, according to the principles of the applied methodology. Fig. 5 shows the graphical representation of a typical parcel from the investigated arable farm. As we mentioned, rectangles represent the various activities, ellipses refer to the components, as well as the dotted, dashed, and solid edges determine the reading, decreasing and increasing of the given components, respecitvely. Graphviz maps the structure automatically into a simple *.DOT format. Having extended this file with the GNU-Prolog declarations, it will describe the characteristics of the investigated process (e.g. the actual extensive quantities, the initial concentrations, etc.). The kernel program of the simulator generates the user and expert modules from this text file, automatically. Next the kernel executes the dynamic simulation or backward simulation according to the prescribed time step, and write the results into a simple *.CSV file. The qualitative tracking and tracing program is based on the final state of the user module.

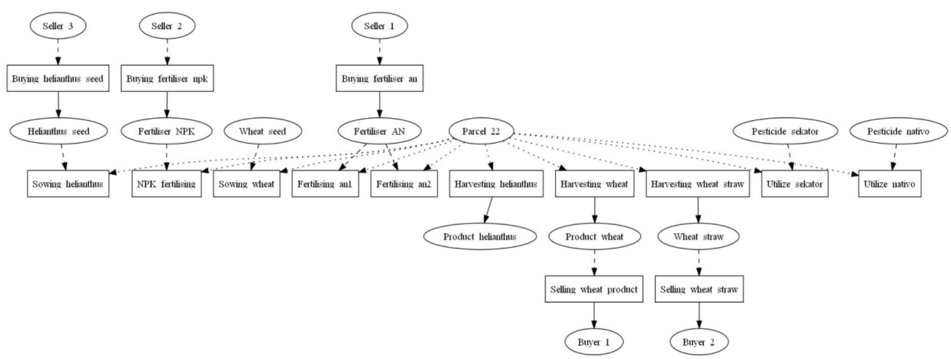



*Figure 5. Graphviz representation of one parcel from the example farm*

Forward and backward simulation (quantitative tracking and tracing)
A representative example for the simulated results of the game management is shown in Fig. 6. Left part of the Figure illustrates the dynamic forward simulation of the changing weight for an individually identified red deer. Right part of the Figure shows the results of the backward simulation for the same individual. As the dynamic simulation provides the basis for the quantitative tracking, inverse simulation supports the quantitative tracing functions. In this case, starting from the saved final state, the kernel executes the backward simulation, with the knowledge of the same equations.

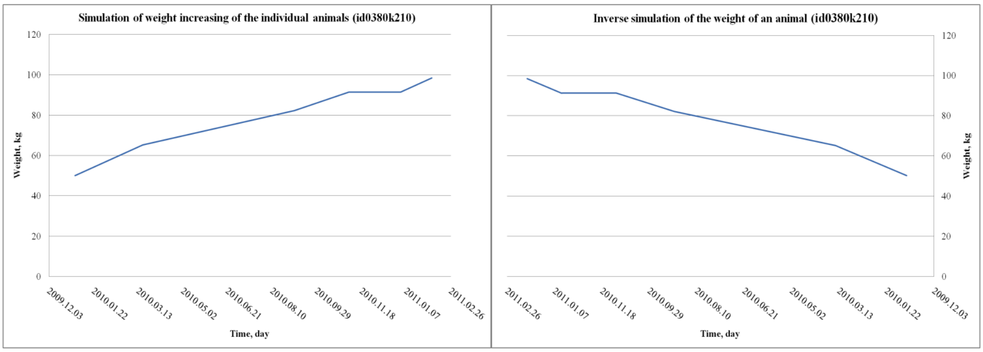



*Figure 6. Forward and backward simulation of the changing weight for an individually identified animal*

**Quantitative tracking and tracing**

Based on the simulation model, with other Prolog algorithms we can search (trace) for the origin of a given component, or we can track the way of a given component, until all of the possible outputs. In this application, the program lists the possible routes into Tables. Table 2 illustrates an example search. As it shown in the Table, we can follow backward the way of a hypothetical harmful component that appeared at Buyer_1. In more sophisticated applications the model database with the search algorithms help to identify the most informative measurement points to identify the possible origin of the examined component.

| Possible origins of a harmful component in Buyer_1 from 600 day before 31 July 2010 | | |
|---|---|---|
| seller_3 | measure | 2009 year 4 month 2 day |
| helianthus_nkferti | measure | 2009 year 4 month 3 day |
| seller_1 | measure | 2009 year 4 month 6 day |
| fertilizer_npk | measure | 2009 year 9 month 24 day |
| wheat_lupus | measure | 2009 year 10 month 13 day |
| fertilizer_an | measure | 2010 year 3 month 10 day |
| nv_sekator | measure | 2010 year 4 month 9 day |
| nv_nativo | measure | 2010 year 5 month 22 day |
| produced_wheat_lupus | measure | 2010 year 7 month 25 day |

*Table 2. Search results for the origins of an optional component in the model database of the arable farming*

Simplified illustration of the multiscale, trans-sectorial process model.
The capabilities of the methodology are summarized in Fig. 7. In this simplified example there are three connected actors: the arable farming, the game management and the slaughterhouse (three fields on the left). They are connected by many other actors, of course. In the example we illustrate one of the trans-sectorial ways, highlighted by red color.

The more detailed process model structure of a given parcel in the arable farm is enlarged on the right top part. The elements, involved in the investigated pathway, are marked with red.
The diagram in the middle right shows the weight change of an individually identified animal. Here we illustrated also the optional backward simulation (see the symmetric curve with the decremented time in the abscissa).

The simplified scheduled pathway of an investigated ingredient through the process network is summarized in the Table below. Besides it, in the right bottom corner a simulation based diagram shows the concentration of an example component in the subsequent stages of the process.

The individual model databases of the cooperating actors start with the installation of the initial state, next the model is stepwise upgraded by the additional connections, representing the actual changes. Also the new state and transition elements have to be added, as needed.
The actual sector-spanning problem solving models are generated from the process model database automatically. It means that considering the problem to be solved, the upper scale model opens and executes the necessary lower scale models.

The model based problem solving for the trans-sectorial processes covers qualitative tracking, qualitative tracing, quantitative tracking by simulation and quantitative tracing by backward simulation. As useful byproducts, the methodology supports also the determination of the hidden resources, the identification of the wastes and the value chain analysis.

The experiences about the examples can be summarized, as follows:

- The major problem is that the continuously changing mandatory data supplies are overlapping, but not complete;
- Another problem is that the existing data exchange with the official authorities does not support the observability and controllability of the processes.

The experiences about the method are the followings:

- It makes possible the unified management of the completely different processes;
- It solves the scalable storage of process models;
- The declarative, logical programming supports the algoritmization of spatial and temporal tracing and tracking;
- The dynamic mass balances can be associated with the actually needed intensions (e.g. concentrations, prices).

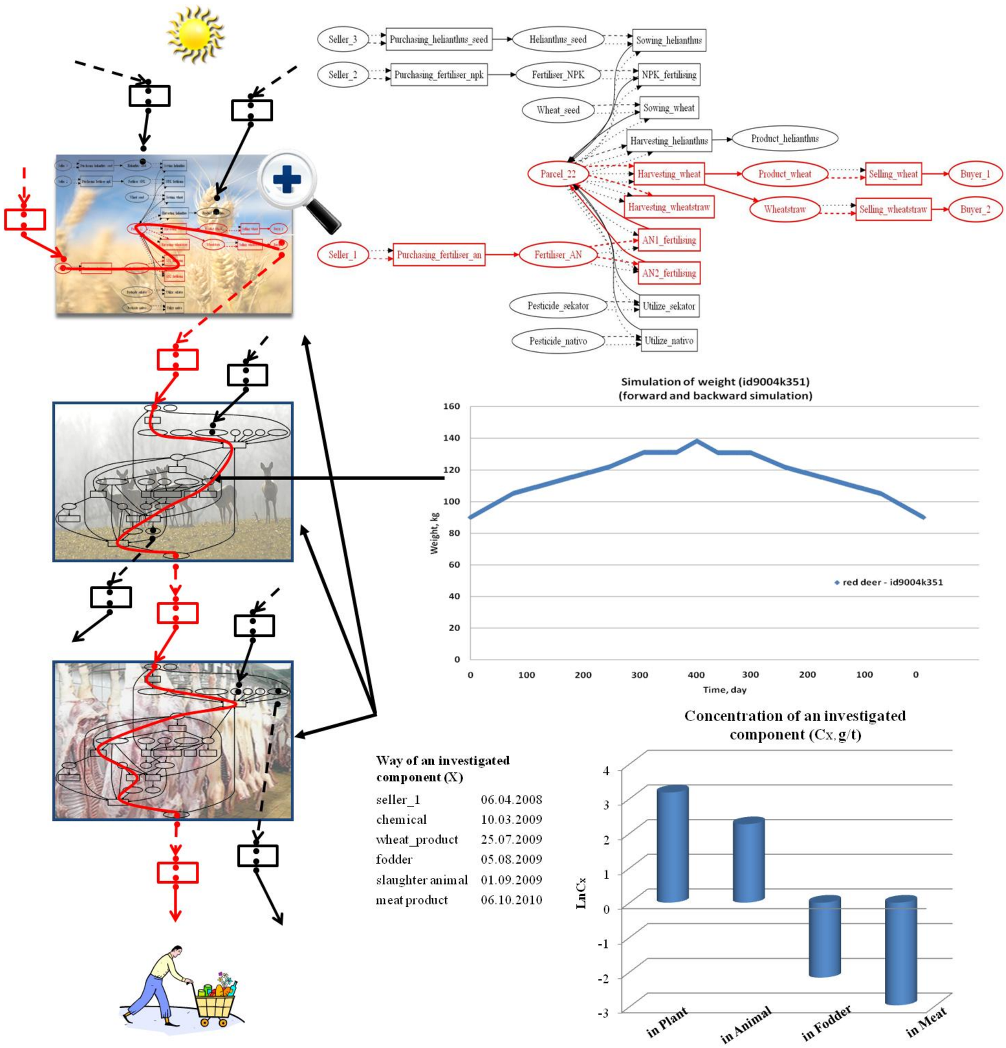



*Figure 7. Illustration of the problem solving capabilities of a trans-sectorial model*

## Outlook

Architecture of the planned Agrifood Interoperability Centers. The effective implementation of the above described methodology can be solved by the cooperative system of the actors, coordinated by the so-called Interoperability Center. The schematic architecture of the Agrifood Interoperability Service is illustrated in Fig. 8.
First (orange arrows in Fig. 8) the process models and the communication interfaces have to be generated by the support of the Interoperability Center. The suggested method of the stoichiometric mass balances makes possible to generate uniform process models from the same building blocks for the quite different technologies and activities.

The applied process modeling methodology makes possible the simultaneous generation of the user interfaces. It is to be noted that the majority of the systematically reported process data is limited to the new "connections", describing the up-to-date transportations and transformations. Nevertheless the method supports the assertion, modifying and deletion of state or transition elements, too.

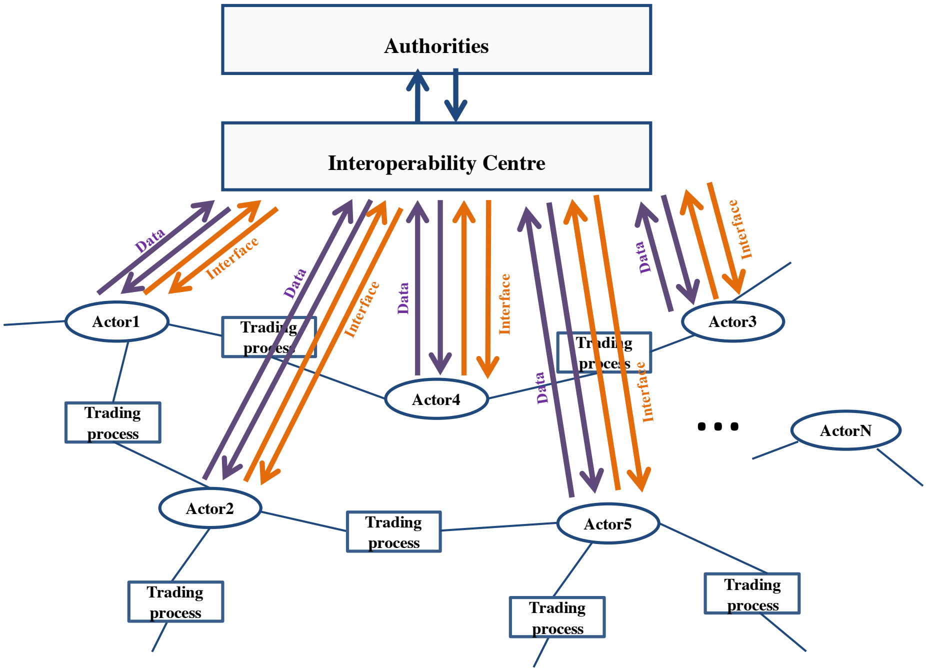



*Figure 8. Architecture of planned Agrifood Interoperability Services*

In the case of the actors, having an appropriate ERP system, the model based interface can be adapted to the existing ERP software. It is worth mentioning that the required system of data is very familiar with the capability of the usual ERP modules. For the frequently used ERP systems easily configurable and uniform applications can be generated.
For the smallest actors (e.g. minor private companies), who do not have ERP systems, a special user-friendly application can be generated by the Interoperability Center.
Having installed the models in the Interoperability Center and the interfaces at the actors, the systematic data reporting, as well as the in-demand problem solving can start (purple arrows in Fig. 8).
The data reporting from the actors to the Interoperability Center means reporting about

- input and output mass flow of TRUs (Kim, Fox and Gruniger, 1995);
- storages of TRUs;
- calculated or estimated stoichiometries between the TRUs;

- additional knowledge and measurement abilities that can be utilized in special investigations.

Problem solving services of the Center for the actors and for the authorities
Interoperability Center serves to actors:

- sector spanning tracing and tracking information; and
- suggest and/or realize additional measurements in actual problem solving.

In addition, Interoperability Center can serve the authorities special, case specific search for the possible origin and dispersion of decontaminations, as well as information about resources, wastes and value chains.
The problem solving services of the Interoperability Center are based on the simplified, dynamic, stoichiometric mass balance models, stored in the central database. The multiscale storage of the models helps to manage the complexity, because only the actually necessary sub-processes have to be opened. The most important tasks can be summarized as follows:

- Qualitative tracking by multiscale search along the forward balance routes;
- Qualitative tracing by multiscale search along the backward balance routes;
- Dynamic simulation based quantitative tracking for the concentration of the known or *ad hoc* appearing components to be studied;
- Backward dynamic simulation based quantitative tracing for the concentration of the known or *ad hoc* appearing components to be studied;
- Interactive, measurement supported search for the possible origin of the various contaminations (combining the above methods with genetic algorithm);
- Reporting about hidden resources or wastes on the basis of balance calculations;
- Trans-sectorial value chain analysis;
- Analysis of the basket of typical consumers' groups.

The problem solving of the Interoperability Center can be based on real data and/or on simulated, hypothetical case studies.

**Conclusion**

We conclude that organizing an agrifood model database according to the building elements of the macro level dynamic mass balances, makes possible the case specific, ad hoc generation of the dynamic models for solving of various agrifood interoperability problems.
There are an appropriate set of general building elements for the unified macro-level modeling of quite different agrifood processes. The database of these model elements makes possible to implement a methodology for the unified generation and execution of the macro level models. Based on this process models, scalable methods can be developed for qualitative tracking and tracing, as well as for dynamic simulation based quantitative tracking and backward dynamic simulation based quantitative tracing. The model database based algorithms can support also

the determination of the hidden resources, the identification of wastes and the value chain analysis.

According to the preliminary results, we can state, that the applied methodology can generally be adapted for the various actors of the agrifood network.

The agrifood process interoperability is a good example for the future economic paradigm, because:

- agriculture utilizes the the solar energy, i.e. the single outer resource of Globe;
- food is indispensable for mankind; and
- interoperability can support the strategically and economically straightforward local supply.

This tasks can be solved by the cooperative system of actors with the planned Agrifood Interoperability Centers, that may be out of the individual interests.

**Acknowledgement**

The project was supported by Baross Gábor R&D Project, #REG-DD-09-2-2009-0101.

**References**

Bertolini, M., Bevilacqua, M., Massini, R. 2006. FMECA approach to product traceability in the food industry. *Food Control,* 17, pp.137-145.

Brauer, W. ed., 1980. *Net Theory and Applications*. Lecture Notes in Computer Science, 84, Springer-Verlag.

Bryan, M. 1998. Guidelines for using XML for Electronic Data Interchange http://www.eccnet.com/xmledi/guidelines-styled.xml

Bunte, F.H.J., Dijkxhoorn, Y., Groeneveld, R.A., Hofstede, G.J., Top, J.L., Vorst, J.G.A.J.van der, Wolfert, J. 2009. Thought for Food: the impact of ICT on agribusiness. In: *Proceedings of the 7th EFITA Conference*. Wageningen: Wageningen Academic Publishers, pp. 913-922.

Csukas, B., Balogh, S. 1998. Combining genetic programming with generic simulation models in evolutionary synthesis. *Computers in Industry,* 36, pp.181-197.

Csukas, B., Balogh, S., Bankuti, Gy. 2005. Generic Bi-layered Net Model – General software for simulation of hybrid processes. In: Li, D., and Wang, B. eds. 2005. *Artificial Intelligence Applications and Innovations II.*, USA: Springer, pp.701-710.

Csukas, B., Varga, K., Lakner, R. 1996. Dynamic simulation by algorithmic generated structural model I., Principles of the Model Generator. *Hung J Ind Chem,* 24(2), pp.107-130.

Diaz, D. 1999-2010. The GNU Prolog web site, http://www.gprolog.org/

Dupuy, C., Botta-Genoulaz, V., Guinet, A. 2005. Batch dispersion model to optimise traceability in food industry. *Journal of Food Engineering,* 70, pp.333-339.

ETPFL, 2005. ftp://ftp.cordis.europa.eu/pub/technology-platforms/docs/foodforlife_etp_visiondoc.pdf

Fritz, M., Schiefer, G. 2009. Tracking, tracing, and business process interests in food commodities: A multi-level decision complexity. *Int. J. Production Economics,* 117(2), pp.317-329.

Füzesi, I. 2010. A húsipari termékpályák minőségmenedzsmentjének IT támogatási lehetőségei (*in Hungarian*) (IT support possibilities in Quality Management Systems of Meat Product Chains). In: Herdon, M. and Kapronczai, I., eds. 2010. *Agricultural Informatics Studies I.*, pp.4-41.

GraphViz, 2009. www.graphviz.org

GS1, 2011. www.gs1.org

IFPRI. 2010. http://www.ifpri.org/publication/new-vision-agriculture-2020-conferences

Jansen-Vullers, M.H., van Dorp, C.A., Beulens, A.J.M. 2003. Managing traceability information in manufacture. International Journal of Information Management, 23, pp.395-413.

Kalman, R., Falb, P., Arbib, M. 1969. *Topics in Mathematical System Theory*. New York: McGraw Hill.

Kim, H.M., Fox, M.S., Gruniger, M. 1995. An Ontology of Quality for Enterprise Modeling. In: *IEEE Proceedings of WET-ICE*. Los Albamitos, CA, USA, pp.105-116.

Minegishi, S., Thiel, D. 2000. System dynamics modeling and simulation of a particular food supply chain. *Simulation Practice and Theory*, 8, pp.321-339.

Qt, 2010: http://qt.nokia.com

Schiefer, G. 2008. Tracing and Tracking – A Challenge for System Organization and IT. *Journal of Information Technology in Agriculture*, 3, pp.19-25.

Skoglund, T., Dejmek, P. 2007. Fuzzy traceability: a process simulation derived extension of the traceability concept in continuous food processing. *Food and Bioproducts Processing,* 85(4), pp.354-359.

TraceFood, 2008. www.tracefood.org/index.php/Tools:Standardized_Analytical_Methods

Varga, M. 2009. Melléktermékeket hasznosító complex körfolyamat gazdasági optimalizálása (Economic optimization of sustainable complex processes). PhD Theses, Kaposvár University

Varga, M., Csukas, B., Balogh, S. 2010. Sector spanning agrifood process transparency with Direct Computer Mapping. *Agricultural Informatics,* 1(2), pp.73-83.

Wolfert, J., C. N. Verdouw, C. M. Verloop, A. J. M. Beulens. 2010. Organizing information integration in agri-food – A method based on a service-oriented architecture and living lab approach. *Computers and Electronics in Agriculture,* 70(2), pp.389-405.

# A Business Intelligence Approach to Support a Greenhouse Tomato Crop Grey Mould Disease Early Warning System

Miguel de Castro Neto[1], Fátima Baptista[2], Luis Manuel Navas[3], Gonçalo Ruiz[3]
*[1] ISEGI – Universidade Nova de Lisboa, [2] Department of Agricultural Engineering - University of Évora, [3] Department of Agricultural and Forestry Engineering - University of Valladolid*
*E-mail: mneto@isegi.unl.pt*

## Introduction

This paper presents a Business Intelligence architecture proposal, including data sources, data warehouse, business analytics, and information delivery, to launch an early warning system for greenhouse tomato crop grey mould disease.

Tomato is a very important crop in the Mediterranean region in general and in Portugal in particular being the production for fresh consumption made essentially in greenhouses.

Botrytis cinerea Pers.: Fr. is the causal agent of grey mould disease and is one of the most important diseases affecting greenhouse tomato crops, high relative humidity and the presence of free water on the plant surfaces have been recognized as favourable to the development of this disease.

The availability of a early warning system providing to the tomato grower alerts with information of the potential favoured conditions for the disease appearance in its early stages or even before can have a very positive impact in reducing the economic and environmental impacts due to a more rational and efficient disease control management.

Today we have the necessary technology to build and launch an Internet based early warning system for grey mould disease in greenhouse tomato crop supported by a wireless sensor network adopting a Business Intelligence approach.

From the research conducted until the moment the proposed solution is viable and the next step will be to validate it in the field in different locations and with distinct greenhouses conditions.

## Greenhouse Tomato Crop Grey Mould Disease

In the Mediterranean region greenhouse areas are significant with tomato being the most commonly grown vegetable (Castilla 2002). In Portugal according to available statistical information greenhouse tomato crop was cultivated in 1,550 ha with a total production of 98,646 ton in 2003 (GPP 2007).

Botrytis cinerea Pers.: Fr. is the causal agent of grey mould disease and is one of the most important diseases affecting greenhouse tomato crops. This disease could be responsible for production losses of 20% and treatments could represent about 60% of the total fungicides used over a cropping season (Prieto *et al.* 2003). Conditions inside greenhouses; warm, humid, high plant density and frequent handling are conducive to the establishment and spread of the pathogen.

High relative humidity and the presence of free water on the plant surfaces have been recognized as favourable to the development of grey mould. Due to the common occurrence of

grey mould, its potentially high rate of spread and high production losses it causes, growers usually apply large amounts of chemical fungicides. This practice may lead to chemical residues on tomato fruits which impede the commercialization, increase production costs and the risk of developing fungicide resistances (Abreu *et al.* 1994).
Environmental and health concerns have increased public attention and pressure to reduce chemicals use in agriculture over the last decade. The European Commission in a communication to the European Parliament in 2002 encourages agricultural practices that reduce or eliminate pesticide use. In response the Parliament recommended a 50% reduction in the use of these chemicals over 10 years (Resolution of the European Parliament 2002/2277). Also, only a few fungicides are now labelled for use in greenhouse tomatoes, and their high costs, have encouraged growers and scientists to find alternative methods to manage grey mould for sustainable and profitable greenhouse production.
In Mediterranean countries most greenhouses are very simple constructions, covered with polyethylene films, being ventilation the main environmental control technique used. Baptista (2007) studied the effect of nocturnal ventilation in the greenhouses climate parameters and on the B. cinerea occurrence, and showed that reducing relative humidity it reduces the occurrence of grey mould.
The more sophisticated facilities now being utilized for greenhouse crops have opened new opportunities for disease control. Most commercial greenhouses are equipped with sensors to measure air temperature and relative humidity. It turns possible to consider the use of automatic control systems based on disease risk levels, a useful tool for growers, providing the opportunity to decide what to do in order to avoid favourable conditions. This approach would help to reduce the number of chemical sprays, with unquestionable economical and environmental benefits.

## Greenhouse environment control

Recent evolution in the information and communication fields, namely in the mobile computing and remote sensing, made available in the market devices with growing processing capacities and smaller sizes which are able to offer sensing functionalities, wireless communication, integrated energy source and action capacities, which are posing a very interesting challenge to the agricultural sector (Hart and Martinez 2006; Wang *et al.* 2006).
In this context, wireless sensors usage in agriculture is considered to have high potential in the near future with applications ranging from environmental monitoring, precision agriculture, machine and process control, facility automation and traceability systems (Wang *et al.* 2006).
Particularly, in the context of greenhouse environmental conditions monitoring, the use of wireless sensors network (WSN) can be very interesting since this new technology can provide processed real-time field data from sensors physically distributed in the greenhouse. WSN are networks of small sensor nodes, with limited processing capacity, including sensors and their specific conditioning circuitry, communicating over short distances, normally using radio frequencies.
Usually most growers follow a chemical treatments calendar based on their experience and also rely on recommendations from the supplier’s technicians. Nevertheless commercial greenhouses are often equipped with sensors to measure and record air temperature and relative humidity. With this information and applying simple rules, like those based on the total time per day with relative humidity higher than 90%, growers could reduce the number

of chemical sprays. This will make it possible to act in time to reverse those conditions, by increasing ventilation or in cases when the risk is too high, by applying preventive fungicides. Other control measures such as cultural or biological should also be considered.

It is precisely in this field that we believe a Business Intelligence (BI) approach could be applied since its purpose is to integrate data from different sources, optimize its organization for analytical procedures through the usage of data warehouses, and support information delivery for decision-making.

In that sense the main objective of this research was to develop an early warning system prototype for B. cinerea in a tomato greenhouse supported by a wireless network of air relative humidity and temperature sensors installed inside the greenhouses.

With that purpose a sensor network with the underlying communication infra-structure was tested in a representative greenhouse in order to collect the data to feed an analysis control algorithm that we develop, test, correct and implement through a web based information system. This warning system combines the cost-effectiveness and friendly use with common communications technologies used by the majority of growers, which will contribute for an easier implementation in commercial greenhouses.

This system takes advantage of the new information and communication technologies to deliver the warning, like SMS, e-mail and Web access, widespread between greenhouse growers. Our goal is to demonstrate the potential of this BI approach and the utility of the prototype we built as a tool for growers and technicians to improve climate and disease control.

In fact Internet and related technologies are increasingly being used to support agriculture advisory and warning services (Hørning 2005, Jensen *et al.* 2004, Zazueta 2003) and with the generalization of ICT adoption and use in the primary sector this will be a standard platform to support agricultural information services in the near future.

With the future validation of the proposed Early Warning System prototype supported by the mentioned wireless sensor network integrating the rules identified in the field work and built in the business intelligence solution we will be able to deliver to the greenhouse growers a very powerful solution that can lead to important economic and environmental benefits.

## Information System Architecture

Considering that we can identify the disease risk level by monitoring the air relative humidity and temperature, we intend to focus on this problem and build a early warning system that can send real time information to growers through SMS, e-mail and web reporting supported by a wireless sensors network (WSN) installed inside the greenhouse and integrated in a Business Intelligence (BI) architecture.

After installing in the greenhouse a real time data recording system, in this case a air relative humidity and temperature WSN, we proceed with the construction of a rules based engine in a BI approach to deal with the recording, storage, processing of data and delivering of information to the end user as well as offering manipulation and analysis functionalities.

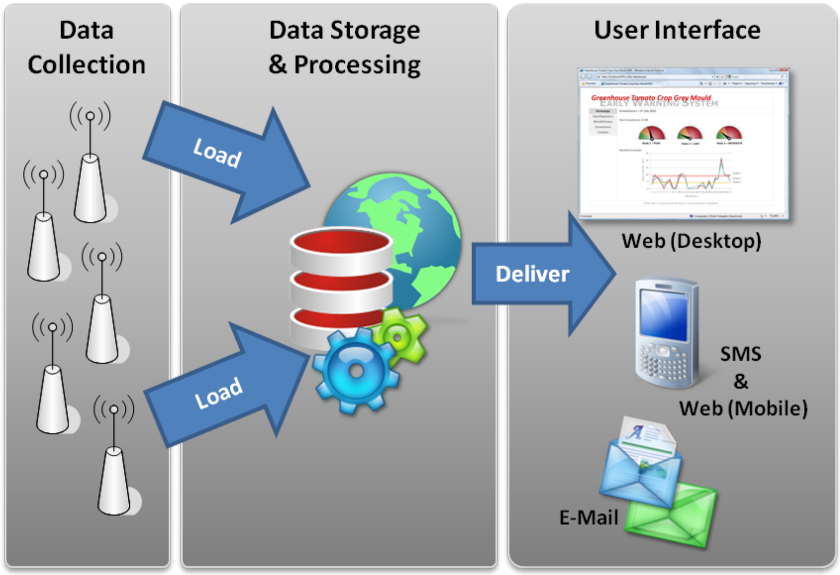



*Figure 1. Information system architecture overview*

BI is a broad term including tools, architectures, data bases, data warehouses, business analytics, performance management, methodologies, etc., all integrated in a single computer application in order to make available to the organization managers and analysts a fast and simple way of dynamic access to all business data in real time and also to offer the possibility of performing manipulation and analysis operations over it (Turban *et al.* 2007).

According to Watson (2009), BI is a broad category of applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users make better decisions.

This definition is broad since it encompasses not only applications, but also technologies and processes.

Through the analysis of historic and present data as well as with metrics and performance indicators, the decision makers obtain a knowledge that supports better and less risky decision- making. From the traditional functionalities offered by BI solutions we can refer: reporting and querying, complex analysis, data mining, forecasting, etc. BI solutions in an organizational environment have four main components: data sources; a data warehouse loaded from the data sources; a business analytics component which is a collection of tools to manipulate and analyse the data warehouse records; and a user interface such as a dashboard.

### Data collection and storage

The technological infrastructure used to implement the data collection and storage component of the early warning system prototype was a wireless sensors network.

Wireless sensors networks appeared in the 1970's for military and industrial use. The main advantages for wireless sensors networks are: low cost, easy use for temporary locations (Aakvaag *et al.* 2006), good signal to noise relationship (Pottie *et al.* 2000) and energy saving in multihop networks (Intanagonwiwat *et al.* 2000; Zhao *et al.* 2004). But they also have some disadvantages: low life-cycle batteries, components with limitations (low speed, memory…) (Hill *et al.*, 2000) and networks with a huge number of sensors.

Since then they have undergone a major evolution, particularly since the 90's, thanks to the improvements in wireless communications, these changes have allowed a wide variety of applications in different sectors such as agriculture and environment.

In this project we developed a SCADA application (Navas *et al.* 2010) programmed with LabVIEW® 8.6 (National Instruments), which allows management of data received by wireless sensors networks through a friendly user interface. For the application shown in this paper we have worked with a MEP 510 sensors network (Crossbow) and the collected data was stored in a Microsoft SQL Server database installed in a Web server used for the information delivery componnent.

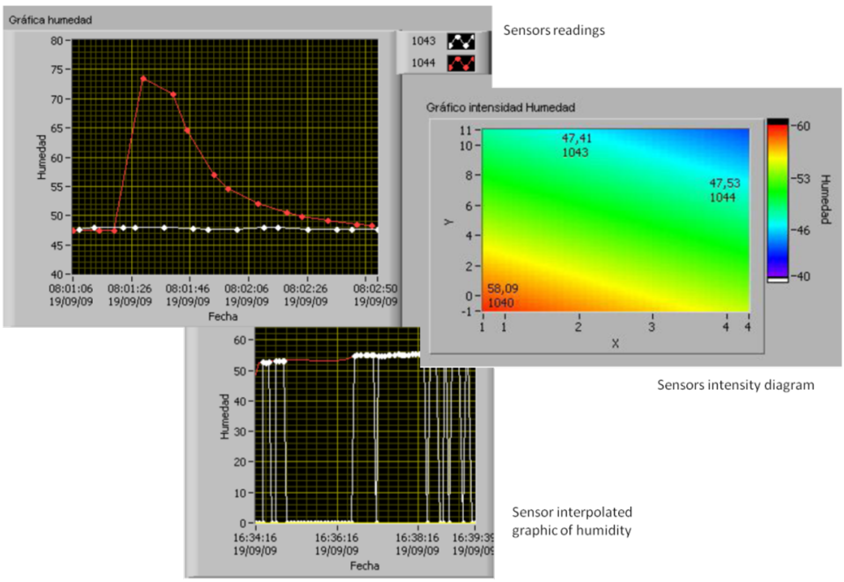



*Figure 2. Screenshots of the SCADA application (Navas et al. 2010)*

The functionalities implemented were the following: Network configuration; Data storage into database; Statistical processing of historical data with polynomial adjustment and spline interpolation; Visualization by data graphics in real time and historical data; Visualization of 2D intensity diagrams from the spatial distribution of sensors and Creation of a users registry system that allows, depending on the category assigned, receiving or not access privileges in the application.

The implemented sensors network and the application developed have been checked by operational tests for each functionality, as well as sensors joining and leaving the network situations, range of variables and working modes. The results obtained showed the robustness of the SCADA application and the limitations of wireless sensors networks operating on field conditions.

Data storage

Following a BI approach the data collected by the SCADA application and stored in a SQL Server relational database was used as a data source for a data warehouse modelled following a star schema optimized for the analytical procedures it would have to support in the next steps in order to deliver the desired early warning system.
In Figure 3 we can observe the data warehouse conceptual data model with the fact table in the center and the different dimensions around it.

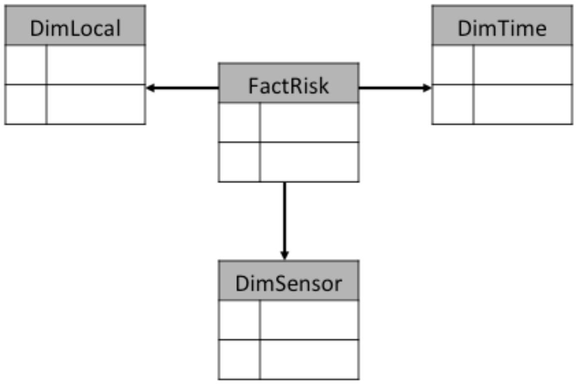



*Figure 3. Data Warehouse conceptual data model*

Data processing and analysis

The most important part of the developed early warning system will be the data processing and analysis component.
Only a few references can be found for disease control in unheated vegetable greenhouses and most of them are based on outside weather (Jewett & Jarvis 2001). Yunis *et al.* (1994) studied the effects of air temperature, relative humidity and canopy wetness on grey mould in cucumber greenhouses. It was concluded that the temperature effect was more significant than relative humidity or leaf wetness, which was attributed to the wet winter season.
Elad *et al.* (1992) studied the epidemiology of grey mould in cucumber greenhouses. They made an attempt to construct quantitative models relating the percentage of infected fruits with microclimate parameters, but the results were unsatisfactory.
However, a qualitative approach allowed the development of a model to predict grey mould epidemics based on daily averages over the week preceding the disease observation. The durations of wet foliage and temperatures between 9 and 21°C during the night were found to be the most significant factors.
Yunis *et al.* (1994) developed a model for predicting outbreaks of grey mould in cucumber greenhouses using outside weather data.
Outbreaks of grey mould occurred following weeks when the average period of leaf wetness exceeded 7 h day-1 and night temperatures were between 9 and 21°C for more than 9.5 h day-
Baptista (2007) developed and validated a Botrytis model (BOTMOD) for unheated tomato greenhouses based on inside climate data. It was presented a practical application, disease risk levels were defined as a function of the time duration with RH > 90%.

These results enabled to make a qualitative analysis concerning the risk of infection with B. cinerea causing grey mould on a tomato crop: HIGH RISK, RH > 90% for more than 9 h day-1; MODERATE RISK, RH > 90% for periods between 4 and 9 h day-1; LOW RISK, RH > 90% for less than 4 h day-1 or RH < 90%.
These were the rules that were implemented on the information system analytical procedures and will support the conversion of the data collected by the wireless sensor network into information to be delivered to the users, as we will see next.

Information delivery

The user interface is the key component of the system since it will be the interaction contact point of the user with the early warning system. We include in this area the digital dashboards and information transmission tools that offers the users an integrated and comprehensive vision of the early warning system metrics, trends and exceptions, combining information from multiple sources.
In the particular case of the digital dashboards they offer an intuitive graphical visualization similar to a car dashboard. The secret of dashboard design consists in capturing the metrics and the performance indicators that, when compared with the actual performance and combined as graphics, give us a picture of the process under analysis "health condition".
An information dashboard is a single-screen display of the most important information a certain person needs to do is job, presented in a way that allows him to monitor what's going on in an instant, is a powerful new medium of communication (Few 2006). According to this author, the fundamental challenge of dashboard design is to display all the required information on a single screen, clearly and without distraction, in a manner that can be assimilated quickly.
Amongst the characteristics we look in a dashboard we can refer (Turban 2007):

- Use of visual components (e.g., charts, performance bars, sparklines, gauges, meters, toplights) to highlight, at a glance, the data and exceptions that require action;
- Transparent to the user, meaning that they require minimal training and are extremely easy to use;
- Combine data from a variety of systems into a single, summarized, unified view of the business;
- Enable drill-down or drill-through to underlying data sources or reports;
- Present a dynamic, real-world view with timely data refreshes, enabling the end userto stay up-to-date with any recent changes in the business;
- Require little, if any, customized coding to implement, deploy, and maintain.

In the present case the information delivery will support two different approaches. In one hand in a push logic mobile telephony short messages service and e-mail will be used to send alerts to the user when any event considered not normal is detected. On the other hand and in a pull logic the access to the information system can be done in a web friendly user interface for two different platforms desktop/laptop computers and personal digital assistants (PDAs).
The main goal of this web interface is to make available a set of key performance indicators (KPIs) displayed in a digital dashboard where in a glance we can get a global and unified picture of the greenhouse status. Besides giving this instantaneous graphical and easy to

capture vision of the greenhouse's health in real time, some selected KPIs are presented with context information, namely information about what is considered to be a good, average and bad reading for the metric under analysis, in this case information about mould disease risk (low, moderate or high). Finally it also supports drill-down operations in order to present more detailed information allowing time granularity manipulation for example if we select a specific day we can see the hourly evolution of the KPI over that specific day.

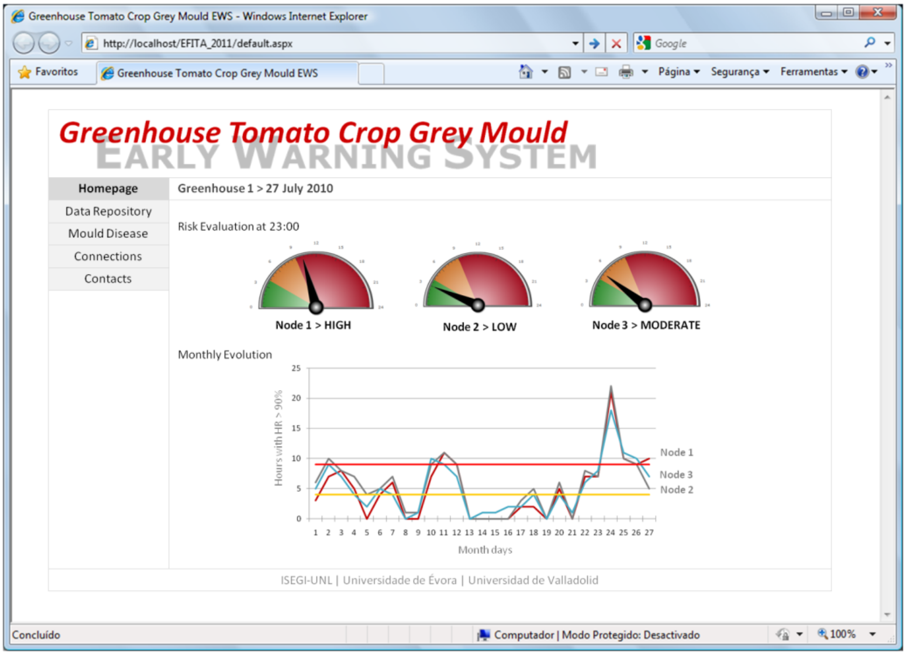



*Figure 4. Information delivery web interface*

Following the above referred digital dashboard approach we have in single screen information covering KPI present state and evolution over time. In the case of PDAs use in the field and if a wireless local/Internet connection is available it supports the information system access directly with a web interface adapted to this type of devices.

## Conclusions and Future Work

Microclimatic parameters have been recognized as key factors in the development of diseases caused by fungal pathogens on aerial plant surfaces. The study of their effects has been used to develop risk prediction models and warning systems mainly for fieldcrops in order to help the grower. In a greenhouse environment, the grower has some ability to intervene on the regulation of climatic parameters and the availability of epidemiological models can help and be useful to limit the occurrence of the conditions favourable to disease development.

Disease warning and integrated control systems are management decision aids that could help growers to apply chemicals more efficiently and economically than traditionally. It results in substantial reduction of spray frequency, which contributes to the reduction of the production costs, impact of pesticides in the environment and can delay the occurrence of fungicide resistance.

The early warning system prototype developed for greenhouse tomato crop mould disease presented in this paper demonstrated the validity of the proposed approach but now it is necessary to carry out field experimentation to validate the obtained results in different locations and distinct greenhouse conditions before making the information system available for public usage.

## References


Aakvaag, N & Frey, J 2006, "Redes de sensores inalámbricos", *ABB Magazine*, no. 2, pp. 39-42.

Abreu, PE, Monteiro, AA & Meneses, JF 1994, "Response on non heated plastic covered greenhouse tomatoes during the cool season and under two different natural ventilation methods", *Acta Horticulturae*, no. 366 pp. 195-200.

Baptista, FJ 2007, Modelling the Climate in Unheated Tomato Greenhouses and predicting Botrytis cinerea infection. PhD Thesis, Évora University.

Castilla, N 2002, "Current situation and future prospects of protected crops in the Mediterranean region", *Acta Horticulturae* no. 582 pp. 135-147.

Few, S 2006, Information Dashboard Desing – The Effective Visual Communication of Data, O'Reilly, USA.

Gabinete de Planeamento e Políticas 2007, *Horticultura*, Ministério da Agricultura do Desenvolvimento Rural e das Pescas, Portugal.

Hart, JK & Martinez, K 2006, "Environmental Sensor Networks: A revolution in the earth system science?", *Earth-Science Reviews*, no. 78, pp. 177-191.

Hill, J, Szewcyk, R, Woo, A, Culler, D, Hollar, S & Pister, K 2000. "System arquitecture directions for networked sensors", *Proceedings of the 9th International Conference on Architectural Support for Programming Languages and Operating Systems*, Cambridge, MA, USA.

Hørning, A, Hansen, JP, Hansen, NF & Schultz, E 2005, "A Common Internet Platform for the Danish Agricultural Advisory Service", *Proceedings of EFITA/WCCA 2005 Joint Conference on IT in Agriculture*, Vila Real, Portugal.

Intanagonwiwat, C, Govindan, R & Estrin, D 2000, "Directed diffusion: A scalable and robust communication paradigm for sensor networks", *Proceedings of the 6th Annual International Conference on Mobile Computing and Networks*, pp. 56–67, Boston, ACM Press.

Jensen, AL & Thysen, I 2004, "Agricultural information and decision support on smartphone", *Proceedings of AFITA/WCCA*, Bangkok, Thailand.

Jewett TJ & Jarvis WR 2001, "Management of the greenhouse microclimate in relation to disease control: a review", *Agronomie*, no. 21 pp. 351-366.

Navas, LM, Bravo, E, Ruiz, G, Correa, A, Gómez, J, González, F, Baptista, F, Cruz, V & Neto, M 2010, "Development of a SCADA System for Access, Processing and Supervision of Data Coming From a Wireless Sensors Netowork in Agro-Environmental Applications", *Proceedings of the XVIIth World Congress of the International Commission of Agricultural Engineering (CIGR) Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)*, Québec City, Canada.

Pokorny, J 2006, "Database architectures: Current trends and their relationships to environmental data management", *Environmental Modelling & Software*, no. 21, pp.1579-1586.

Pottie, G & Kaiser W 2000, "Wireless integrated network sensors", *Communications of the ACM*.

Prieto R, Rodrigues S & Henriques S 2003, "Protecção integrada de tomate em estufa", Relatório Projecto Agro 4 - Desenvolvimento de técnicas de produção integrada na horticultura protegida e de ar livre na região Oeste.

Turban, ER, Sharda, R, Aronson, J & King D 2007, *Business Intelligence* – A Managerail Approach, Prentice Hall, New Jersey.

Wang, N, Zhang, N & Wang, M 2006, "Wireless sensors in agriculture and food industry—Recent development and future perspective", *Computers and Electronics in Agriculture*, vol. 50, pp. 1-14.
Watson, HJ 2009, "Tutorial: Business Intelligence – Past, Present, and Future", *Communications of the Association for Information Systems*, vol. 25, article 39.
Yunis H, Shtienberg D, Elad Y & Mahrer Y, 1994, "Qualitative approach for modelling outbreaks of grey mould epidemics in non-heated cucumber greenhouses", *Crop Protection*, no. 13 pp. 99-104.
Zazueta, FS & Vergot, P 2003, "Use of handheld computers in agricultural extension programs", *Proceedings of EFITA2003*, Debrecen-Budapest, Hungary, July.
Zhao, F & Guibas, LJ 2004, *Wireless Sensor Networks*, Morgan Kaufmann.

# Towards a new theoretical framework: exploring the dynamics of using ICT for farming purposes

Csótó Mihály
*Budapest University of Technology and Economics - Information Society Research Institute (BME–ITTK)*
E-mail: csoto.mihaly@ittk.hu

## Introduction

Information has a central role in our modern way of living and agriculture is no exception: success in farming requires gaining, processing, using and evaluating a huge amount of information (policy, markets, new methods etc.). Farmers are working in an information-intensive environment and numerous studies have showed that information and communication technologies (ICT) can play a vital role in realizing benefits with more effective information management in the farm level.

The supply chain in agriculture not only means the flow of products and income but also that of information (Niederhauser et al, 2008). The toolkit of the information society offers new opportunities for efficient operation, decision-making and adaptation to the environment (Herdon, 2009). This, however, can only work successfully if certain conditions are met. The potential opportunities can only be exploited to an optimum under the right circumstances, and it is also important that they are harmonised with the previously used farm management practice, or else the much-awaited success will suffer.

## Information flow and the farm manager

Modelling the agricultural decision-making process and information flow is nothing new. The primary objective of studies aimed at creating models is to determine why a given group of farmers behave the way they do (Gladwin, 1989). With the help of models, support and knowledge that contribute to more efficient operation and better decision-making, can be provided for farmers.

Information and knowledge are inseparable, since the most important input in decisions originates from the stratum of information (i.e. given potentials, already acquired knowledge, and finding additional information necessary for making decisions). Information flow, or the lack of it, as well as its intensity and knowing about the circumstances of a given farm and its manager might provide a comprehensive picture about the decision-making process and thus the character of a farm.

Sörensen *et al.* (2010) conducted interviews and drew the following information cloud around the farm manager, a so-called rich picture of factors that influence and shape decisions (Figure 1.).

Important factors for the decision-maker of a farm, to find his way in the above information environment, not only include the numerous opportunities for interaction and information sources, but also the used channels and the quality of information.

In a traditional approach information channels can be grouped into:

- Personal network (suppliers, extension service, other farmers, markets, events)
- Printed media (books, encyclopaedias, maps, almanacs, newsletters, advertisements)
- Electronic media (TV, radio, websites, databases, services on cell phone, remote sensing)

In the last group the Internet must be highlighted as a special entity that can equally be a number one source (databases, consultancy web services, electronic administrative services) or a secondary intermediary of printed or electronic media (e.g. IPTV, video TV) and a means of personal communication.

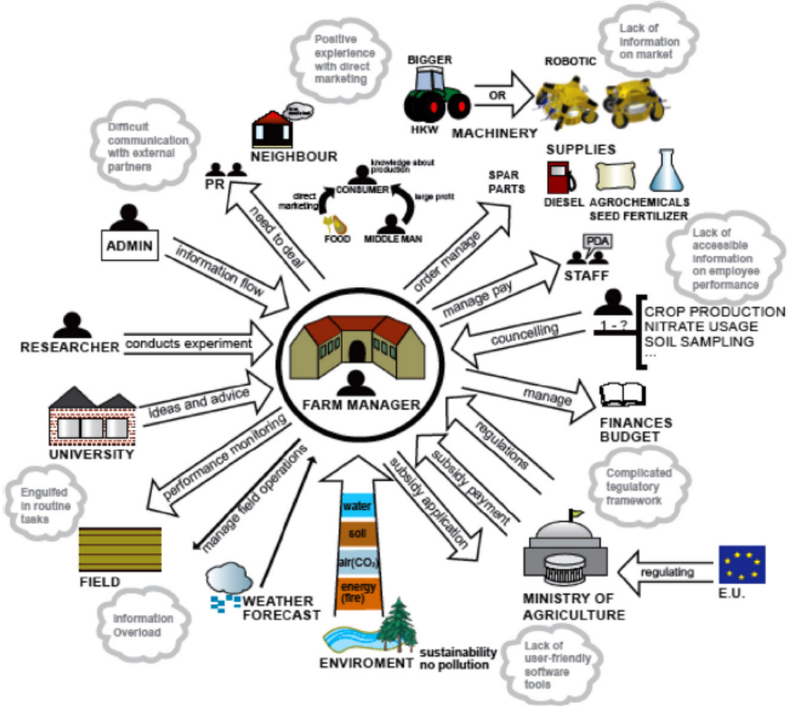



*Figure 1. Farm management, the "rich picture" (Sörensen et al, 2010)*

In addition to the type of channel used, the quality of information is of vital importance since its "usefulness" is determined by several criteria which are mostly determined by the perception of the decision-maker. There are several types of knowledge and information that are irrelevant under given circumstances (the diverse social, economic and environmental conditions of the rural communities might require specific knowledge/ and competences (Knight & Wilkin, 2004)), and this can be a problem, for example with decision support systems based on generic models possibly offering solutions that are varying from the farmers' own experience. Generalised information is often unsuitable, while relevant knowledge is often inaccessible to farmers since it is not converted into a consumable form, or an accessible information system. There might be asymmetries in the supply chain between the knowledge of the various players, which can even be intentional.

The main characteristics of quality information are therefore relevance, accuracy, comprehensiveness and timeliness. Of these timeliness/up-to-datedness is regarded as being the most important by many people: information that arrives late is often useless, even if it would otherwise satisfy the needs of the farmer. The timing of information is of critical importance especially in the case of daily or other short-term decisions, but less so during strategic, long-term planning.

The survey of Yongling (2004) highlights the importance of requirements towards information, stressing that poor quality, outdated, inaccurate or incomplete information poses a problem, mainly because farmers cannot distinguish between "good" and "bad" information, which the author mainly links with a low level of education.

**Dealing with information – an everyday activity or a special challenge?**

Models used in the area of agricultural informatics confirm that in a "text book type" farm, the farm manager approaches his information needs, the cost of the information and alternative sources at system level and in a systematic way, and is able to recognise information necessary for decision making. Furthermore, he consciously applies and builds the latest information technologies into his everyday work. Information literacy is the skill to access information and use it for value creation, and an individual can be regarded as information literate if he recognises when he needs information. Information literacy is possessed by those who have learnt how to learn. It can be seen that information literacy requires and demands several skills from the conscious citizens of the digital world. Never before in human history have these skills been expected of everyone; such expectations typically concerned the members of the intellectual elite (Rab, 2008). However, since the world of today is determined by the presence, lack, value and authenticity of information and the speed of its flow, these skills have become important for everybody, including those working in agriculture.

As Hill (Hill, 2009) states: "As individuals, farmers have their favoured information sources, which they use depending on the specific information being sought. The amount of information collected depends on the complexity of the task and the importance of the decision." The context, the channel, the type of information, the access to information and the farmer's experience and personal preferences are in close relation. Therefore, information flow must be studied in its complexity, at different levels and in detail. (e.g. Fountas et al, 2006; Nash et al,2009). The significant differences between European farms (with regard to size, type, culture and numerous other factors) each influence decision making processes and information strategies (e.g. Öhlmer et al, 1998).

As a rule of thumb, based on Nuthall (2004) it can be stated that with great probability farmers will keep seeking and collecting information until they actually feel that the cost incurred by continuing the search will exceed the benefits that can be secured by the information attained. However, this "marginal utility" approach is made more complex if we consider the knowledge base and experience of a farmer at any given time (e.g. if we apply this logic young farmers might need more information). At the same time it must be remembered that since it is difficult to attach economic value to information, the search process will mostly depend on the perception of the farmers, which means that the proportion of intuitive decisions is likely to increase significantly with experience (at least under the given circumstances).

**The diffusion of ICT**

Harkin's research (n.d.) shows that the diffusion of ICT opens up the following opportunities. Information technologies are regarded in the following list as mediating channels and a vehicle for new services i.e,.

- Timelier, more comprehensive information
- Accessing a new type of information
- Several competing information sources
- Single window access
- Easy exchange and discussion of information ideas
- Easier cooperation and access to other farmers and experts

Despite these opportnities, ICT take-up in agriculture remains low. It is important to research the various aspects of adoption of new information technology, and also not only the use, but the "effective use" of it (Gurstein, 2003). It is important to distinguish between the opportunities for digitally-enabled activity presented by ICT access and the actual realization of those opportunities in the form of "effective use".

According to on Harkin, Offer (Offer, 2005) and others, it can be concluded that since the late 1970s and early 1980s service and technology developers have been making the same mistake, i.e. they fail to satisfy user needs, which is a key challenge in the area. Access to information/knowledge must be made more flexible. This should include gathering information about the potential users' needs and problems as well as offering them consultation opportunities. This problem is not specific to agricultural IT developments. It spans the close to ten years (i.e. almost the entire period) of European eGovernment development projects, for example.

Farmers have not yet been encouraged as obviously and successfully, to change their approach to information management as it could have been expected based on the technical "potentials" (Alvarez and Nuthall, 2006). Nuthall also points out that, since the farmer is an essential component in the "information system" of a farm, the decision to use ICT often depends on the personal characteristics of the farmer, such as his or her personality, experience, age, education, goals and objectives. The main factors cited in most of the literature on the problems of adopting ICT among farmers only include the typical characteristics of smallholders.

Mainly based on Apps and Iddings (1990) as well as Taragola and Van Lierde (2010), a more detailed picture can be drawn:

- Farmer's characteristics (age, experience, personality, education)

- Farm characteristics (size, type, geography)
- Goals and objectives (attitude towards learning)
- Decision making and information management style (time, information sources ( number, intensity in use), extension usage, support from the outside)
- Trust
- Community culture (network, associations)

Surprisingly enough, the factors influencing ICT use have barely changed in the last 30 years, since personal computers have become commercially available. One can find more or less the same old stories regarding computers, decision support systems (DSS) and internet-usage: the background of the farmer (demographic and other variables), the characteristics of the farm (size, complexity etc.), the cultural background and the local community. Every farm is different and the information strategy of every farmer is different. So it is really hard to find solutions that fit in with the processes of the majority of the farms and farmers.

Earlier studies, the aims of which were to discover the factors related to ICT-use among farmers, mainly concentrated on the use of the tool itself. If the problem is considered from a different perspective, the best way to describe it is within the framework of information and communication: the information management style of the farmer, the communication network of the farmer, the channels and flow of information and its intensity.

Information flow, or the lack of it, as well as its intensity and being informed about the circumstances of a given farm and its manager could provide a comprehensive picture about the decision-making process and thus the characteristics of a farm. Recently we have found models that contain the well-known factors and also "communication" as a variable (Figure 2.). Therefore, information flow must be studied in all its complexity, at different levels and in detail. If we do so we will find ourselves in a new territory: network-research.

**"The surprising power of social networks" of farmers?**

In the last decade or so, many authors (mainly with a background in the natural sciences, e.g. Barabási, Watts, Buhannan, Christakis and Fowler) have written famous bestsellers about social networks and how they are shaping people's everyday lives. However, this is not a new phenomenon in the literature of the diffusion of innovation. One of the best descriptions comes from Valente (1995): "*The diffusion of innovation occurs among individuals in a social system, and the pattern of communication among these individuals is a social network. The network of communication determines how quickly innovations diffuse and the timing for each individual's adoption.*" In other words: a chain of innovation is a social network, defined by offering or adopting an economic innovation (Letenyei, 2001).

The idea of networking in agriculture innovation suggests that the farmers' position in social networks determines their access to resources needed to implement the innovation as they are part of networks where people exchange information, discuss experiences of adoption and find solutions for a better adaptation in specific contexts (Hartwich and Scheidegger, 2010).

| Category | Variables |
| --- | --- |
| 1. Adoption of innovation, technologies, knowledge etc. | Use of an innovation (yes or no)<br>Use of an innovation package (from 0 and 100%)<br>Share of land to which the innovation is applied<br>Time lag between persuasion and adoption |
| 2. Access to and endowment with resources | Household or farm size (property and rented)<br>Wealth of household or farm<br>Access to credit<br>Knowledge and technical capacity |
| 3. Socio demographic factors | Education<br>Literacy<br>Gender<br>Age |
| 4. Socio-psychological behavior | Risk aversion<br>Attitude to credit<br>Attitude to change<br>Attitude to external information<br>Exposure to mass media |
| 5. Communication | Contacts to other farmers<br>Contacts to research, extenionists and other change agents<br>Exposure to media<br>Opinion leadership |

*Figure 2. Frequently used adoption predictors (Hartwich and Scheidegger, 2010)*

The characteristics of the diffusion are determined by a number of factors (collected by Letenyei, 2001):

- Granovetter (1973) argued that weak ties (people loosely connected to others in the network) were necessary for diffusion to occur across subgroups within a system.
- Burt (1987) argues that structural equivalence (the degree of equality in network position) influences the adoption of innovations.
- Other personal and social network characteristics which might influence the diffusion of innovations include centrality, density and reciprocity (Valente, 1995)

The network-concept raises a very interesting question: if we accept that innovation and adoption is a communication process how can we evaluate the spread of information and communication technologies?

**The double edged sword: an object or an enabler of adoption?**

As observed earlier, mobile phones and the Internet are not simple tools: they can form a primary source, a secondary intermediary of printed or other electronic media and a means of personal communication. Just to make things more complex, they can be combined

(speaking on a mobile ph one with the extension service employee who is using an online tool to help the farmer).
The personal and some of the electric channels (unlike TV and radio) can be used for two way interaction, and by so doing the (electronic) services (banking, making claims for farm subsidies and so on) can also be involved in the model. Personal/oral information is as two-faceted as the electronic kind: a person (family member, other farmer, extension service employee etc.) can be the source of information or the intermediary as well. An intermediary can also modify the information (or can put it into context thus helping it become knowledge).
It is clear that on the one hand ICT itself is an object of adoption: new and innovative tools to make something better and more efficient. On the other hand the thing is that that 'something' is communication which is at the forefront of any diffusion process as it can also be an enabler.
The real question is how the farmers' social network and personal communication changes with the use of ICT.

**More food for thought – further questions to ask**

It must be stated that a model based on this approach cannot be fully comprehensive. This happens because communication is not the only way to obtain relevant information. Even without a sound information background it is possible to reconstruct meanings identical to those produced with the help of such a background. It is also proven that the same information can be collected in various ways (from whole different 'pieces') (Z. Karvalics, 2003). It must not be forgotten that the context, the channel, the type of information, the access to information and the farmer's experience and personal preferences are closely linked. In addition, the farming environment is crucial: tacit knowledge can also play a role (Polányi, 1993) and make mapping somewhat more difficult. These difficulties do not mean that this approach should be rejected. Mapping the flow of information and the social network of the farmers can be used to make more targeted actions to stimulate the use of ICT technologies for more effective farming, because it can demonstrate the intensity of the flow – or the lack of it
The timeframe is another crucial element. Because of the cumulative nature of communication (that is why gaining information is not always carried out in order to gain new knowledge but to refresh or verify that which already exists) while the continuous adaptation to the environment, the various elements of the flow of information (the dominant channel, the intensity of the communicative moment, or its periodicity) are undergoing constant change. The farmer is always evaluating the scale of energy input required for the acquisition of information and whether it is in proportion to its importance. So mapping these changes also lead to more valuable results than a simple picture which lacks the fourth dimension.

**Conclusion**

Nowadays we know a lot about the factors affecting the use of ICT. An approach that uses a normative picture of these technologies means, that ICT automatically makes farming more efficient. But this has not yet been realized. Improving the efficiency of farming by the means of ICT needs a more detailed approach and this can be achieved by making communication and information flow more efficient. Using or not using (ICT) is no longer an issue. The appropriate questions are rather: who uses ICT, how do they use it, for what, and in what context.

If ICT use is not coupled with critical thinking and a clear concept, the desired results and a positive impact on farm management cannot be achieved, what is more, damage can be caused as happened in the case of industrial methods adopted without a critical approach. One of the key issues with regard to the diffusion of IT applications is the assessment of the information needs at any given time (Rockart, 1979). Only by targeting these needs (possibly after latent needs are explored and defined) can developments bring success. If this is accompanied by easy usability, trust and economic advantages, it can be expected that these applications and tools will be used, however, the characteristics of the given farm and the personal characteristics and attitude of the farm manager and especially his/her network will play a significant role.

Understanding the take up of innovative information technologies and their role in farm management needs some new methods, because the known factors do not fully describe the real situation. Studying the dynamics of networking contributes to a better understanding of how and when smallholder farmers make decisions to adopt innovations. Analysing the network structure in a given field has to be the basis for any kind of development of local and regional economic activities.

## References


Alvarez, J., Nuthall, P. 2006. Adoption of computer-based information systems: The case of dairy farmers in Canterbury, NZ, and Florida, Uruguay. Computers and Electronics in Agriculture 50(1) pp. 48-60.

Apps, J., Iddings, R. 1990. What influences farmers' computer use? J. Ext. 28 (1) pp. 19–21.

Burt, R. S. 1987. Social contagion and innovation: cohesion versus structural equivalence. American Joumal of Sociology Vol. 92, No. 6. (1987), pp. 1287-1335.

Fountas S., Wulfsohn D., Blackmore B.S., Jacobsen H.L., Pedersen S.M. 2006. A model of decision-making and information flows for information-intensive agriculture Agricultural Systems 87 (2006) pp. 192–210.

Gladwin C.H. (editor) 1989. Ethnographic Decision Tree Modeling, Sage Publications

Granovetter, M. 1973. The strength ot weak ties. American Jornal of Sociology Vol. 78, No. 6. (1973), pp. 1360-1380.

Gurstein, M. 2003. Effective use: A community informatics strategy beyond the Digital Divide In: First Monday, Volume 8, Number 12 - 1 December 2003 Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1107/1027 Accessed 5 January 2011

Harkin, M (no date): ICT Adoption as an Agricultural Information Dissemination Tool – An historical perspective http://departments.agri.huji.ac.il/economics/gelb-harkin-3.pdf Accessed 10th September 2010.

Hartwich, F., Scheidegger, U. 2010. Fostering innovation networks: the missing piece in rural development? Rural Development News 1/2010. Agridea Available at: http://www.agridea-international.ch/fileadmin/10_International/PDF/RDN/RDN_2010/14_innovation_networks.pdf

Accessed 20th October 2011.

Herdon, M. 2009. Impacts of e-collaboration tools for development of rural areas. In: AVA Congress 4: International Congress on the Aspects and Visions of Applied Economics and Informatics. Nábrádi A., Lazány J., Fenyves V. ed. Debrecen, Magyarország, 2009.03.26-2009.03.27. Debrecen: Agroinform Kiadó, 952-959.(ISBN:978-963-502-897-9)

Hill, M. 2009. Using farmer's information seeking behaviour to inform the design of extension, Extension Farming Systems Journal 5/2 (2009) pp. 121-126.

Knight, J.D., Wilkin, D.R. 2004. A novel method for the delivery of information on the storage of malting barley Computers and Electronics in Agriculture 44 (2004) pp. 133–144.
Letenyei, L. 2001. Rural Innovation Chains. Two examples for the diffusion of rural innovations Review of Sociology Vol. 7 (2001) 1, pp. 85-100.
Nash, E., Dreger, F., Schwarz, J., Bill, R., Werner, A. 2009. Development of a model of data-flows for precision agriculture based on a collaborative research project. Computers and Electronics in Agriculture 66 (1) pp. 25–37.
Niederhauser, N., Oberthu T., Kattnig S., Cockj. 2008. Information and its management for differentiation of agricultural products: The example of specialty coffee In: Computers and electronics in agriculture 6 1 (2008) pp. 241–253.
Nuthall, P.L. 2004. Case studies of the interactions between farm profitability and the use of a farm computer In: Computers and Electronics in Agriculture 42 (2004) pp. 19–30.
Öhlmer, B., Olson, K., Brehmer, B. 1998. Understanding farmers' decision making processes and improving managerial assistance. Agricultural Economics 18 pp. 273–290.
Polányi, M. 1992. Filozófiai írások I-II. (Philosophical essays I-II.) Atlantisz, Budapest, 1992, (ISBN 963-7978-275)
Rab, Á. 2008. Digital culture – Digitalised culture and culture created on a digital platform In: Information Society From Theory to Political Practice. Coursebook, pp. 183-201. Pintér R. ed. Gondolat – Új Mandátum, Budapest
Rockart, J. 1979. Chief Executives Define Their Own Information Needs. In: Harvard Business
Sörensen C.G., Pesonen L., Fountas S., Suomi P., Bochtis D., Bildsře P., Pedersen S.M. 2010. A user-centric approach for information modelling in arable farming In: Computers and Electronics in Agriculture 73 (2010) pp. 44–55.
Taragola N.M., Van Lierde D.F. 2010. Factors affecting the Internet behaviour of horticultural growers in Flanders, Belgium Computers and Electronics in Agriculture 70 (2010) pp. 369–379.
Valente, Th. W. 1995. Network Models of the Diffusion of Innovations New Jersey: Hampton Press
Yongling, Z. 2004. Information service in rural China field surveys and findings. Bangkok: FAO.
Z. Karvalics, L. 2003. Információ, társadalom, történet (Information, society, history) Typotex, 2003, Budapest (ISBN 963-9326-91-7)

# Data quality concepts and methods applied to biological species occurrence data

Allan Koch Veiga, Antonio Mauro Saraiva, Etienne Americo Cartolano Junior
*University of Sao Paulo, School of Engineering, Computer Engineering Dept., Agricultural Automation Laboratory*
E-mail: allan.kv@usp.br

## Introduction

Establishing a balance between economic and social development and environment and biodiversity conservation is a non-trivial mandatory challenge of the modern society. Strategies are needed to enable development without affecting future generations, especially when it comes to environmental damage, which can be irreversible (Brundtland, 1987).
In the midst of this global biodiversity crisis of yet unknown magnitude, the scientific community has joined efforts to coordinate the dissemination and use of biodiversity datasets. The availability of these data is an important resource for environmental management (GBIF, 2010).
Species occurrence records are widely used in many studies on conservation and sustainable use of natural resources (Chapman, 2005a). Species occurrence can be defined as an observation or collection of an organism in a particular point in space and time. From this type of data one can carry out analyses and create models that can be used to support strategic decision making (Chapman, 2005b).
However, in order to obtain accurate results data should be of high quality. The indiscriminate use of such data without considering possible errors, can lead to erroneous results, misleading information and, consequently, to decisions that can adversely affect the management of the environment (Chapman, 2005b).
Therefore, Data Quality (DQ) in the processes of production, collection, analysis and dissemination of data is an important factor in many industries that use them, as in agriculture, for example.
In this chapter[15] we discuss concepts of DQ applied to species occurrences data. Based on these concepts we present methods that can help reduce errors that negatively affect some aspects of DQ in the process of data digitization. These methods are being deployed and implemented in a web-based Information System (IS) of species occurrence data management, called Biodiversity Data Digitizer (BDD) (Cartolano *et al.*, 2010; Saraiva *et al.*, 2011).

## Species occurrence data

Observations of nature are fundamental to studies that will result in scientific publications, biodiversity management plans and conservation policies. The sharing and interoperability of data are needed to improve conservation of the global biodiversity. In many cases complex researches that are performed in this area require the availability of species occurrence data in

[15] Based on Veiga et al. (2011a)

a broader geographic and temporal scale (Chapman, 2005a). In order to achieve it multiple data sources are used to integrate a more complete database (Kelling, 2008).

However, for different data sources to interoperate many efforts have been undertaken by national and international initiatives. One such initiative is the Biodiversity Information Standards (TDWG), which has developed standards for management, integration and access to information on biodiversity (Kelling, 2008). One of the most important standards is the Darwin Core Metadata Schema (DwC), which enables the exchange and interoperability of species occurrence data.

The DwC schema (http://rs.tdwg.org/dwc/terms/) is organized into subsets of related data elements, which can be called data domains (Dalcin, 2005). DwC is composed of seven data domain standards. Other specific data domain, called extensions, can be added to the standard schema to complement it, if necessary. Among those domains, three stand out in importance because they are essential for the development of biological species distribution models:

- Location data domain: These data refer to geographical information of the species occurrence. The elements of this data domain are, among others, country, state or province, municipality and locality.
- Geospatial data domain: Spatial data are related to a georeference. They are composed by the attributes of latitude, longitude and altitude, among others, and provide further detail to the occurrence location.
- Taxon data domain: The domain of taxon data comprises information on nomenclature and taxonomic hierarchy of the organism observed or collected. Kingdom, phylum, class, order, family, genus and scientific names are part of this data domain that defines the classification of the species.

Both geospatial and location domains represent the occurrence location, but in different ways. They are important components of occurrence data, since they are mandatory for many studies and computer models, such as species distribution models (Chapman, 2005a). These models can be used, for example, to develop scenarios for a species in relation to environmental changes (GBIF, 2010).

Taxonomy, or Systematics, is the theory and practice of classifying organisms (Chapman, 2005b; Dalcin, 2005). Classification is the creation and definition of hierarchical systematic groups of organisms known as taxa (taxon, in singular). The nomenclature in this context is the assignment of names for each taxon (Dalcin, 2005). Every known living organism is part of a taxon. Therefore, each occurrence should receive identification. An occurrence without taxonomic identification is not useful in most studies on biodiversity (Chapman, 2005a). The taxon data domain in Darwin Core refers to the classification and nomenclature of biological organisms.

Thus, given the myriad of applications that make use of these data, it is essential to ensure its fitness for use from the point of view of quality (GBIF, 2010; Chapman, 2005a). It is therefore of paramount importance to develop policies, methods and tools to measure and improve DQ.

## Concepts of Data Quality

It is widely accepted that quality can be defined as compliance with requirements. It means that the concept of quality changes as requirements change. This definition can also be applied to the concept of DQ (Wang *et al.*, 1993).

Regarding the various definitions of DQ, it is a consensus that the concept of quality can only be defined when the data are used for some purpose (Dalcin, 2005). A widely accepted definition is that high quality data are fit for use (Strong *et al.*, 1997), in other words, the data should serve the purposes of the user.

Thus, by definition, DQ studies should be based on the application context. These studies can be conducted with at least two approaches: DQ assessment and DQ management (Ge & Helfert, 2007).

DQ assessment can be defined as the process of assigning a numeric or category value to a DQ aspect in a given context. The DQ management approach is to apply methods and techniques in order to improve the DQ, considering the assessment approach.

The following section describes the methodology used to conduct the assessment and management in location, geospatial and taxon data domains.

Data Quality Assessment

Related literature explores three main components in DQ assessment (Ge & Helfert, 2007):

- DQ problems;
- DQ dimensions;
- DQ methodology;

These components can be organized in three layers as depicted in Figure 1.

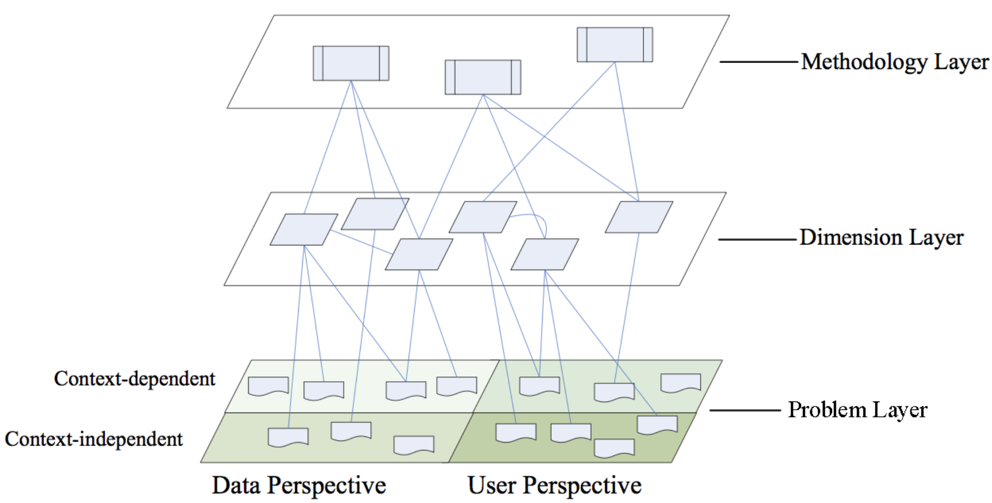



*Figure 1. Components for DQ assessment. Based on Ge & Helfert (2007).*

The problem layer refers to error patterns found in the context of the IS. The problems can be classified according to their context-dependence or context-independence and according to either a data perspective or a user perspective. Each problem can affect one or more DQ dimensions.

DQ is defined in the literature as a multidimensional concept (Pipino *et al.*, 2002), in which each dimension represents an aspect of quality. These dimensions allow the measurement and management of DQ in a more objective and specific way. The value of each DQ dimension is measured by means of the elements of DQ assessment methodology layer.

The methodology layer consists of models, frameworks and methods used to measure the DQ dimensions according to a specific data domain.

Accordingly, to assess the DQ in a specific context it is necessary to identify which elements of each layer will be evaluated, what they mean in each IS context and how these elements relate to each other. Thus, the following section describes the DQ assessment study in the context of the Biodiversity Data Digitizer information system.

Problem Layer

The error patterns proposed by English (1999) were applied by Dalcin (2005) to taxonomic databases:

- Domain value redundancy: occurs when data values are not standardized or are synonyms, in other words, when two or more different values represent the same thing in the real world.
- Missing data value: occurs when required data are missing. This includes both mandatory and non-mandatory fields (in terms of a data standard), which are required to be used in certain applications.
- Incorrect data values: these errors can be caused by transposition of characters at the time of typing, by entering data in the wrong places, for not understanding the meaning of information, or when an information system requires inclusion of data that is currently unknown.
- Non-atomic data values: occurs when data fields that should contain atomic values contain multiple values.
- Domain schizophrenia: occurs when fields are interpreted and used in different ways depending on context.
- Duplicate Occurrences: Occurs when multiple records are a single entity in the real world.
- Inconsistent data values: the inconsistencies may occur due to the heterogeneity of standards and management procedures adopted by different institutions, collections, or individuals. These errors are defined by contradictions.
- Information quality contamination: contamination of information quality occurs when using incorrect data in the production of new data, thus propagating the errors.

These error patterns can have different meanings in different data domain (Dalcin, 2005).

In the location data domain, errors in domain value redundancy and duplicate values can be related to the language in which the information was digitized. For example, Brasil (Portuguese) and Brazil (English), both are correct and refer to the same entity in the real world, however, are different data. Incorrect data values in the location data domain are often caused by spelling mistakes. Non-atomic data values can occur when the field of municipality, for example, is typed "New York, NY". Information quality contamination happens when a data with error is reused to produce new data; for example, reuse the name of the municipality "New York, NY" to register a new record.

In the geospatial data domain, error of missing data value is very strongly related to the fields of latitude and longitude. The absence of one of these fields usually has the same effect as the absence of both fields, as together they represent the geospatial coordinate. Typing errors are also common in this data domain. The transposition of a comma or absence of a minus sign (where there should be) in the fields of decimal latitude or longitude, for example, are errors of incorrect data values. The insertion of latitude and longitude in the same field, for example, "Latitude: -23.834, -59.984" is an error of type non-atomic data values. Domain schizophrenia and inconsistent data value errors can occur when geospatial coordinates are filled in the form of degrees (minutes, seconds) in fields where they should be filled in decimal format.

In the taxon data domain, errors in domain value redundancy and duplicate values can occur due to the fact that the nomenclature of the taxa can change over time, and so synonymous arise. For example, a class may be named as Insecta or Hexapoda. These two names are synonymous and represent the same entity in the real world.

The problem of missing value is very common in more specific taxa such as scientific name. This is because the identification at more specific levels of taxonomic hierarchy is the most difficult task that may require experience and specific knowledge. Thus, these data are omitted when their identification is doubtful.

The error of incorrect data values is very common in this data domain and is caused by spelling mistakes. The fact that taxon names are written in Latin can contribute to increase the amount of data errors in this data domain.

Non-atomic data values may occur, for example, on inclusion of synonyms for a taxon in the same field. One example might be entering the name of a class as "Insecta, Hexapoda".

The domain schizophrenia error occurs when a field is used for a purpose for which it was not designed; in this case an example could be using the field scientific name to register "sp1" to indicate a morphospecies. Inconsistent data values can be related to not using nomenclature patterns and taxonomic hierarchies or the inadequacy of the digitized data to the standard adopted.

The error information quality contamination occurs when a data with any of the errors mentioned above is used as a basis to produce a new data or register.

Since these errors patterns can be frequent and are very relevant to biodiversity management, we have developed some resources that can be used to reduce errors and improve the DQ in location, geospatial and taxon data domains.

**Dimension Layer**

In this paper we cover five dimensions of DQ:

- Completeness - is a manageable and measurable dimension that indicates the sufficiency of data for use to perform a specific task (Pipino *et al.* 2002; Dalcin, 2005);

- Consistency - measures and manages the absence of contradictions in the data;
- Credibility of Source - refers to the measurement of aspects related to the reputation of the data source (Dalcin, 2005) and is used to measure how reliable the data is for their use (Wang *et al.* 1995);
- Accuracy - is considered in many DQ studies as a key dimension, and can be defined as the measure of correctness or veracity of the data (Pipino *et al.*, 2002);
- Precision - is often confused with accuracy. However, accuracy is related to errors, while precision is related to the resolution or granularity of the data (Chapman, 2005a).

In the geospatial and location data domains, completeness of data is an important factor, since the absence of some geospatial data (e.g. latitude or longitude) or geographic location (e.g. municipality name) of a species occurrence precludes the use of such data for many applications (Chapman, 2005a). The cause of incompleteness of geospatial data is often due to the unavailability of geo-resources, such as a GPS receiver (Global Positioning System), at the registration of the occurrence. One technique that can be used to obtain geospatial coordinates, when location data was correctly filled, is to use the coordinates of the center of mass of the municipality where the occurrence was recorded. However, this technique can affect the quality dimensions of accuracy and precision.

Consistency in these data domains can be identified by the absence of contradiction between geospatial coordinates data in relation to location. An example of inconsistency would be to indicate that the site of occurrence was the municipality of São Paulo, Brazil, but the coordinated geospatial refer to a position on the African continent.

In the location data domain, accuracy may be related to the correct spelling of geographical locations names, while precision may be related to the presence of data on more specific locations, such as municipality name, for example.

Regarding the geospatial data domain, the dimensions of accuracy and precision are strongly related (Dalcin, 2005). Accuracy and precision are regularly confused and their differences are often not understood (Chapman, 2005b). Figure 2 illustrates the concepts.


*Figure 2. Examples of precision and accuracy (Dalcin, 2005).*

Accuracy refers to the gap between the real value of the position and the value reported. Precision (or resolution) can be divided into two main approaches: statistical and numerical. Statistical precision refers to the relative compliance of geospatial positions of a set of occurrences. As shown in Figure 1, the positions of geospatial occurrences can be precise but not accurate. Numerical precision is relative to the amount of significant digits used to represent a position in space. For example, the decimal latitude and longitude can be represented with 10 digits (about 0.01 mm resolution), however, the real resolution is not more than 1 to 10 m (Chapman, 2005b).

The definition of quality in the taxon data domain differs considerably from the geospatial data domain, because taxonomic data are usually more abstract and harder to qualify (Chapman, 2005b). These data are the main identifiers of species occurrence. The task of making a taxonomic identification of specimens requires experience and very specific knowledge about particular groups of taxa. Often, the researcher must consult bibliographies, multimedia resources and taxonomic keys to assist in his decision. Thus, the completeness of the data depends primarily on the researcher's knowledge about the species collected or observed.

The consistency of taxonomic data is related to the lack of contradiction on definitions of hierarchies of taxa and their classification. However, each institution can adopt its own nomenclature or taxonomic hierarchy (Kelling, 2008). Even if there is this "inconsistency" in a global context, it is necessary that, at least, each institution adopts a pattern of hierarchy and nomenclature to be used in its databases. In this context, there are authorities that define standards of taxonomic hierarchies and nomenclatures. The adoption of a standard of one of those authorities can influence the credibility of source.

The accuracy of data in this data domain is related to the correct spelling of taxa names. Precision can be defined as the presence of taxa data in the most specific taxonomic hierarchy, such as genus, subgenus, specific epithet and scientific name, for example. The accuracy can also be related to the taxonomic identification correctness of a species, or to accurately report that a species belongs to a taxon X and not to a taxon Y.

Based on these concepts, we observe that the quality of a dataset depends on a number of issues (English, 1999, Wang *et al.*, 1995). A major factor is related to the data source, because the process of data digitization is inherently prone to errors (Dalcin, 2005). Accordingly, to reduce errors and improve DQ it is necessary to identify which types of errors can occur in the digitization process and how to avoid or correct them.

**Assessment Methodology Layer**

The DQ assessment methodology proposed is based on the premise that the quality in each dimension is affected by the presence or absence of problems. Nevertheless, as the problems and the dimensions have different meaning in each data domain, the elements of the methodology layer are based on the data domain. Therefore, the measurement of DQ is defined based on the presence or absence of specific errors in a specific data domain – Location (L), Geoespatial (G) or Taxon (T) – according to Table 1.

Table 1 shows which problems (rows) affect and degrade the dimensions (columns) of DQ for each data domain (T, G or L). The presence of a data domain symbol (L, G or L) in a cell indicates that the error on that row, affects the dimension of that column for the domains present in the cell. For example, the cell highlighted in grey, which has “T| - |L” shows that the error ”missing data value” affects “precision” for the domains Taxon (T) and Location (L), but not for Geospatial (-).

| | Completeness | Consistency | Accuracy | Precision | Credibility of Source |
|---|---|---|---|---|---|
| *Domain value redundancy* | - | T \| - \| L | T \| G \| L | - | T \| G \| L |
| *Missing data value* | T \| G \| L | - | - | T \| - \| L | T \| G \| L |
| *Incorrect data values* | T \| G \| L | T \| G \| L | T \| G \| L | - \| G \| - | T \| G \| L |
| *Nonatomic data values* | T \| G \| L | T \| - \| L | T \| G \| L | - | - |
| *Domain schizophrenia* | T \| G \| L | T \| G \| L | T \| G \| L | - | T \| G \| L |
| *Duplicate occurrences* | - | T \| - \| L | - | - | - |
| *Inconsistent data values* | - | T \| G \| L | T \| G \| L | - \| G \| - | T \| G \| L |
| *Information quality contamination* | T \| G \| L | T \| G \| L | T \| G \| L | T \| G \| L | T \| G \| L |

*Table 1. Matrix of DQ dimensions degradation by problems in data domains.*

The result of the DQ assessment methodology layer, represented by Table 1, shows how DQ can be improved in a given dimension and in a given data domain by reducing the occurrence of specific problems.

### Data Quality Management

It is a consensus among most experts that quality management applied to products can also be applied to data (Dalcin, 2005). There are two approaches to improving DQ: error prevention, and error detection and correction (Embury, 2001 apud Dalcin, 2005). Error prevention is considered a superior approach to detection and correction, since the latter is often expensive and does not guarantee total success, neither on the detection nor on the correction (Dalcin, 2005). We, then, present two features implemented to help prevent errors while digitizing species occurrence data through the Biodiversity Data Digitizer, BDD (Veiga *et al.*, 2011b).

### Georeferencing Data Quality Resource

To digitize geospatial and location data, a georeferencing resource can help reduce errors in these data domain. The BDD implements a web-based resource, called BDD Georeferencing Tool (BGT) (Veiga *et al.*, 2010), which allows the user to obtain the name of the country, state, city, latitude, longitude and altitude by means of a click on the desired location, through an interactive map in three dimensions. Moreover, it is also part of this tool an indicator of uncertainty/error, which allows the user to report the level of uncertainty about the geospatial information.

Using this tool we can improve the completeness of data. The user does not need necessarily to have the exact geographic coordinates of the occurrence to fill in the coordinates fields. The user can locate a known region on the map, such as a park or a mountain, for example, and use a more specific reference, such as a river or a highway to obtain location data and approximate geospatial coordinates. In some cases, this resource may contribute also to improve the accuracy and precision in the geospatial data domain; for example, when the geospatial coordinates are obtained from the center of mass of the municipality where the occurrence was recorded. Moreover, the problem of missing value is potentially reduced.

The DQ dimensions of consistency and accuracy (in the location data domain) may have a potential improvement with the use of the BGT. Considering that the geospatial data and location are automatically completed by the tool, the user is unable to enter conflicting and incorrect data, for example, "Country=Brazil, State=Rio de Janeiro, City=London, Latitude=0.0, Longitude=12:32.93". Errors of information quality contamination, inconsistent date values, domain value redundancy, incorrect date value and non-atomic data values can be reduced.

Another important aspect on the DQ in the geospatial data domain is put forward in GBIF (2010). The accuracy and the precision doe not need necessarily to be perfect (Wang *et al.*, 1995). The use of geospatial data in some applications admits low accuracy and precision. However, there are cases where the accuracy and precision of the data must be high. Thus, it is necessary to report how precise and accurate data is to assess its fitness for use (GBIF, 2010).

Aiming to this need, BGT enables its users to indicate the uncertainty of the reported values. It is possible to report how precise or accurate data are. If a user knows that a particular specimen was collected in a specific mountain, but do not know the exact geospatial position, it is possible, through an indicator of uncertainty/error, to report that the informed geospatial position may contain an error of up to 10 km, for example. Thus, it is easier to assess the fitness for use of the data. This factor allows for a possible improvement on the credibility of source.

For implementation of this tool we used PHP and Javascript programming languages. For the rendering of the map we used the Google Earth API, therefore is necessary to install the plug-in for Google Earth on the user browser. Geospatial and location data were obtained from the Google Maps API Web Services. Knowing that the data source is Google, the credibility of source can also be partially evaluated.

**Taxonomic Data Quality Resource**

The DQ in the taxon data domain is very closely linked to the compliance of taxa names and hierarchies to a particular standard, whether it is internationally recognized or adopted only internally by an institution. In order to try to improve the taxonomic DQ we are implemented a web-based resource in BDD, called BDD Taxonomic Tool (BTT). This tool aims to help users to digitize taxa data.

This resource has five features: (1) suggestions for classifications and hierarchies based on a local database, (2) suggestions for classifications and hierarchies based on a taxonomic authority, (3) support to decision making regarding taxonomic identification, (4) support digitization of morphospecies data and (5) indicator of uncertainty.

In BTT, a taxon field is presented to the user. As characters are being typed in by the user in this field, valid taxonomic name suggestions are shown to the user in a list. This feature is called autocomplete. The suggestions displayed are retrieved from the database of the institution, using a fuzzy matching technique. This technique allows retrieving textual data which is orthographically similar. For example, if the user enters Apes melifera, the system may suggest Apis mellifera (the correct scientific name for honeybee). When selecting a taxon, its hierarchy is automatically filled in accordance with the database of the institution.

If the taxon name typed does not exist in the local database, a query is performed to taxonomic authorities. The authorities consulted are the Integrated Taxonomic Information System (ITIS) (http://www.itis.gov) and the Catalog of Life (COL) (http://www.catalogueoflife.org). Also using fuzzy matching, the system suggests valid nomenclatures and taxonomic hierarchy according to these authorities.

These first two features can greatly reduce errors of redundancy domain value, incorrect data values, non-atomic data values, duplicate data values, inconsistent data values and information quality contamination. Moreover, DQ can be affected on the dimensions of consistency, credibility of source, accuracy and precision.

The features discussed below are related to the identification of specimens. The resources developed will allow the user to obtain bibliographic information, multimedia resources and taxonomic keys related to a certain species to help decide which taxon a certain specimen belongs to. This information is previously registered and related to the taxa. Since this feature can be useful for the taxonomic identification, errors of missing value can be reduced. Moreover, the DQ dimensions of completeness, accuracy and credibility of source can be positively impacted.

In some cases, there are doubts about the identification of a set of specimens, however, it is recognized that all specimens belong to the same taxon. When this occurs, these specimens may receive a temporary identification, called morphospecies, such as "sp. 1", for example. This indicates that all specimens identified as "sp. 1" belong to the same taxon though the definite specific epithet is still missing. Thus, when performing the final taxonomic identification of one individual of that set, all the other individuals that received the same morphospecies identification will consequently receive the same taxon name. Thus, in BTT, when the user types the string "sp. " followed by a number, for example, "sp. 1", "sp. 5" in the field taxa, the system automatically identifies this taxon as a morphospecies. Later, when one of these specimens is identified, the other records with the same identification will be automatically changed. This feature can cause, therefore, a reduction of errors in schizophrenia domain.

The fifth BTT feature allows the user to report the uncertainty degree about identification. For example, normally if a user has doubts about the identification of a specimen, he can take two decisions: not register, and thus decrease the completeness, or register; or, if the identification is incorrect, reduce accuracy. In this sense, an indicator of uncertainty allows the user to indicate that the information has a probability of being incorrect and needs a validation of an expert. Thus, this indicator can be used to assess the fitness for use of such data. Therefore, this feature can impact the credibility of source.

For implementation of these resources we used, as in BGT, the PHP and Javascript programming languages. The data of taxonomic authorities are periodically updated and maintained in a database on a dedicated server. The BTT retrieves the data via a web service implemented on this server.

## Conclusion

The methods and tools discussed in this chapter have the potential to greatly improve data quality in species occurrence in the dimensions of completeness, consistency, source credibility, accuracy and precision by preventing the user from incurring inadvertently in errors such as domain value redundancy, missing data value, incorrect data values, non-atomic data values, domain schizophrenia, duplicate occurrences, inconsistent data values and information quality contamination in the process of digitization of geospatial, location and taxa data. Moreover, these tools facilitate the digitization of data, making the process more interactive and productive.

These concepts and methods applied on geospatial, location and taxon data domain may be appropriate to and used in other data domain as well.

## References


Brundtland, G. 1987. Our common future: The world commission on environment and development. Oxford University Press, Oxford.

Cartolano, E. A., Saraiva, A. M, Veiga, A. K., Krobath, D. B., Saraiva, L. G. P., Tavares, G. 2010. Biodiversity Data Digitizer – BDD. In: Proceedings of the TDWG, Woods Hole, Massachusetts, USA.

Chapman, A. D. 2005a. Uses of Primary Species-Occurrence Data, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen.

Chapman, A. D. 2005b. Principles of Data Quality, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen.

Dalcin, E. C. 2005. Data quality concepts and techniques applied to taxonomic databases. PHd thesis, University of Southampton, United Kingdom.

Embury, S. M. 2001. Data quality issues in information systems. Pages 41 in Database Systems Cardiff University, School of Computer Science, Cardiff.

English, L. P. 1999. Improving data warehouse and business information quality: methods for reducing costs and increasing profits. John Wiley & Sons, Inc., New York.

Hill, A. W., Otegui, J., Ariño, A. H., Guralnick, R. P. 2010. GBIF Position Paper on Future Directions and Recommendations for Enhancing Fitness-for-Use Across the GBIF Network, version 1.0. GBIF, Copenhagen.

Ge, M., Helfert, M. 2007. A review of information quality research-develop a research agenda. In proceedings of the 12th International Conference on Information Quality.

Kelling, S. 2008. Significance of organism observations: Data discovery and access in biodiversity research. Report for the Global Biodiversity Information Facility, Copenhagen.

Pipino, L., Lee, Y., Wang, R. 2002. Data Quality Assessment. Communications of the ACM. pp. 211-218.

Saraiva, A. M., Veiga, A. K., Cartolano, E. A. 2011. Biodiversity Data Digitizer: A tool for a richer biodiversity data content digitization. In: The Proceedings of TDWG: Abstracts of the 2011 Annual Conference of the Taxonomic Databases Working Group. New Orleans, LA, USA.

Strong, D., Lee, Y., Wang, R. 1997. Data Quality in Context. Communications of the ACM. pp. 103-110.

Veiga, A. K., Saraiva, A. M., Cartolano, E. A. 2010. A Georeferencing Tool to Improve Biodiversity Data Quality. In: Proceedings of the TDWG, Woods Hole, Massachusetts, USA.

Veiga, A.; Saraiva, A.M.; Cartolano Junior, E.A. 2011a. Methods and tools to improve data quality in biodiversity specimen occurrence data. In: Proceedings of the European Federation for Information Technologies in Agriculture, Food and the Environment (EFITA) Conference. Prague. pp. 471-480.

Veiga, A. K., Cartolano Junior, E. A., Saraiva, A. M. 2011b. Data Quality Resources in Species Occurrence Digitization. In: Proceedings of TDWG: Abstracts of the 2011 Annual Conference of the Taxonomic Databases Working Group. New Orleans, LA, USA.
Wang, R., Redata domainy, M., Kon, H. 1995. Toward quality data: An attribute-based approach. Journal of Decision Support Systems, Volume 13, Issue 3-4. pp. 349-372.
Wang, R., Kon, H., Madnick, S. 1993. Data Quality Requirements Analysis and Modeling. Proceedings of the Ninth International Conference of Data Engineering. pp. 670-677.

# The central role of use cases in enhancing data exchange and interoperability in agriculture

Rob Lokers[1], Liisa Pesonen[2], Cor Verdouw[3], Benedicte Fusai[4], Mario Schmitz[5]
*[1]Alterra Wageningen UR, [2]MTT Agrifood Research Finland, [3]LEI Wageningen, [4]Institut de l'Élevage, [5]KTBL*
E-mail: liisa.pesonen@mtt.fi

## Introduction

Information sharing is widely addressed as a crucial issue in the European agri-food industry (AMI@Netfood 2006, Food4Life 2007, GCI/IBM 2009, amongst others). Information becomes a competitive factor that enables a more demand-driven and knowledge-based production, needed to meet the various and changing needs of consumers and society (Kinsey 2001). In describing their vision on the future value chain for 2016, the Global Commerce Initiative (GCI 2006) conclude that multi-partner information sharing among key stakeholders will be one of the main characteristics of the future supply chain.
The FP7 agriXchange project addresses the identified need for information sharing and information exchange in the agri-food sector by coordinating and supporting the setup of a network for developing a system for common data exchange in the agricultural sector through:

- establishment of a platform on data exchange in agriculture in the EU;
- development of a reference framework for interoperability of data exchange;
- identification of the main challenges for harmonizing data exchange.

This paper describes the methods used for the development of the reference framework for interoperability of data exchange and focuses on the role of use cases in that process.

## A reference framework for interoperability

Considering the fact that information sharing will be a key competitive factor in the current and future agri-food chain, it is essential for the European agricultural sector to improve on data exchange. This requires a common ground to analyze current information exchange practice and to subsequently harmonize data exchange between parties. It also demands for coordinated action between system developers. To facilitate this process, a common framework that guides this process in the right direction, is needed.

From an information-theoretical perspective, a reference framework is a combination of an architectural framework and a reference model. An architectural framework is a systematic taxonomy of concepts of how to organize the structure of information models (Sowa & Zachman 1992). It is a meta-architecture, which defines the required model types in different views and at various levels of abstraction and shows how these are related (Verdouw, Beulens, Trienekens & Verwaart 2010). A reference model is a predefined information model that captures 'recommended practices' and that is used as a 'frame of reference' (i.e. blueprint, template) to construct company-specific information models (Verdouw, Beulens, Trienekens

& Wolfert 2010). Consequently, a reference framework both defines the structure of information models (meta-model) and provides content in the form of predefined models of specific use cases. The purpose of the framework, discussed in this paper, is to support the implementation of interoperability for exchanging data in the agri-food sector. Interoperability is broadly defined as the ability of two or more systems or components to exchange information and to use the information that has been exchanged (IEEE 1990).

The approach to be followed for the set up of a common framework is based on the method proposed by Wolfert, Verdouw, Verloop & Beulens (2010) and that is depicted in Figure 1.

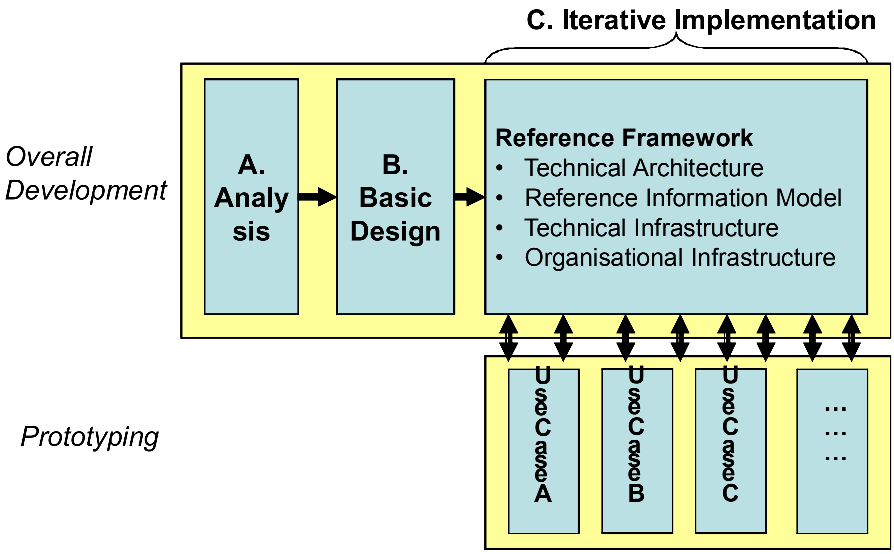



*Figure 1. Overall method for development of the reference framework (based on Wolfert et al. 2010).*

This approach is characterized by a combination of overall development and incremental prototyping and consists of the following phases:

1. Analysis: In the analysis phase, the existing state of information integration is investigated and required improvements are planned.
2. Basic Design: In the basic design, the core structure for data exchange harmonization reference framework is set up. This core structure ensures the consistency of the shared information and pilot implementations.
3. Iterative Implementation: In the iterative implementation phase, parts of the shared information are implemented in pilots. Pilot results are abstracted into the models (or solutions) of the reference framework, which serves as an incrementally growing knowledge base of 'best practices'.

All these phases feed the development of the reference framework. In the analysis phase, a conceptual model is developed defining integration and possible levels of integration (among others based on Giachetti 2004). Based on this so-called generic integration framework, more specific content is developed in the basic design and implementation phases.

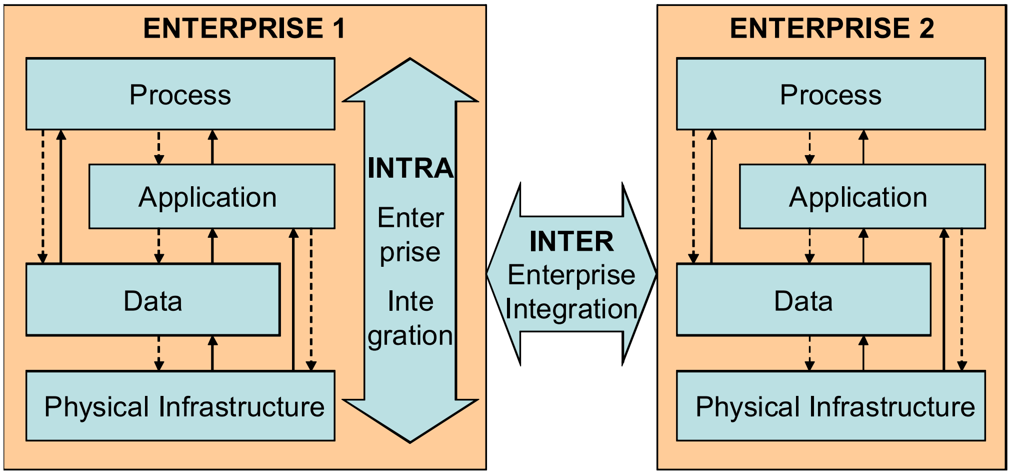



*Figure 2. Generic Integration Framework (adapted from Giachetti, 2004)*

Following the schematization of the generic integration framework as depicted in Figure 2, the reference framework distinguishes between:

Different integration levels:

*Intra-Enterprise*: within enterprises to overcome fragmentation between organizational units (functional silos) and systems;
*Inter-Enterprise*: between enterprises to move from operating as an isolated company towards a virtual enterprise that is integrated in multi-dimensional networks.

Different integration types:

*Process Integration*: alignment of tasks by coordination mechanisms;

*Application Integration*: alignment of software systems so that one online system can use data generated by another one;

*Data Integration*: alignment of data definitions in order to be able to share data;

*Physical Integration*: technical infrastructure to enable communication between hardware components.

A reference framework for interoperability of data exchange in agriculture in the EU should facilitate the harmonization of data exchange by effectively supporting the three phases of the development process. This demands a uniform and formal methodology that can be used as a sound basis to develop evolving implementations and to concurrently develop the reference framework by allowing feedback of evolving knowledge and experiences gained through the development work to again be integrated into that framework. The reference framework should therefore support the involved community in:

1. Structured documentation of the status of information integration for specific business cases relevant in the European agricultural sector. This methodology should also support the structured analysis and documentation of variants of a business case. Such variants can be of a very diverse nature. They can originate from: a) geographical differences, b) different implementations of (part of) the business case on any of the integration levels (process, application, data, physical) mentioned in Figure 2.
2. Fitting the documented business case with the concepts laid out in the basic design as the basis for designing systems that support sound information integration.
4. An easy translation of the documented business cases to a real-life implementation, following the core structure set up in the basic design.
5. Feedback of the evolving knowledge and experiences gained in developing these real-life implementations in order to be picked up in following iterations, improving the existing implementation or to be re-used in other business cases.

The role of use cases in the reference framework
One of the key elements used in the reference framework for interoperability to be developed in the agriXchange project is the integrated use of use cases as a ways to document business cases throughout the three phases of the common framework development. Use cases are utilised commonly in software development as a specification technique for capturing functional requirements (Jacobson 1992, Cockburn 2001). In the agriXchange project the work focuses on Inter-Enterprise level data exchange harmonization. Thus, instead of a development of individual, local ICT systems, the scope of the use case descriptions is in the use cases where the end-user uses dispersed system consisting of several sub-systems. The sub-systems have many independent owners and sub-users. The investigated field is very large and plenty of use case analyses are needed to carry out.

To document the defined use cases that are relevant in the context of information integration in agriculture, a use case template is developed. Actually, there are two levels of use case descriptions needed. The first one is a "wide scope" use case description covering a whole domain specific procedure, e.g. fertilizing procedure from planning to execution, consisting of a chain of processes, actors and data exchange transactions. These use case descriptions are in interest of researchers, domain specific associations, advisory services, etc. From business point of view, this level of use cases is interesting for actors in the SOA business service layer. Table 1 shows the metadata fields used in the "wide scope" use case template.

| METADATA | DESCRIPTION |
|---|---|
| Name of the use case | Short name (verb+object) plus context, country or region if relevant. |
| Short description | Short description of the use case in "natural language" to document the goal of the use case and the main actors and activities. |
| High-level graphical description | A simple and easy to understand graphical representation of the use case (e.g. in the form of a UML use case diagram) |
| Relevant countries or regions | Defines the geographical domain for this use case. |
| Relevance for European agri-sector | Percentage of EU countries where the use case is relevant at the moment and in the near future (within next 5 years) |
| Relevant parties | List of actors in the use case:<br>The parties that are actively involved in the information exchange for the use case.<br>List of stakeholders of the use case (parties that are not actively involved, but have a stake in efficient, harmonized data exchange). |
| Description of information exchange | Structured description of the information exchange processes and exchanged data (using BPMN) |
| Relevant "conditions" | a. Standards<br>b. Dictionaries<br>c. Regulations and legislation<br>d. Technologies |
| Use case variants | Description of different implementations of the business case by BPMN graphs and short written descriptions. |
| Known issues and bottlenecks for harmonization | Description of known issues that hamper efficient information exchange, using "natural language" and graphical presentation, if necessary |
| Proposed recommendations and solutions for harmonization | Description of ways to improve and harmonize information exchange, using "natural language" and graphical presentation, if necessary |

*Table 1. Metadata fields of the agriXchange "wide scope" use case template*

In the other use case description level the scope is narrower and it covers only a single interface between two actors in a "wide scope" use case. The "narrow scope" description serves as a metadata model of the interface, where the context of the data transaction of that specific interface is described shortly.

- Name of the Interface
- Short description of the purpose of the communication; actors, processes
- Description of the information and exchanged data
    - Relevant "conditions"
    - Standards used
    - Dictionaries
    - Regulations and legislation
    - Technologies
- Possible issues to consider when applying in other conditions

This use case description is needed when sharing existing interface solutions, which can be used in several “wide scope” use cases or are variants for a single interface solution in a certain “wide scope” use case. The “narrow scope” use cases are of interest of system developers, and they can be linked to the “wide scope” use case descriptions created by e.g. business modelers (bottom up). Also, parties composing novel “wide scope” use case descriptions can utilize existing interface-specific use case description in their work (top down).

The uniform documentation of business cases and interfaces in the form of use case descriptions supports the development of the common framework in a number of ways. It allows the involved user and stakeholder community to communicate about specific business cases and the information integration aspects, captured in a structured manner. The use cases assist the involved stakeholders to also identify the indirect business opportunities, which may be difficult to come across otherwise. It provides a uniform and easy to use method to analyze and describe the aspects relevant for integration purposes, supported through natural language descriptions and dedicated graphical notations that are easy to use and understand. Thus, it allows such uses cases to be used as a means of communication not only among system developers, but also with parties and stakeholders that lack the technical background to fully understand the formal languages and notations used by ICT developers (e.g. UML, Unified Modelling Language). Moreover, it supports the three identified phases of the described approach for the development of the common framework:

*Analysis* – In this phase, the current status of the business processes and especially the relevant aspects from the viewpoint of information integration are uniformly described in the format of the use case template. Besides, the template allows describing relevant use case variant. Geographically determined variants can be described either as variants within the same use case description, or as separate use cases focusing on a specific country or region. Variants that are related to differences regarding on the identified integration levels (see Figure 2: process, application, data, physical infrastructure) can be addressed by referring e.g. to relevant dictionaries (data level) or data exchange standards (application level) or by the inclusion of multiple process diagrams (process level). Existing issues and required improvements that originate from the analysis phase can also be documented.

*Basic Design* – The basic design for the common framework is based upon the use of Business Process Management (BPM) and reference information models. The technical infrastructure is based upon the principles of a Service Oriented Architecture (SOA) (Figure 3).

The use of Business Process Modelling Notation (BPMN) to model processes to be integrated, provides a sound method for documentation of the processes and the information exchange between parties acting in the use case. BPMN provides a common language for describing process behaviour, shareable by business and IT (Silver 2009). It is easy to understand by non technicians and thus provides an excellent means of communication regarding process and information integration. Moreover, it allows IT experts, in later stages, to fully specify the details of processes even to the level of executable processes running in SOA architectures. In that way it also provides the means to translate process design to the level of applications and implementation of inter-enterprise information exchange by means of services.

The BPMN models define the sequence of activities of different involved actors and the exchange of data objects, in particular between the different involved actors. In the Basic Design, attention is paid to the description of standards and dictionaries used in the data transfer interactions between the actors in investigated use cases. In real world, there are and there will be many variants of business cases competing with each other to fulfil the user requirements. This is an essential feature of the markets. This is why data exchange harmonizing work concentrates on the single data transfer interfaces between the actors, identified in the use case models. These interfaces can be freely utilised according to different business cases. The business service layer in Figure 3 represents the focus of data exchange harmonizing work. It may be sometimes difficult for ICT developers to identify relevant business cases in distributed systems, which may hinder the motivation to actively develop new type of services for end-users, and thus, needed data exchange. SOA-based system design and its business service layer (Figure 3) provide an approach to develop new intermediating services with harmonized data exchange interfaces. This provides possibilities to enhance creation of new services for the ICT developers and to improve service availability and system usability for the end-user.

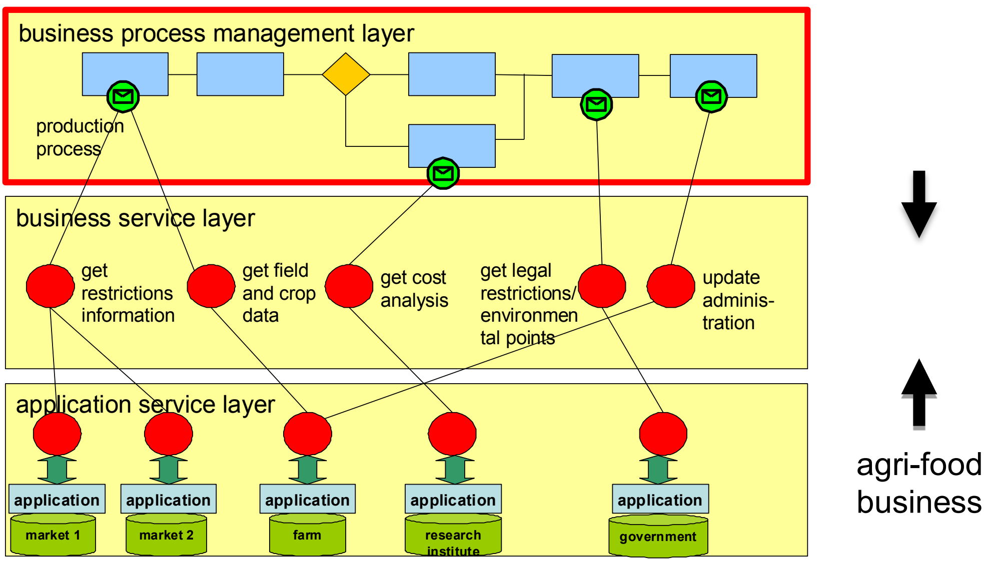



*Figure 3. Service Oriented Architecture as an essential part of the Basic Design (based on Wolfert, Verdouw, Verloop & Beulens 2010).*

*Iterative Implementation* - In the implementation phase, the modelled use cases will be used as a guideline for the detailed design and subsequent implementation of pilots. In the previous phases, the emphasis in framework development has been on the technical architecture and information modelling (in particular process and data models). Implementation demands also an organisational method, suitable for iterative, open and user-driven development of the framework. For this, the so-called 'living lab approach' has been adopted. The Living Lab (LL) is a user-centric innovation approach for sensing, prototyping, validating and refining complex solutions in multiple and evolving real life contexts (Wolfert, Verdouw, Verloop & Beulens 2010). It can be characterized by an iterative, design-oriented approach in which incremental solutions are re-used, sometimes leading to unexpected results or changes in intended innovation directions. LL promotes an alternative innovation paradigm: the end-user's role shifts from research object to a pro-active position where users, researchers and other stakeholders collaboratively document, learn and develop. To facilitate this process, the agriXchange platform will be used as the technical infrastructure that facilitates joint communication, documentation, learning and development (http://www.agrixchange.eu).

In the agriXchange project, four use cases are selected to be elaborated. Three of these use cases will be integrated as pilots supporting the iterative development of the reference framework. The fourth use case will be used as a test case in order to verify the completeness and usability of developed framework. The four selected use cases are the following:

*LPIS update* – describing the process of updating the LPIS (Land Parcel Identification System) between farmers and government.
*Geo-Fertilizer* – describing the process of variable rate application by farmers, supported by data served by government and service providers.

*Animal registration* – describing the (registration) processes involved in importing and exporting cattle.

*Animal identification* – describing the processes involved in collecting and registering animal production information throughout the production chain, supporting tracking and tracing and consumer information.

It is important to notice that the added value for the implementation phase lies in the fact that also new implementations and the experiences gathered while setting up these implementations can again be documented in the form of use cases, leading to an evolving knowledge base. Therefore, according to the LL approach, the project encourage users, experts and any other stakeholders to join the agriXchange community, to make use of the framework, to comment and discuss work in progress, to re-use developed concepts and components and to bring forward such additional use cases.

In the following section the agriXchange use case "Geo-Fertilizer" is described, focussing on just a few of the relevant fields of the described use case template. It mainly serves to illustrate how specific parts of a use case can be documented.

**An example use case: Geo-Fertilizer**

- Short description / Variant

The Geo-Fertilizer use case describes the business case in which a farmer optimizes his fertilizing practice through variable rate application, considering the local conditions on a field parcel. As such, this use case functions as a general model for on-farm precision agriculture practices where spatial information is used to optimize in-field processes and where external parties are involved to provide the required information and knowledge.

One of the possible implementations, or variants, of the business processes and the related information exchange in this business case involves the farmer, commercial service providers and national or regional government as data exchanging parties and could have the following process steps.

1. The farmer requests and receives the geometries of his LPIS parcels from the LPIS and parcel registration system provided by his national government.
2. The farmer requests a farm advisory service to provide a fertilizer advice. A request for advice, including the spatial information of the parcel, is sent to the system of the service provider.
   - The farm advisory service provider sends out a request for the soil information required for the specified parcel to a soil data service provider. The soil information is delivered to the farm advisory service provider.
   - The service provider requests information on the measured leaf area index (LAI) data for the specified field from a remote sensing data provider. He receives a LAI map covering the specified field.
   - An agronomic model processes the available datasets (parcel data and geometry, soil data, LAI map) and generates a fertilizer application advice for the farmers' field.
   - The fertilizer application advice is sent back to the farmers FMIS
3. The farmer receives the fertilizer advice and stores it in the data repository of his FMIS.
4. The farmer, through his FMIS, transforms the fertilizer advice to an application map that can be entered into the board computer of his machinery. He loads the dataset onto the machinery and fertilizes the field. During the operational process, the local application rates are logged and stored.
5. The farmer transfers the logged data from the machinery to his FMIS. The FMIS processes and stores the dataset in the repository.

- Description of information exchange

Figure 4 shows how the implementation variant of the business process in question can be visualized in the form of a BPMN graphical diagram. The BPMN concepts of pools (in Figure 4: farmer, government, N-advice provider, LAI service provider and Soil data service provider) and swimming lanes (in Figure 4: farm management system and farm machine within the pool farmer) are used to model the participating parties and relevant local components, respectively. The model clearly shows how processes are built up of interconnected tasks and how information flows from task to task (intra-enterprise) and between different parties (inter-enterprise). As such it provides an efficient means to further elaborate on use case variants and specific details on information exchange among ICT and agriculture domain experts.

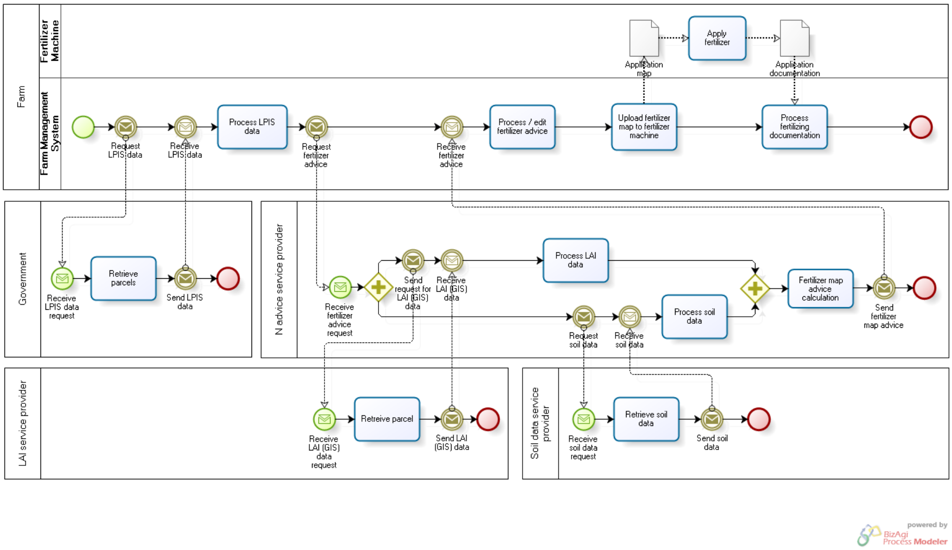



*Figure 4. Description of information exchange in BPMN of the Geo-Fertilizer use case.*

- Relevant standards

The most relevant standards that support information integration for the Geo-Fertilizer use case are:

*ISO-11783 or ISOBUS* - a world-wide recognized and widely implemented standard for on farm data communication. It supports standardized data communication and automated process execution on farm machinery. Moreover, it supports the exchange of (operational) data from FMS to farm machinery and vice versa by defining the data elements and code lists required in on-farm data communication.

*AgroXML* - the German standard for data exchange in the agricultural sector. Although it still has a low implementation grade, it seems to be the most evolved standard in the area of pre-harvest data exchange between FMIS and other business parties. Current implementation supports FMIS to government and FMIS to soil data provider.

*EDI-teelt* - a Dutch standard for data exchange in arable farming which is mostly used in post harvest agri-business and tracking and tracing applications. Although it's fairly wide spread in the Netherlands, it does not support the Geo-Fertilizer use case. A new version of the EDI-teelt standard is currently developed that will support data exchange between farm and other parties for pre-harvest business processes.

*UN/Cefact* - the standards developed by UN/Cefact support "electronic business". They cover a wider range of domains (financial, administrative, technical). In the technical working group on agriculture, a few agricultural business cases are covered or in development. Although they do not directly cover (parts of) the Geo-Fertilizer use case, it is good practice to at least consider reuse of generic core components and of relevant code lists.

- Known issues and bottlenecks for harmonization

Regarding the bottlenecks for harmonization, a few conclusions can be drawn based upon the analysed use case. First of all, focussing on the Geo-Fertilizer use case, no generally accepted set of standards is available at the European level that supports this use case. Commonly used standards in agriculture in general (UN/Cefact, AgroXML, EDITeelt, (E)Daplos) have a strong focus on aspects like trade, logistics and tracking and tracing and not so much on on-farm data exchange and process integration. The ISO-11783 standard is widely accepted and supported for standardized data communication and automated process execution on farm machinery and for the exchange of data from FMIS to farm machinery. However, implementation of these standards by machine manufacturers and FMIS providers is still under development in many cases. In general, the support of spatial data standards (e.g. the standards developed by the Open Geospatial Consortium) as part of standardized data exchange is still immature.

## Conclusions

Information sharing will be a key competitive factor in the future value chain of agri-food sector. In order to facilitate information sharing and data exchange in agriculture, The FP7 agriXchange project develops a reference framework to guide the improvement of information sharing and data exchange in agriculture. The approach is characterized by a combination of overall system development and incremental prototyping. Use cases have a central role in this approach throughout three different phases of development of the generic framework; 1) analysis of present situation of data exchange in different countries and system implementations, 2) constructing the Basic Design for the generic reference framework and 3) implementation of results in LivingLab actions to get feedback for further harmonization work. Use cases capture functional features of dispersed systems composing the value chain, and assist to define and analyse specific business cases and data exchange interfaces. They are uniformly structured and comparable and they provide a good means of communication among the community of the network in workshops, living labs, WEB2.0, etc. As such, they provide a way to discuss and analyze business cases from an integrated business and technical perspective. The use of business process modelling pinpoints to the domain typical data exchange interfaces for further harmonizing work. Use cases provide a means of further research, but via the BPMN modelling also the first step to implementation of information sharing and data exchange systems based upon a Service Oriented Architecture.

## Acknowledgements

The work leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 244957.

## References


AMI@netfood 2006. European ICT Strategic Research Agenda for Agri-food & Rural development; a vision for 2015.

Cockburn, A 2001. *Writing Effective Use Cases*, Addison-Wesley.

Food4Life 2007. European Technology Platform on Food for Life - Strategic Research Agenda 2007-2020, CIAA.

GCI 2006. *2016: The Future Value Chain*, Global Commerce Initiative / Capgemini / Intel.

GCI/IBM 2009. Global Commerce Initiative Information Sharing Report.

Giachetti, RE 2004. *A framework to review the information integration of the enterprise.* International Journal of Production Research 42(6), 1147-1166.

IEEE 1990. IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. New York: Institute of Electrical and Electronics Engineers.

Jacobson, I, Jonsson, P, Christerson, M and Overgaard, G 1992. *Object-Oriented Software Engineering – A Use Case Driven Approach*, Addison Wesley Longman, Upper Saddle River, N.J.

Silver, B 2009. *BPMN Method & Style*, Cody-Cassidy Press.

Sowa, JF & Zachman, JA 1992. Extending and Formalizing the Framework for Information-Systems Architecture. IBM Systems Journal 31(3): 590-616.

Verdouw, CN, Beulens, AJM, Trienekens, JH & Verwaart, T 2010. *Towards dynamic reference information models: Readiness for ICT mass customization*. Computers in Industry 61(9): 833-844, http://dx.doi.org/10.1016/j.compind.2010.07.008

Verdouw, CN, Beulens, AJM, Trienekens, JH & Wolfert, J 2010. *Process modelling in demand-driven supply chains: A reference model for the fruit industry*. Computers and Electronics in Agriculture 73(2): 174-187, http://dx.doi.org/10.1016/j.compag.2010.05.005.

Wolfert, J, Verdouw, CN, Verloop, CM. & Beulens, AJM 2010. *Organizing information integration in agri-food - a method based on a service-oriented architecture and living lab approach*. Computers and electronics in agriculture 70:389-405, http://doi:10.1016/j.compag.2009.07.015.

# Competence centre SenGIS – exploring methods for georeferenced multi-sensor data acquisition, storage, handling and analysis

[1]M. Keller, C. Zech, [2]M. Jackenkroll, [1]M. Weis, [2]J. Link-Dolezal, [1] R. Gerhards, [2]W. Claupein
[1]*University of Hohenheim, Department of Weed Science,* [2]*University of Hohenheim, Department of Crop Science*
*E-mail: Martina.Keller@uni-hohenheim.de*

## Introduction

At the beginning, precision farming was defined as an agricultural production system technique, which takes the natural heterogeneity of agricultural sites into consideration (Ehlert,1995). Its first main field of application was site-specific fertilisation. Today, precision agriculture is more widely defined and includes techniques such as automatic guidance systems, product traceability and on-farm research (Gebbers and Adamchuk, 2010). This chapter deals with precision farming in the sense of site-specific production and management practices.

Site-specific farming considers the actual local growth conditions and adapts the management strategies and treatments accordingly to obtain an optimum system output. By doing so, negative impacts of agricultural practices on the environment may be reduced (Pierce and Nowak, 1999) and economically valuable savings can be achieved (Timmermann *et al.*, 2003). The key to precision farming for sub-field level management is the collection of data with a high spatial resolution.

Information about existing field heterogeneity can be obtained either by hand or by automated sensor measurements. The former is mainly restricted to research due to its labour intensity and high costs. During recent decades, several sensor principles have been identified and tested for precision farming purposes, to be used in research and practice. Sensors can reveal spatial variability in terms of plant nutrition status, leaf area, green biomass, soil texture and many other parameters (Oerke *et al.*, 2010, part II). Nevertheless, only few sensors have succeeded so far as management tools in agricultural production systems, e.g. the Yara N-sensor (Lammel *et al.*, 2002). The main reason is the complexity of the data handling and in particular the interpretation and transformation of the sensor readings into decisions and application maps (Dobermann *et al.*, 2004). A requirement for the latter is, that the sensor data show high correlation with the relevant parameters in the decision process. Furthermore, the correlation needs to be robust under varying field conditions and the sensor data should be sufficiently specific for the parameters to be assessed.

Basic research is necessary to evaluate different sensors under field conditions and their capability to track the variability of major crop growth condition parameters. This is essential to develop sound sensor based decision systems for site-specific management. With increasing amount of georeferenced sensor data, its long-term storage and availability to different users become issues to be addressed. Metadata sets, describing the properties of the measurement data, are important to have valuable and reusable data for this kind of research.

To face these challenges the "Competence Centre for Sensors and Geoinformation Systems" (SenGIS) was established in 2009 at the University of Hohenheim, Germany. The focus of the Competence Centre is on *a)* the development of sensor platforms to do basic research on the spatial acquisition of data from multiple sensors and *b)* the establishment of a geodatabase to describe and store georeferenced data for long-term use and maintain their availability for different users. The objective of this chapter is to present the concept and structure of SenGIS. Therefore details of the sensor platforms and the geodata handling are given and three examples of different applications with interdisciplinary collaboration are briefly presented.

**Sensing platforms**

A central aspect of SenGIS is the integration of various sensor measurement techniques and analysis methods, which together form the basis of precision farming, comprising biological issues from climate, soil, fertilisation up to pest control and technical issues as well as applied mathematics. Building a multi-sensor platform means to enforce synergistic effects. The competence centre aims to bridge the gap between researchers of different disciplines and institutes and involves industry into the development process of new precision farming technologies. To establish new measurement techniques in practical approaches and products, these have to be tested under field conditions. Consequently, calibration as well as the development of methods for data analysis, interpretation and evaluation are part of the project.

A ground-based sensor platform and two low-air-based sensing platforms are currently developed and in use at SenGIS. Each sensor system needs to be calibrated and operated differently and provides its data via custom interfaces. Most of the sensors can be connected to a field computer to trigger measurements and receive the measured values for storage. In a first approach sensor data can be logged together with a GPS-position and then be analysed 'offline' in the office. The second step is the implementation of 'online' data analysis algorithms to control treatments in the field. Analysis techniques and decision components based on the sensor measurements are developed.

Ground-based sensor platform

The sensor platform is suitable to measure simultaneously with different sensors. Available sensors to measure properties of plant parameters are mounted on a vehicle, located on the left and right side in rows (Figure 1; left). Sampling locations are defined along the way and each sensor is triggered to measure at these locations. Equipped with a GPS-RTK receiver, all measured data can be precisely georeferenced. At present, various optical sensors are installed (bi- and multi-spectral cameras, spectrometers and fluorescence sensors, Table 1). Such sensors can either be passive or active (with illumination) and conduct non-destructive measurements. Thus, repeated measurements for crop monitoring are possible. In the near future, additional sensor types will be used to acquire information about local environmental conditions as temperature or soil moisture. The integration of new sensors for comparison and test purposes is possible due to a modular system design.

| Sensor | Sensor principle | Examples of determined parameters |
|---|---|---|
| Multiplex (Force-A, Orsay, France) | Fluorescence (B/R/IR) | Photosynthesis activity, flavonol content, nitrogen state |
| MiniVegN (Fritzmeier, Großhelfendorf, Germany) | Fluorescence (R/IR) | Photosynthesis activity |
| FieldSpec HandHeld (Analytical Spectral Devices Inc., Boulder, USA) | Spectrometer (325–1075 nm) | Reflectance spectrum |
| Yara N-sensor (Yara GmbH & Co. KG, Dülmen, Germany) | Spectrometer (400–1000 nm) | Reflectance spectrum, nitrogen state |
| MMS1 enhanced (Tec5 Corp., Oberursel, Germany) | Spectrometer (310–1100 nm) | Reflectance spectrum |
| Bi-spectral camera (Asentics, Siegen, Germany; Thecon, Liebenburg, Germany) | Near-infrared and red spectrum | Coverage of plants, weed identification |
| Optio A10 (Pentax Corp., Tokio, Japan) | RGB-camera | Visible light image |
| Mavolux 5032 (Gossen GmbH, Nürnberg, Germany) | Luxmeter | Ambient illumination (380–700 nm), calibrated to eye sensitivity |
| PAR meter (Spectrum Technologies, Inc., Illinois, USA) | Light intensity | Photosynthetically active radiation (400–700 nm) |

*Table 1. Overview of the available sensors, their distributors, sensor principle and actual applications. In brackets the measured wavelengths and spectra are indicated: Blue (B), infra-red (IR), green (G) and red (R).*

Low-air-based sensor platforms

Unmanned aerial vehicles (UAVs) provide a flexible platform for small-scale remote sensing surveys and are becoming popular due to their relatively low costs and flexibility (Grenzdörffer *et al.*, 2008). A further advantage is, that flights can be undertaken, even if the field is not accessible to a ground-based sensor platform due to weather conditions. Furthermore, soil compaction is not an issue. In contrast to high-flying remote sensing platforms like planes or satellites, UAVs are navigated on a lower operating altitude (up to 150 m) and thus can be used for data collection also under cloudy conditions. For takeoffs and landings outside of a declared airfield the available payload of the platform is limited to a maximum weight of 5 kg (German regulations). This given threshold limits the application to light-weight sensors, combined with efficient power supply. Besides, an experienced pilot is needed to conduct the operation in the field and still, flights are highly dependent on wind conditions.

Currently, two optical sensor systems can be mounted on the available UAVs: the spectrometric device MMS 1 enhanced and a digital RGB camera (Tab. 1). One of the UAVs is equipped with an autopilot system (MicroPilot Inc., Stony Mountain, Canada) and can fly along prepared way-points to cover a certain area completely.


*Figure 1. left: Sensor platform "Sensicle" for simultaneous field measurements with different sensors. The sensors are mounted at the left and/or right of the vehicle to be able to measure the same locations along the track; right: UAV with spectral sensor.*

**Geodatabase**

Most operations in precision farming are based on spatial information. This is the case for many tasks ranging from field measurements, planning and decision-making to site-specific applications with appropriate implements, e.g. variable rate technology. The sound organisation of the data flow is therefore of high importance for precision agriculture. Figure 2 gives an overview of the data flow within SenGIS. In SenGIS different types of data are distinguished: measurement data (Sensing), the analysed data (Analysis), the background and control data for decision support (Decision support) and the application data (Application). All data should be described by metadata. The storage, query and analysis of data can be performed by a number of different scientists, i.e. the project partners or external partners or users in general. Acquired sensor measurements are stored in a database and described by metadata within the geodatabase. External data sets can either be included or referenced via their metadata descriptions. Analyses based on the data can then be conducted and their results be added to the database or used in decision components, e.g. for the creation of application maps and application control in the field.

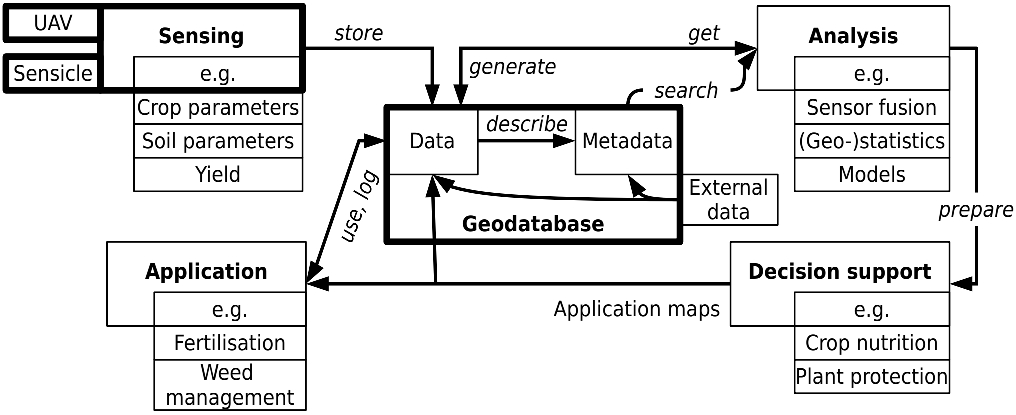



*Figure 2. The data flow in the SenGIS structure is organised around the central geodatabase. Special focus is on the sensing inputs and their storage. The data sets are used by analysis and decision steps to derive an optimised application procedure. The same data flow is used for field trials.*

**Organisation of data and services**

Measurement data and additional geographic information like remote sensing imagery, field boundaries and trial maps are the input to the analyses steps (Figure 2). There are different data sources to be considered in terms of data types and their providers. Providers of data and users are often dispersed geographically and use different information technology systems. Data acquired within the project can be made available directly by SenGIS, data of partners and external sources can also be integrated into the system. Since there is a heterogeneous structure of users and applications of the system, the data handling has to be flexible and based on common standards to achieve optimum interoperability.

There are still gaps in the seamless handling of such data between different systems and for multiple purposes. These gaps are identified by recent research efforts of projects on the European level such as FUTUREFARM (Blackmore *et al.*, 2009) and AGRIXCHANGE (Nikkilä *et al.*, 2010). Findings and standardisation efforts of these projects are the guidelines for the implementation of the infrastructure components of SenGIS.

This improves the usability of the data for precision farming applications and makes them available for the users as well as other information systems, e.g. farm management systems and application control units. A server based system is able to handle data access by the use of standardised services. Working with easily operable, normative web-browsers is possible as with professional software (Desktop GIS) for more complex analyses. The advantages of online map-viewers by browsers are manifold. They do not need special software-licenses, they are fast, low-maintenance and can be used from nearly every computer with access to the network. Spatial data infrastructure (SDI) components of external data providers can provide additional data like aerial and satellite images, soil maps and cadastre data. SDI components are set up to tackle the handling of geodata and metadata. These components have interfaces to communicate and transfer data between each other.

One of the goals is the standardisation of the measurement data sets in a way that the data can be used seamlessly in automated analyses. A complete documentation of all measurements, trial maps and auxiliary data sets help researchers to develop precision farming applications and to address new research topics. Besides the raw measurements and trial data a documentation of the management strategies is maintained over the years, enabling long-term research in precision farming. Due to the sound organisation of the data and description by metadata, an unstructured data overflow as described by Zhang *et al.* (2002) can be avoided.

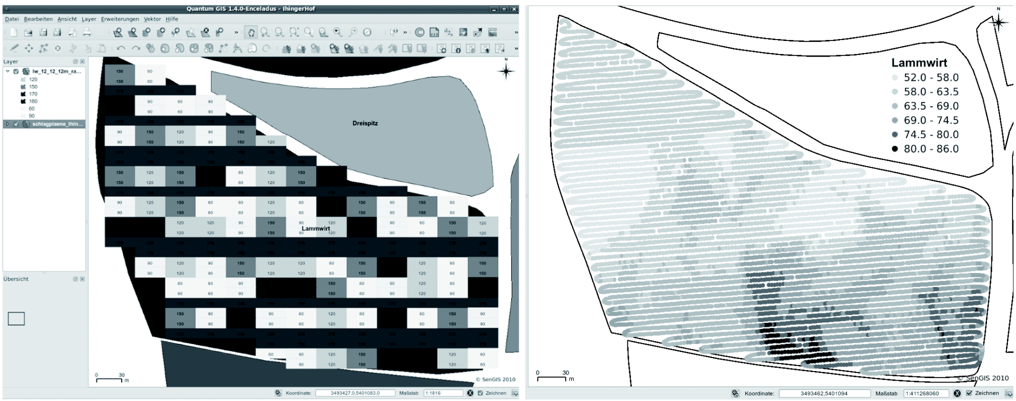



*Figure 3. Data from the geodatabase visualised in a desktop GIS. Field boundary data and trial map (N levels) on the left and raw measurements (e.g. EM38 soil conductivity measurements) on the right.*

The central place to store all precision farming data is a geodatabase (Figure 2). A geodatabase combines the advantages of relational database management system (RDBMS), which offers fast and location-independent access to 1-dimensional (1D) data, with spatial extensions for the efficient handling of multidimensional data. The spatial extensions provide data types and functions for the processing and indexing of 2D and 3D vector geometries. Raster data (images) are not stored within the database due to the amount of data, these are usually handled within a file system. Caching methods are used to increase performance of raster data. PostgreSQL was chosen as RDBMS and extended by the PostGIS-software, which offers spatial functionality. In the RDBMS, databases for different project partners can be created and used in parallel to increase performance, security and structure. A regular backup strategy minimises the likelihood of data loss. The central storage place therefore has an advantage for the users in terms of data security.

All data sets are delivered through standardised geodata web services, as defined and standardised by the Open Geospatial Consortium. The most relevant implemented services are:

**WFS(-T)**

– web feature service (transactional), providing access to the vector data with geometries and attributes, either read-only or with write permissions (transactional)

**WMS**
– web map service, delivering rendered maps of the data, mainly for visualisation purposes

**WCS**
– web coverage service, access to raster data sets (imagery)

**CSW**
– catalogue service-web, providing access and search facilities for metadata, describing the available data sets

**WPS**
– (geospatial) web processing service, implementing algorithms and complex models to manipulate geodata via standardised request and response structures (input/output formats)

An example of the data usage from the geodatabase via WFS-T interface in a GIS is given in Figure 3.

Initially, a web based system-environment is open for any user. There are no standards for authentication to OGC-services, although considerations and test environments exist (OGC, 2011). For a user/group-based management individual wrappers around the services are needed to implement authentication and authorisation (A&A). The user's permissions to view and especially to write and change data have to be restricted by authorisation. For this purpose the OWS-proxy-module (Rothstein, 2011) was added, acting as intermediate server component. This way different views of the user's own area and more generalized information about foreign parcels can be accessed. A gracefully built right management on all levels protects the interests of farmers, for example sensitive information regarding status and quality of their land.

Additionally to A&A the data transfer itself should be secured in insecure networks like the internet. If required in future, encryption via secure sockets layer (SSL) has to be implemented by additional proxy servers, as proposed and implemented by Intevation (2011). The proxy software provides access to a web service by Hypertext Transfer Protocol Secure (HTTPS) even for clients not supporting transport security related features. These components add a certain level of complexity to the SDI and therefore possibly lead to difficult-to-debug problems for the users, but are nonetheless regarded as necessary for a larger interoperability of the services.

### Metadata

Users and applications of the infrastructure can find available data sets based on metadata descriptions. The type of data, their attributes, the geographical extent, licensing information, details about the quality and responsible persons are parts of the metadata description. Based on standardised information, not only humans, but also automated analysis processes can discover and use the data. There are several metadata standards available for geodata, none of them particularly addressing precision farming needs. Furthermore, external metadata-databases can be added to the SDI. The scanning of this data and implementation in the project by items or keywords is possible through CSW services. These can be cascaded, such that the information layers available in external systems can be queried alongside the local information.

Korduan (2004) developed extensions to the Content Standard for Digital Geospatial Metadata (CSDGM) in the precision farming domain. On the European level the INSPIRE initiative defines profiles for the interoperability of geodata, based on "Geographic Information - Metadata/Services", the ISO 19115 and ISO 19139 standards. In this project the latter standards are used to describe the geodata sets and provide them via a catalogue service component (CSW) as implemented by Geonetwork (OSGeo, 2011). Access information to the data sets are given, linking to the services providing the data. Templates for the description of sensors and data attributes ease the creation of consistent metadata. However, still many fields of the metadata files are free text entries for textual description. This limits the automatic interpretation of the metadata information. Common, general terms can be taken from thesauri like the GEMET (EEA, 2004), but detailed and technical descriptions are not possible based on them. Ongoing research developments towards a standardised description of sensors and models result in Sensor Web Enablement (SWE), defining a Sensor Model Language (SensorML) and service structures for queries (Botts *et al.*, 2007). Thessler *et al.* (2011) review the user requirements for sensor networks and their application in the agricultural domain for optimised crop production. In the long term, such descriptions are desirable for the sensors, but their integration into the metadata scheme is still to be accomplished. ISO 19115 contains a limited number of describing attributes for sensors. By cascading this information a coherence between CSW and sensor description could be made.

Case studies

The following subsections illustrate briefly research activities which benefit from the infrastructure of SenGIS. The first and second example are part of the data analysis, the development of decision support systems and the application component within the dataflow scheme (Figure 2). The third case study presents the adoption of the geodatabase in scientific work.

### Measurement of different nitrogen levels

Nitrogen is an essential plant nutrient and of primary importance for the production of high yields at a good quality level. To optimise fertiliser use and derive management zones for a site-specific application, the current plant nutrition status has to be assessed. Spectrometry and fluorescence measurements are efficient and suitable methods to monitor the nitrogen status of crops (Schächtl *et al.*, 2005; Reyniers *et al.*, 2004). For comparative purposes measurements with the spectrometer FieldSpec HandHeld (HH) and the Multiplex fluorescence sensor (Table. 1) were taken in a continuous mode in winter wheat (Triticum aestivum L.), both sensors were mounted on the ground-based platform. Both sensors measure optical properties of the crop canopy in a non-invasive way. Measurements are taken in field trials with treatments of different nitrogen levels (Figure 3, left). The collected data of both sensors is additionally compared with data of a Yara N-sensor, which is a spectrometer commercially in use for site-specific nitrogen application. Sensor data and inferred indices for crop nitrogen state are compared with nitrogen analysis of biomass taken during the experiments and yield data. One goal of further investigations is sensor fusion of ground-based as well as aerial data samples. Since winter wheat is a well investigated crop regarding nitrogen status assessed by sensors, sensor fusion approaches can be developed for the calculation and calibration of management decision systems. The obtained data in these trials can be combined and presented by the WEB-GIS-viewer. Data from previous trials or data about the specific field can be found by metadata and added to the view in the browser.

The use of sensors to estimate weed–crop interaction

The automatic detection of weeds is a prerequisite to put site-specific weed control into practice. In addition, the automatic detection of weeds eases weed sampling in field trials. The use of bi-spectral camera sensors and suitable classification algorithms allows automatic weed detection (Weis *et al.*, 2008) and is part of the SenGIS infrastructure. Every single image taken is georeferenced. Thus, the weed distribution can be determined with high spatial resolution: The sensors are mounted on the ground-based sensor platform and are currently used in on-farm research trials in winter wheat and maize (Zea mays L.). In these trials the effect of weeds on yield is determined. In contrast to the classical block design, the trial area of on-farm-research trials is generally much larger and the natural heterogeneity of soil and weeds are explicitly part of the design (Ritter *et al.*, 2008). The weed distribution data is combined with georeferenced yield data and with georeferenced soil data. The effect of weeds, soil properties and herbicide application on the yield can then be separately determined with linear mixed models with spatial correlation structure (Gutjahr *et al.*, 2008; Ritter *et al.*, 2008). The spatially increased sampling rate of the automatic weed detection is expected to provide better estimates of the effect of weeds on yield compared to manual weed counts in these field experiments. The determined weed distribution and derived application data are stored in the geodatabase. Application maps can be generated in a viewer in the browser or by GIS software accessing the datasets directly. Archived images, the original measurements, are maintained and linked to these data sets. The spatial distribution of weeds can be tracked over years and is available for long-term research. The data is available to authorised users by services accessing the geodatabase. A user-interface for a metadata-search offers the possibility to find suitable data sets. For example fields with a certain crop or weed infestation could be easily found. Researchers can create tools to analyse the data and provide these analyses tools via SDI components, for example WPS. These increase the transparency as the data and the analysis steps are both accessible.

Using the geodatabase for scientific work

An often encountered constraint in research is the fact, that scientists store their data sets locally on their computers. For other users it is often impossible to understand and interpret the data organization, storage system and the data itself. This problem becomes worse with the large amount of data which can be generated by sensors. A geodatabase offers the possibility of a central, well-structured storage of data, accessible via network. Combined with information about field management, researchers will have an invaluable database, which will include data over a longer period of time and from different locations. For the scientific work data can be exported easily by searching via the metadata-tools and visualising them in the browser interface. Verbal descriptions or map views can show further details. Services like WMS and WFS are used for publishing. By WFS-T users can edit data and add more information. The advantage of using standards along the whole analysis process makes data from different sources comparable and ready for automatic handling. Analysis models published and offered by WPS are flexible like map and feature-services and can be easily shared with other parties.

## Discussion and outlook

The two main components of SenGIS are the set-up and implementation of different sensor platforms and a geodatabase. The former allows comparison, combination and testing of different sensors. The gained knowledge will be used to develop and improve decision systems for site-specific management. For example, current measurements can be used to update state variables of site-specific crop growth models and thus could improve site-specific nitrogen applications (Link *et al.*, 2008). Furthermore the suitability of sensors in new fields of application can be easily studied in an interdisciplinary environment. The availability of a ground-based sensor platform and low-air sensor platforms offers several advantages: The scale of the measurements ranges from the plant scale quantified with the ground-based sensor platform to the canopy scale quantified with the low-air-based platforms. The comparison of these measurements leads to further knowledge about the scale dependencies of such measurements. An advantage of a low-air-based platforms is, that larger field areas can be measured in less time compared with a ground-based sensor platform, though at the expense of spatial resolution. In contrast to low-air-based measurements, the spatial resolution of data sampling can be more easily varied with a ground-based sensor platform. A precise triggering of sensor measurements of the ground-based sensor platform will furthermore allow comparison of measurements taken by different sensors at the exactly same location and with many data points due to the automation. This minimises the influence of locational variance for each data point and increases the comparability of the collected data. With such data sets, sensor fusion approaches are possible.

The geodatabase, the second pillar of SenGIS, serves as a prototype for further developments. The architecture of the geodatabase is modular and offers the possibility to integrate external data, other information and analysis systems based on standards for the data exchange. Within the case studies analysis steps are designed to operate on the data provided by the infrastructure. They are adapted to the current data structure of the measurements. Functional analysis algorithms can later be automated and integrated into the service oriented structure as a component.

There are still several questions to be addressed in the project: The accuracy of the used sensors has to be determined. Most of the sensing principles are not specific to a unique (e.g. biological) parameter of interest, which was defined as a requirement in the introduction for a sensor to be used for site-specific management. Thus, the combination of different measurement principles could provide more specificity to assess a parameter of interest. Furthermore, the performance of the sensor platforms and the detailed architecture of the geodatabase are still to be optimised and adapted to the needs of the users and their applications.

**Acknowledgements**

The authors thank the Carl-Zeiss-Foundation for funding this work.

**References**

Blackmore, S., Apostolidi, K. and Pahlitzanakis, G. (2009). Second periodic report for FutureFarm (2009), Deliverable D8.5, FUTUREFARM – Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards. URL: http://www.futurefarm.eu/node/220

Botts, M., Percivall, G., Reed, C. and Davidson, J. (2007). OGC sensor web enablement: Overview and high level architecture, OpenGIS White Paper OGC 07-165, Open Geospatial Consortium Inc., Wayland, MA, USA. URL: http://portal.opengeospatial.org/files/?artifact_id=25562

Dobermann, A., Blackmore, S., Cook, S. E. and Adamchuk, V. I. (2004). Precision farming: Challenges and future directions, in T. Fischer, N. Turner, J. Angus, L. McIntyre, M. Robertson, A. Borrell and D. Lloyd (eds), New directions for a diverse planet: Proceedings of the 4th Internation Crop Science Conference, Cropscience, The Regional Institute Ltd, Brisbane, Australia, p. 19. URL: http://www.cropscience.org.au/icsc2004/symposia/4/3/217_dobermanna.htm

EEA (2004). GEMET - GEneral Multilingual Environmental Thesaurus. hosted by EEA – European Environment Agency, Eionet – European Environment Information and Observation Network. URL: http://www.eionet.europa.eu/gemet

Ehlert, D. (1995). Technik für die kleinräumige Bestandesführung (Technic for small-scale crop management), Vol. 214 of KTBL-Hefte, Landwirtschaftsverlag, Münster-Hiltrup, ATB/KTBL Arbeitspapier Stand der Technik und Forschungaufgaben zur kleinräumigen Bestandesführung, pp. 50–57. ATB/KTBL-Kolloquium am 13./14. April 1994 in Potsdam-Bornim.

Gebbers, R. and Adamchuk, V. I. (2010). Precision agriculture and food security, Science 327(5967): 828–831. URL: http://www.sciencemag.org/content/327/5967/828.abstract

Grenzdörffer, G., Engel, A. and Teichert, B. (2008). The photogrammetric potential of low-cost UAV's in forestry and agriculture, Int. Arch. of Photogrammetry and Remote Sensing - XXIst ISPRS Congress, Vol. Vol. XXXVII, ISPRS - International Society for Photogrammetry and Remote Sensing, Beijing, China, pp. 1207–1214. URL: http://www.isprs.org/proceedings/XXXVII/congress/1_pdf/206.pdf

Gutjahr, C., Weis, M., Sökefeld, M., Ritter, C., Möhring, J., Büchse, A., Piepho, H. and Gerhards, R. (2008). Erarbeitung von Entscheidungsalgorithmen für die teilflächenspezifische Unkrautbekämpfung (Development of decision rules for site-specific weed management), in R. Gerhards and R. Belz (eds), 24th German Conference on Weed Biology and Weed Control, March 4-6, Vol. Special Issue XXI of Journal of Plant Diseases and Protection, German Phytomedical Society (DPG), Eugen Ulmer KG, Stuttgart-Hohenheim, Germany, pp. 143–148. 24th German Conference on Weed Biology and Weed Control, March 4-6. URL: http://www.jpdp-online.com/special-issueXXI

Intevation (2011). InteProxy - Security extension for unsecure OWS clients to secure spatial data infrastructures, online. Intevation GmbH. URL: http://inteproxy.wald.intevation.org/

Korduan, P. (2004). Metainformationssysteme für Precision Agriculture (Meta information systems for precision agriculture), PhD thesis, Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät, Institut für Management ländlicher Räume. URL: http://www2.auf.uni-rostock.de/gg/korduan/mysite/docs/Dissertation_Korduan.pdf

Lammel, J., Wollring, J. and Reusch, S. (2002). Tractor based remote sensing for variable nitrogen fertilizer application, in W. J. Horst, M. K. Schenk, A. Bürkert, N. Claassen, H. Flessa, W. B.

Frommer, H. Goldbach, H.-W. Olfs, V. Römheld, B. Sattelmacher, U. Schmidhalter, S. Schubert, N. Wirén and L. Wittenmayer (eds), Plant Nutrition, Vol. 92 of Developments in Plant and Soil Sciences, Springer Netherlands, pp. 694–695.

Link, J., Batchelor, W. D., Graeff, S. and Claupein, W. (2008). Evaluation of current and model-based site-specific nitrogen applications on wheat (Triticum aestivum L.) yield and environmental quality, Precision Agriculture 9(5): 251–267.
URL: http://www.springerlink.com/content/l03l5042u278523p/

Nikkilä, R., Seilonen, I. and Koskinen, K. (2010). Software architecture for farm management information systems in precision agriculture, Computers and Electronics in Agriculture 70(2): 328–336.
URL:http://www.sciencedirect.com/science/article/B6T5M-4XMC000-1/2/54b46b3c8aa2a9ee60d022eb392b017a

Oerke, E.-C., Gerhards, R., Menz, G. and Sikora, R. A. (eds) (2010). Precision Crop Protection - the Challenge and Use of Heterogeneity, 1 edn, Springer Verlag, Dordrecht, Heidelberg, London, New York.
URL: http://springer.com/978-90-481-9276-2

OGC (2011). OGC authentication for OGC web services, Engineering Report OGC 10-192, Open Geospatial Consortium.
URL: http://www.opengeospatial.org/projects/initiatives/authie

OSGeo (2011). Geonetwork opensource, online. Open Source Geospatial Foundation.
URL: http://geonetwork-opensource.org/

Pierce, F. J. and Nowak, P. (1999). Aspects of Precision Agriculture, Vol. 67 of Advances in Agronomy, Academic Press, pp. 1–85.
URL: http://books.google.com/books?id=tyqHWDxPIDwC

Reyniers, M., Vrindts, E. and De Baerdernaeker, J. (2004). Fine-scaled optical detection of nitrogen stress in grain crops, Optical Engineering 43(47): 3119–3129.

Ritter, C., Dicke, D., Weis, M., Oebel, H., Piepho, H., Büchse, A. and Gerhards, R. (2008). An on-farm approach to quantify yield variation and to derive decision rules for site-specific weed management, Precision Agriculture 9(3): 133–146.

Rothstein, U. (2011). Ows proxy, online (Wiki page).
URL: http://www.mapbender.org/OWS_Proxy

Schächtl, J., Huber, G., Maidl, F., Sticksel, E., Schulz, J. and Haschberger, P. (2005). Laser-Induced chlorophyll fluorescence measurements for detecting the nitrogen status of wheat (Triticum aestivum l.) canopies, Precision Agriculture 6(2): 143–156.
URL: http://dx.doi.org/10.1007/s11119-004-1031-y

Thessler, S., Kooistra, L., Teye, F., Huitu, H. and Bregt, A. K. (2011). Geosensors to support crop production: Current applications and user requirements, Sensors 11(7): 6656–6684.
URL: http://www.mdpi.com/1424-8220/11/7/6656/

Timmermann, C., Gerhards, R. and Kühbauch, W. (2003). The economic impact of the site-specific weed control, Precision Agriculture 4(3): 249–260.
URL: http://dx.doi.org/10.1023/A:1024988022674
Weis, M., Gutjahr, C., Rueda Ayala, V., Gerhards, R., Ritter, C. and Schölderle, F. (2008). Precision farming for weed management: techniques, Gesunde Pflanzen 60: 171–181.
URL: http://www.springerlink.com/content/566u8x1646331v08/

Zhang, N., Wang, M. and Wang, N. (2002). Precision agriculture–a worldwide overview, Computers and Electronics in Agriculture 36: 113–132.
URL:http://www.sciencedirect.com/science/article/B6T5M-46YXKP5-2/2/da7c7d3218a24a0b263af9994a5cf825

# Condition monitoring in centrifugal irrigation pumps with selforganizing feature visualisation

Dimitrios Moshou, Ioannis Gravalos, Dimitrios Kateris, Athanasios Natsis, Nader Sawalhi, Spyridon Loutridis
*Aristotle University of Thessaloniki, Technological Educational Institute of Larissa, Aristotle University of Thessaloniki, Agricultural University of Athens, The University of New South Wales, Technological Educational Institute of Larissa*
E-mail: dmoshou@agro.auth.gr

## Introduction

Irrigation pumping plants are usually of the centrifugal type. A centrifugal pump plays an important role in industries and it requires continuous monitoring to increase the availability of the pump. The pumps are the key elements amongst others, in the food industry, waste water treatment plants, agriculture, oil and gas industry, paper and pulp industry. In a monoblock centrifugal pump, bearing, seal and impeller are the critical components that directly affect the desired pump characteristics. In a monoblock centrifugal pump, defective bearing, defective seal, defect on the impeller and cavitation cause more number of serious problems such as abnormal noise, leakage, high vibration, etc. (Sakthivel, Suguraman and Babudevasenapati, 2010). Vibration signals are widely used in condition monitoring of centrifugal pumps (Peck and Burrows, 1994). Fault detection is achieved by comparing the signals of monoblock centrifugal pump running under normal and faulty conditions.

The bearing is at the heart of rotating machinery playing a very important role in industrial applications and is mainly used to support and fix the axle in rotating machinery. Bearing failure in practical operation can lead to the breakdown of the whole machine. To increase reliability and reduce loss of production condition monitoring of bearing became more and more important in recent years. The use of vibration signals is quite common in the field of condition monitoring and fault diagnosis of bearings (Xu, *et al.*, 2009).

A recent study shows that more than 40% of induction motor failures are related to bearings (Singh and Kazzaz, 2003). Therefore, this type of fault must be detected as soon as possible to avoid fatal breakdowns of machines that may lead to loss of production (Han, Yang, Choi and Kim, 2006; Zarei and Poshtan, 2007). Researchers have suggested many methods to identify the bearing faults at an early stage (Mathew and Alfredson, 1984; Khan and Williams, 1992; Prabhu, 1996; Loparo, *et al.*, 2000). A machine vibration signal is composed of three parts, stationary vibration, random vibration, and noise.

To inspect raw vibration signals, a wide variety of techniques have been introduced that may be categorized into two main groups: classic signal processing (McFadden and Smith, 1984) and intelligent systems (Paya, Esat and Badi, 1997). To make mention of a few, FFT, Wigner–Ville distribution (Baydar and Ball, 2001), wavelets (Wang and Gao, 2003), blind source separation (Tse, Yang and Tam, 2004), statistical signal analysis (Jardine, Lin and Banjevic, 2006), and their combinations (Fan and Zuo, 2006) are classic signal processing methods.

ANN-based, GA-based, FL-based, various similar classifiers (Saravanan, Kumar Siddabattuni and Ramachandran, 2008), expert systems (Ebersbach and Peng, 2008), and hybrid algorithms (Rafiee, Tse, Harifi and Sadeghi, 2009) can be classified as intelligent systems. Currently, industrial applications of intelligent monitoring systems have increased due to the progress of intelligent systems.

The application of the Self-organizing map (SOM) (Kohonen, 2001) in system health monitoring can be a suitable choice to accomplish this goal. The organization of the paper is as follows: A first section describes the basic SOM algorithm with a deeper insight of its capabilities, citing some applications. The roller bearing health monitoring context for fault diagnosis and the associated methodology are explained in the next section. An approach, based on SOM visualization for health monitoring is suggested in the next section. The last section discusses the obtained results, followed by the conclusion.

## Materials and methods

### The SOM algorithm

The Self-Organizing Map also called SOM (Kohonen, 2001) is a neural network that maps signals from a high-dimensional space to a one- or two-dimensional discrete lattice (M) of neuron units. Each neuron stores a weight. The map preserves topological relationships between inputs in a way that neighbouring inputs in the input space are mapped to neighbouring neurons in the map space. SOM mimics the clustering behaviour observed in biological neural networks, by grouping units that respond to similar stimuli together. Nerve cells, neurons, in the cortex of the brain seem to be clustered by their function. For example brain cells responsible for vision, form the visual cortex and those responsible for hearing form the auditory cortex.

The learning rule of the SOM consists of two distinct phases: when an input $\mathbf{x}$ is presented, search for the best matching unit or bmu through competition, and the update of the codebook patterns of the bmu and its neighbours. In the basic SOM the activations of the units are inversely proportional to their Euclidean distances from the input pattern hence the bmu can therefore be defined as:

$$b(\mathbf{x}) = \arg\min_{i \in M} \|\mathbf{x} - \mathbf{m}_i\| \tag{1}$$

where $b(\mathbf{x})$ is the index of the bmu, $\mathbf{m}_i$ is the codebook vector of unit $i$ and $\mathbf{x}$ is the input pattern vector. The update part of the rule moves the bmu and its neighbours toward $\mathbf{x}$ to slightly enforce maps response to the pattern. The update rule can be written as follows:

$$\Delta\mathbf{m}_i = \gamma \cdot h(b(\mathbf{x}), i)(\mathbf{x} - \mathbf{m}_i) \tag{2}$$

where $\gamma$ is a learning rate parameter and $h(b(\mathbf{x}),i)$ captures the neighborhood interaction between the bmu $b(\mathbf{x})$ and the unit $i$ being updated. We can also write equation (2) as:

$$\Delta \mathbf{m}_i = \gamma \cdot H(\mathbf{x},i)(\mathbf{x} - \mathbf{m}_i) \tag{3}$$

where $H(\mathbf{x},i)$ is a shorthand notation for $h(b(\mathbf{x}),i)$. Equations (1) and (3) define a Hebbian learning rule, where the strength of the training step is determined not only by the learning rate parameter $0 < \gamma \leq 1$, but also by the relationship of the updated unit $i$ with the bmu $b(\mathbf{x})$ on the map.

The inter-unit relationships are captured by the neighborhood $h(i,j)$ which defines how strongly units are attracted to each other. In essence the learning rule of the SOM defines the model as a collection of competitive units that are related through the neighborhood function. In practice, the units are placed on a regular low dimensional grid and the neighborhood is defined as a monotonically decreasing function on the distance of the units on the map lattice, thus creating a latent space, which has the dimension of the map grid and flexibility determined by the neighborhood function. The SOM can produce a flawless, in the sense that the map follows the manifold, embedding when the dimension of the map grid matches the dimension of the input data manifold. A typical choice for the neighborhood function is a Gaussian:

$$h(i,j) = \exp\left[\frac{d_M(i,j)^2}{2\sigma^2}\right] \tag{4}$$

where $d_M(i,j)$ is a distance measure in the map space (M), $\sigma^2$ is the variance of the Gaussian. The radius of the neighborhood is usually but not necessarily decreased during training. Likewise, the learning rate parameter $\gamma$ is normally decreased in accordance to a predetermined cooling schedule, aiming to allow the map sufficient time and freedom to organize before fine tuning the codebook.

The SOM was used to find correlations between the data by labeling the neurons of the SOM using the training set and finding the best-matching-units (bmus) for every example. The labels of the different units correspond to a majority voting procedure, which allocates the label according to the class attribute of the proportion of the hits.

**Diagnosis of faults in rotating machinery**

Today's industry uses increasingly complex rotating machines, some with extremely demanding performance criteria. Attempting to diagnose faults in these systems is often a difficult and daunting task for operators and plant maintainers. Machine failure can lead to economic loss and safety problems due to unexpected and sudden production stoppages. In

rotating machinery, the root cause of faults is often faulty rolling element bearings. One way to increase operational reliability and thereby increase machine availability, is to monitor incipient faults in these bearings.
The use of vibration signals is quite common in the field of condition monitoring of rotating machinery. By comparing the signals of a machine running in normal and faulty conditions, detection of faults like mass unbalance, rotor rub, shaft misalignment, gear failures and bearing defects, is possible. These signals can also be used to detect the incipient failures of the machine components, through the on-line monitoring system, reducing the possibility of catastrophic damage and the down time.
The procedure of fault diagnosis starts with data acquisition, followed by feature extraction, fault detection and identification. Feature extraction is critical for the success of the diagnostic procedure. Rolling element bearings often fail due to spall and cracked defects in the inner and outer races (see an example in Fig. 1), and the rolling elements. This study shows how the measurements obtained for a faulty bearing can be mapped to a two dimensional state representation and state trajectory can be plotted to observe the system trend.

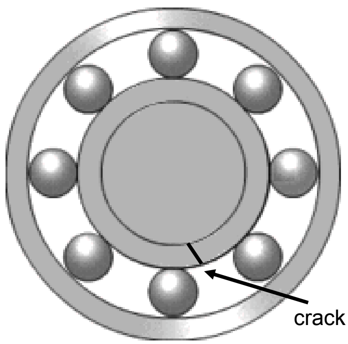



*Figure 1. Example of a crack in the inner race.*

The data were obtained from the web database of Bearing Data Center (URL: Bearing Data Center). Three types of faults conditions were accounted for,: normal bearing, bearings with 0.007 inches fault diameter and 0.021 inches fault diameter at the inner raceway. Experiments were conducted using a 2 Hp and 1Hp Reliance Electric motor, and acceleration data was measured at locations near to and remote from the motor bearings. Vibration data was recorded for motor load of 2 Hp and 1 Hp (motor speed 1750 RPM). Digital data was collected at 48,000 samples per second for drive end bearing faults for a total time of 10s per fault category and for normal bearings. Speed and horsepower data were collected using the torque transducer/encoder.

## SOM approach for monitoring fault evolution

The algorithm of Kohonen for training the SOM is a nonlinear projection method. It maps different characteristic features into the clusters on the map, without performing an explicit modeling of the system. The feature selection stage is one of the most important factors in the success of modeling. In practical situations some measurement values may be missing, due to several possible reasons, but missing data can be handled very well by SOM (Samad and Harp, 1992). Two approaches have been proposed in the literature for the application of SOM for fault diagnosis (Kohonen, *et al.*, 1996). The first approach is a supervised case where data corresponding to almost all the states, including faults, is available for training the SOM. The second approach is an unsupervised case where only the non-faulty data is available. This paper presents the case of a supervised SOM, where the data corresponding to a particular bearing fault in several stages of fault evolution has been used to train the SOM. This data is in the form of a set of two features that have been calculated based on samples from the acceleration signal. Using Kohonen's algorithm the feature data get mapped onto different regions on a 2D topographic map. Once the SOM network is trained, it is exposed to actual data from the system (in the form of the same type of features used for the training set) representing a yet unknown state. The data points are mapped onto the network as they are sequentially fed to the map describing the current state of the system. The current state can be retrieved based on the labeling procedure that has been described in a previous section. A trajectory can be drawn joining all the activated units of SOM which in turn helps to assess the fault evolution. This information can be used for fault severity assessment, for making prognostic assessments towards the remaining useful life of a system and subsequently for maintenance scheduling.

Feature extraction was performed using two features, Kurtosis and a newly proposed feature consisting of the line integral of the acceleration signal. Both provide statistical information about the nature of data, and were found to be reasonably good features for bearing fault detection. The Kurtosis is the fourth moment about the mean normalized with variance and for a sliding window of N sampling points is given by Eq (5):

$$K = \frac{\sum_{i=1}^{N} (x_i - \mu_X)^4}{N\sigma_X^4} \tag{5}$$

The proposed line integral feature for a sliding window of N sampling points is given by Eq (6):

$$\begin{aligned} LI &= \int_a^b ds \approx \sum_{i=1}^{N} \left\| \vec{r}(t_i + T_s) - \vec{r}(t_i) \right\| = \sum_{i=1}^{N} \sqrt{\left(a(t_i + T_s) - a(t_i)\right)^2 + T_s^2} \\ &\approx \sum_{i=1}^{N} \left| a(t_i + T_s) - a(t_i) \right| \end{aligned} \tag{6}$$

Where N is the number of sample points (equal to 500) in the window used to calculate Kurtosis and Ts is the sampling period. Given the high sampling rate of 48 kHz and the domination of the signal from high frequencies (especially due to the presence of faults), the final approximation contains only acceleration values.

The feature vectors are then fed to the SOM for training. To test the effectiveness of SOM, several snapshots of data corresponding to a fault evolution in a similar bearing but at different horsepower of 1 Hp were used to test the generalization of the SOM. The features of the testing set have been calculated over 10000 sampling points instead of 500 in order to increase the robustness of the features to changing conditions like the power variations of the engine. This choice did not affect the range of the Kurtosis feature but the Line integral feature needed normalization with the number of samples (10000) over the number training samples (500) in order to bring it into the same range like the one of the training set. Consequently the testing set was applied to the trained SOM so that similar map to nodes representing similar states as the input that was used for training, and when plotted give trajectories showing the system advancement during fault evolution. The implementation was done using the SOM Matlab Toolbox (URL: SOM Toolbox).

**Kurtogram approach for monitoring fault evolution**

The STFT obtained by shifting a time window along the record is here represented in terms of the amplitude envelope function H(t,f), but its square represents the power spectrum values at each time position. The average of all these short time power spectra (equivalent to the Welch method) would be the power spectrum of the whole record, or in other words at each frequency the mean square value of the output of a filter corresponding to that frequency line. The Spectral Kurtosis (SK) for each frequency f can be calculated by taking the fourth power of H(t,f) at each time and averaging its value along the record, then normalizing it by the square of the mean square value. It can be shown that if 2 is subtracted from this ratio, as given in Eq (7), the result will be zero for a Gaussian signal (Vrabie, Granjon and Serviere, 2003).

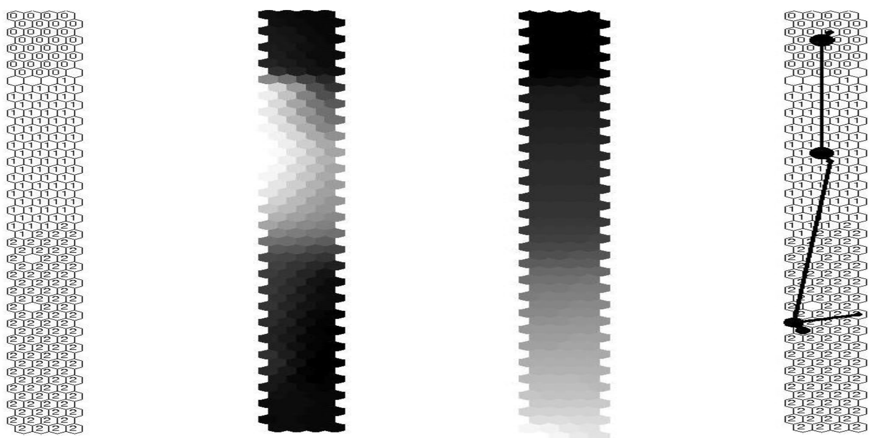
$$SK(f) = \frac{\left\langle H^4(t,f) \right\rangle}{\left\langle H^2(t,f) \right\rangle^2} - 2 \tag{7}$$

Since SK is large in frequency bands where the impulsive bearing fault signal is dominant, and effectively zero where the spectrum is dominated by stationary components, it makes sense to use it as a filter function to filter out that part of the signal with the highest level of impulsiveness. For the hypothetical case of a series of impulses mixed with stationary noise, Antoni and Randall (2006) show that the optimum Wiener filter (maximizing the similarity between the filtered component and the true noise-free signal) is the square root of SK. They also show that the optimum matched filter (maximizing the signal/noise ratio (SNR) of the filtered signal, without regard to its shape) is a narrowband filter at the maximum value of SK. However, the optimum result in a given case may vary with both the centre frequency and bandwidth of the filter; in Antoni and Randall (2006), a display showing the optimum combination was called the 'kurtogram'. Computation of the full kurtogram covering all combinations of centre frequency and bandwidth is very costly and so a number of more efficient alternatives have been proposed. The most detailed study is presented by Antoni (2006), who proposes the 'fast kurtogram', based on a series of digital filters rather than the STFT. The gains in computational speed are based on a dyadic decomposition, rather similar in principle to the FFT, and even more similar to the discrete wavelet packet transform (DWPT). In the most basic version, the frequency range is progressively split into bands that are one-half the width of the previous stage (or scale), a so-called binary tree. The version recommended, however, is the '1/3-binary tree', where the split includes divisions into 1/3, so that the overall division is in the sequence 1/2, 1/3, 1/4, 1/6, 1/8, 1/12, and so on.

**Results**

For the experiments in this paper a map size of 4x67 was used. This specific SOM shape has been selected because it enables the mapping of the evolution of feature combinations corresponding to fault evolution in a way that were different fault classes form distinct zones on the SOM. Following a voting procedure for the allocation of labels, the resulting allocation to units of the SOM in Fig. 2(a) show clear clusters that are formed on the SOM with associated classes corresponding to fault evolution.

The component maps in Figs. 2(b) and 2(c) show the distribution of Kurtosis and line integral values respectively for the units of the SOM. The distribution of Kurtosis as shown in Fig. 2(b) indicates a sensitivity of Kurtosis to minor faults compared to normal bearings and more extended faults (of 0.021 inches width). On the other hand, the proposed feature of the line integral is more sensitive to extended faults as shown in Fig. 2(c). The fusion (by direct concatenation) of the two features, due to its complementary nature, results in a more accurate separation of classes related to the fault extent. This is evident from the mapping of fault classes in the SOM representation which shows the labels of each SOM unit in Fig. 2(a). The testing of the trained SOM with fault data from a similar bearing running at a power of 1 Hp shows that the 30 testing samples form a trajectory on the trained SOM as shown in Fig. 2(d) that falls by 100% in the correct corresponding classes indicating normal and faulty conditions.



(a) (b) (c) (d)

*Figure 2. Different manifestations of the features are shown: (a) Assigned labels corresponding to different fault widths (0: no fault, 1: width 0.007 inches, 2: width 0.021 inches). (b) Component map corresponding to Kurtosis. (c) Component map corresponding to Line integral. (d) Formation of trajectories for 30 data vectors for testing correct classification of faults. All faults have been assigned to clusters having same class as real fault.*

The same data as in the SOM training were used for calculating fast kurtograms for each category of fault. The corresponding fast kurtograms for intact bearings and bearings with fault widths of 0.007 and 0.021 inches respectively are shown in Figs. 3,4 and 5. From the figures it is evident that the transient character of the signal is stronger when faults are present. The SK is minimal with a maximum of 1.1 in the case of intact bearings (Fig 3) while it acquires high values of maximum 7.9 when faults are present (Fig 4). In the case of a larger fault the value of SK reaches values of 858.5 which means that very large transients dominate the signal.

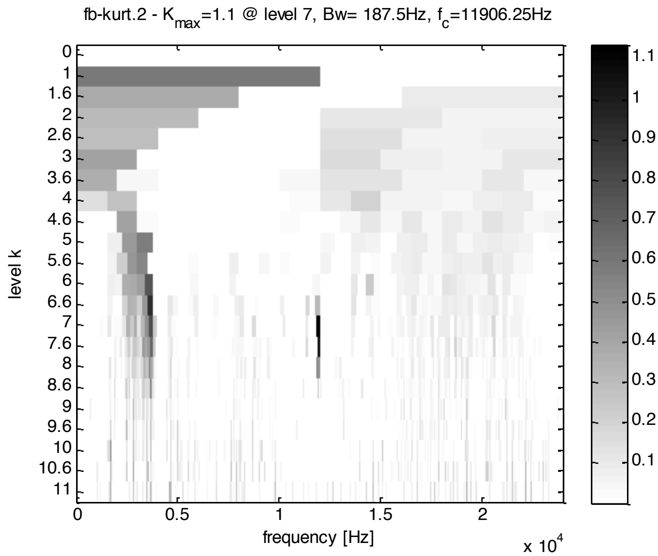



*Figure 3. Fast Kurtogram for intact bearings.*

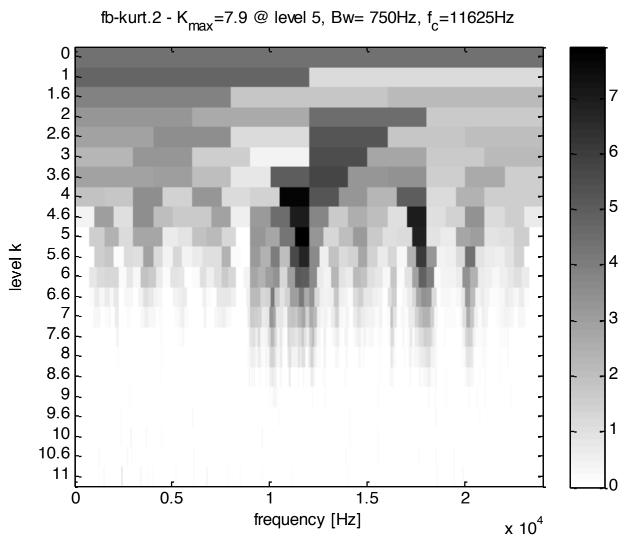



*Figure 4. Fast Kurtogram for bearings with faults of 0.007 inches.*

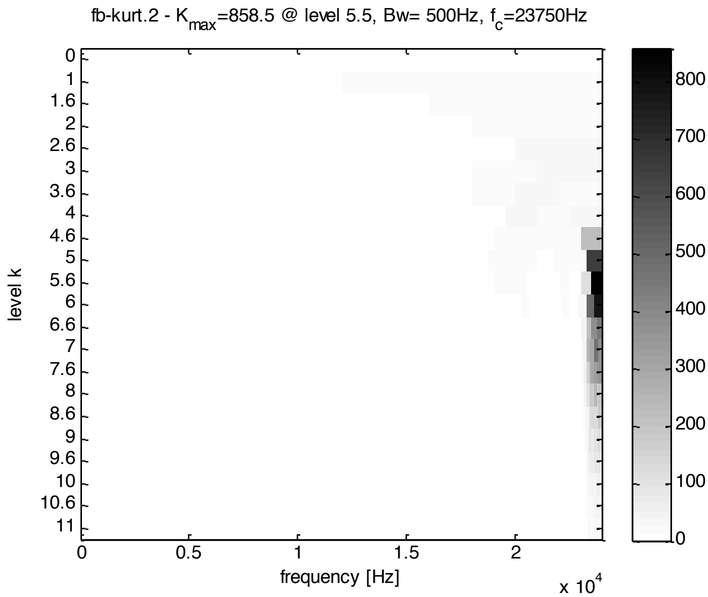



*Figure 5. Fast Kurtogram for bearings with faults of 0.021 inches.*

## Discussion

The methods of SOM and Spectral Kurtosis (in the form of a kurtogram) have been presented for monitoring the development of faults in bearings of centrifugal irrigation pumps. The kurtogram shows a clear development of large transients due to fault development. The SOM provides a means of visualizing fault development through fusing different features seamlessly in a visualization interface. Not all features are sensitive to specific faults in a similar fashion. Data from different features, like kurtosis and line integral map differently on a SOM, depending upon how they perceive the change in the vibration signal. Some features that can not detect a fault can be recognized by the fact that even for small fault levels, they show high feature values and for large faults they do not necessarily approach high values. Moreover, this idea can also be extended to the cases where a-priori information about faults is not available. A suitable measure like the quantization error can be defined, if it increases beyond a certain threshold for new input vectors, an alarm can be raised to indicate that the system is approaching a faulty state. Other approaches implement growing SOM variants, determining the number of nodes on the map based on the input vectors.

## Conclusion

Bearing faults are a common reason of failures in centrifugal irrigation pumps and the prevention of such failures can enable predictive maintenance. It has been shown that the SOM can be used to detect faults in roller bearings and assess the severity of the faults, and can therefore prove to be a powerful tool for bearing health monitoring, especially regarding monitoring of fault evolution. This approach can enable maintenance scheduling. The proposed technique can be used for both supervised and unsupervised learning. User-friendly and fairly comprehensible visualizations are easy to monitor and provide an indication as to which features (from the same or different sensors) respond to different fault stages, and under what conditions. Different stages of fault evolution may be indicated by a collective response or fusion of several features or sensors, which may not be obvious by just looking at the data using other diagnostic techniques. An area of future research is to use hierarchical SOM models to discover if the data contains any inherent hierarchy.


## Acknowledgements

The support of the Greek Scholarships Fund is gratefully acknowledged.


## References


Antoni, J. and Randall, R.B., 2006. The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. *Mechanical Systems and Signal Processing*, 20(2), 308–31.

Antoni, J., 2006. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1), pp. 108–24.

Baydar, N. and Ball, A., 2001. A comparative study of acoustic and vibration signals in detection of gear failures using Wigner–Ville distribution. *Mechanical Systems and Signal Processing*, 15(6), pp. 1091–107.

Ebersbach, S. and Peng, Z., 2008. Expert system development for vibration analysis in machine condition monitoring. *Expert Systems with Applications*, 34(1), pp. 291–99.

Fan, X. and Zuo, M. J., 2006. Gearbox fault detection using Hilbert and wavelet packet transform. *Mechanical Systems and Signal Processing*, 20, pp. 966–82.

Han, T., Yang, B.S., Choi, W.H. and Kim, J.S., 2006. Fault diagnostic system of induction motors based on neural network and genetic algorithm using stator current signals. *International Journal of Rotating Machinery*, 1, pp. 1–13.

Jardine, A. K. S., Lin, D. and Banjevic, D., 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), pp. 1483–510.

Khan, A.F. and Williams, E.J., 1992. Predicting the remaining life of rolling element bearings. In: *International conference on vibration in rotating machinery.* Bath, pp. 403–7.

Kohonen, T., 2001. *Self-organizing Maps*. Berlin, Heidelberg, New York: Springer-Verlag.

Kohonen, T., Oja, E., Simula, O., Visa, A. and Kangas, J., 1996. Engineering Applications of the Self-Organizing Map. *Proceedings of the IEEE*, 84(10), pp. 1358–84.

Loparo, K.A., Adams, M.L., Lin, W., Abdel-Magied, M.F. and Afshari, N., 2000. Fault detection and fault diagnosis of rotating machinery. *IEEE Transactions on Industrial Electronics*, 47(5), pp. 1005–14.

Mathew, J. and Alfredson, RJ., 1984. The condition monitoring of rolling element bearings using vibration analysis. *Transactions of the ASME Journal of Vibration Acoustics Stress and Reliability in Design*, 106 (5), pp. 447–53.

McFadden, P. D. and Smith, J. D., 1984. Vibration monitoring of rolling element bearings by the high-frequency resonance technique – A review. *Tribology International,* 17(1), pp. 3–10.

Paya, B. A., Esat, I. I. and Badi, M. N. M., 1997. Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor. *Mechanical Systems and Signal Processing*, 11(5), pp. 751–65.

Peck, J. P. and Burrows, J., 1994. On-line condition monitoring of rotating equipmentusing neural networks. *ISA Transactions,* 33, pp. 159–64.

Prabhu, B.S., 1996. Rolling bearing diagnosis. In: *Indo–US symposium on emerging trends in vibration and noise engineering.* Delhi, India, pp. 311–20.

Rafiee, J., Tse, P. W., Harifi, A. and Sadeghi, M. H., 2009. A novel technique for selecting mother wavelet function using an intelligent fault diagnosis system. *Expert Systems with Applications*, 36, pp. 4862–75.

Sakthivel, N. R., Sugumaran, V. and Babudevasenapati, S., 2010. Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. *Expert Systems with Applications*, 37, pp. 4040–49.

Samad,T. and Harp, S. A., 1992. Self-Organization with Partial Data. *Network: Computation in Neural Systems*, 3(2), pp. 205-12.

Saravanan, N., Kumar Siddabattuni, V. N. S. and Ramachandran, K. I., 2008. A comparative study on classification of features by SVM and PSVM extracted using Morlet wavelet for fault diagnosis of spur bevel gearbox. *Expert Systems with Applications*, 35(3), pp. 1351–66.

Singh, G.K and Kazzaz, AS.AK., 2003. Induction machine drive condition monitoring and diagnostic research—a survey. *Electric Power System Research*, 64(2), 145–58.

Tse, P. W., Yang, W. X. and Tam, H. Y., 2004. Machine fault diagnosis through an effective exact wavelet analysis. *Journal of Sound and Vibration*, 277, pp. 1005–24.

URL: Bearing Data Center: http://www.eecs.case.edu/laboratory/bearing/CWRUFlash.htm

URL: SOM Toolbox: http://www.cis.hut.fi/projects/somtoolbox/

Vrabie,V.D., Granjon, P. and Serviere, C., 2003. Spectral kurtosis: from definition to application. In: IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing. Grado, Italy, 8–11 June.

Wang, C. and Gao, R. X. 2003. Wavelet transform with spectral post-processing for enhanced feature extraction. *IEEE Transactions on Instrumentation and Measurement*, 52(4), pp. 1296–301.

Xu, Z., Xuan, J., Shi, T., Wu, B. and Hu, Y., 2009. Application of a modified fuzzy ARTMAP with feature-weight learning for the fault diagnosis of bearing. *Expert Systems with Applications*, 36, pp. 9961-68.

Zarei, J. and Poshtan, J., 2007. Bearing fault detection using wavelet packet transform of induction motor stator current. *Tribology International*, 40, pp. 763–69.

# VLIT NODE – new technology for wireless sensor network

Zbynek Krivanek, Karel Charvat, Jan Jezek, Marek Musil
*Czech Centre for Sceince and Society*
E-mail: krivanek@ccss.cz

**Introduction**

In the last decade an increasing importance of meteorological measurements in agriculture are stressed due to rapid weather changes. There is a need for immediate access to relevant local weather information and the possibility of its on-line analysis. There are currently no meteorological data of sufficient extent available for managing crop production. The data should be included in specific agro-meteorological models. There is also lack of adequate systems and support of the management of agricultural production based on agro meteorological models. There is a need to establish a network of local sensors and weather stations in order to support farmers. Advanced technology of in-situ monitoring must be supported by the development of skills of assimilation with the agro-meteorological models in real time. At the same time, integrating measurements from meteorological sensors and data from existing agro-production networks opened new opportunities for farmers to increase the quality of their production. This will increase their market competitiveness. Monitoring of agro meteorological elements has a strong influence on the management of growth and development of crops, but also reflects the dynamics of other important biological features including plant disease and pest incidence. It is necessary to develop methods to monitor the development of climate data collection and to assess weather conditions. Representative samples from meteorological stations located on land monitoring data must then be integrated into the production database.

Typical applications for crop growth control are aimed to assess the availability of nutrients in the soil (e.g. Prefarm system) and the control of diseases and pests in the soil and crops throughout the growing season until harvest. Integrating measurements of soil and environmental conditions in the agronomic decision support model for optimal planning of activities is missing. Assimilation of real data (from sensors and satellites) in the growth model is currently the subject of research (e.g. the Future Farm project in the 7th Framework Programme for Research and Development). In recent years, there were proposed several approaches to data assimilation in real meteorological models. They are using methods including 3D-Var, 4D-Var and Kalman filter-Ensamble. Each of these methods has its advantages and disadvantages that should be evaluated in small scale models. In practice this means that the farm management system (FDMS) used by farmers and / or service organizations (Prefarm) are supported by a comprehensive system of data collection (CDC). They integrated end-user through the visualisation of geographic information system (GIS). The effectiveness of all FDMS can be evaluated based on their ability to provide relevant, accurate and timely information (e.g. weather, soil conditions and crop condition in real time).

Basic requirements to be addressed are:

- To design an optimal sensor network for collecting agro meteorological data at farm level density, which can monitor local impacts;
- To design optimization parameter measurements guaranteeing the possibility of local modelling;
- To ensure communication between sensors and data transmission via mobile unit to the server;
- To standardise interface for integration of sensor measurements with other data;
- To integrate with Prefarm through open OGC protocols;
- Modularity solution.

**Sensor networks**

Currently, there is a number of technologies, protocols and standards for building wireless sensor networks. Sensor networks are generally seen as cloud of mutually communicating measurement units that are capable of measuring one or more physical parameters. Each measuring unit consists of a communication node ensuring the communication with other units of measurement and its own sensor. The communication nodes are built on different platforms. Their drawback is that they are able to guarantee the communication between sensors of only tens of meters. This reduces the network ranges and the networks are not affordable.

Sensor Network Systems provides a novel paradigm for managing, modelling and supporting complex systems requiring massive data gathering. It has pervasive and persistent detection/monitoring capabilities. In recent years, a growing emphasis has been steered towards the employment of sensor networks in various technological fields e.g. aerospace, environment monitoring, homeland security, smart buildings. A significant amount of resources has been allocated for national (e.g. USA, France, Germany) and international (e.g. European Commission) research programmes targeted at developing innovative methodologies and emerging technologies in different application fields of wireless sensor network. The main features that a sensor network should have are:

- each node should have a very low power consumption, the capability of recharging its battery or scavenging energy from the environment, and very limited processing capabilities;
- each node should be allowed to go in stand-by mode (to save as much battery as possible) without severely degrading the connectivity of the whole network and without requiring complicated re-routing strategies;
- the estimation/measurement capabilities of the system as a whole should significantly outperform the capabilities of each sensor and the performance should improve as the number of sensors increases, with no mandatory requirement on the transmission of the data of each single sensor towards a centralised control/processing unit; in

other words, the network must be scalable and self-organising, i.e. capable of maintaining its functionality (although modifying the performance) when the number of sensor is increased;

- a sensor network is ultimately an event-driven system, important is to guarantee that the information about events of interest reaches the appropriate control nodes possibly through the simplest propagation mechanism not necessarily bounded to the common OSI protocol stack layer;
- congestion around the sink nodes should be avoided by introducing some form of distributed processing;
- the information should flow through the network in the simplest possible way, not necessarily relying on sophisticated modulation or multiplexing techniques.

Summarising, the fundamental requirements of a sensor network are:

- Very low complexity of elementary sensors, associated with a low power consumption and low-cost;
- High reliability of the decision/estimation/measurement of the network as a whole;
- Long network life-time for low maintenance and stand-alone operation;
- High scalability;
- The resilience to congestion problems in traffic peak conditions.

An extensive research and development work has been done in the past. It includes ensuring information technology use in agriculture; long range wireless sensor network creation for specific agricultural use, ensuring a PA technological leap, solving pressing problems for agriculture and making PA widely available for farmers, even for low scale use (cranberry fields, fruit gardens, bee-gardens etc.). However, for existing solutions the following problems remain:

- Existing WSN solutions are in experimental development phase; their implementation is not possible without the specific WSN technology developers' assistance.
- Existing WSNs have a short working range (ability to guarantee communication between sensors only at a range of several tens of meters); therefore their implementation in large area is very expensive.
- Existing WSN technology application programming is not possible without deep WSN operating system (open source Tiny OS, commercial ZigBee etc.) knowledge, that is possible only in specialised development centres;
- Presently known WSN physical node technologies with several hundred meters working range don't support available operating systems;
- Existing WSNs are not suited for climatic and geographical factors as well as production manufacturing problems;

- Realistic WSN implementation is unthinkable without specific WSN technology that includes physical nodes, sensors, operating system, application programming environment and competence centre support.

It is clear, that new development is necessary. Development would include:

- New sensor nodes with communication ranges of 200-800m depending on environment, weather conditions and sensor location, that are suited for use in most of the European countries;
- Development of operating system programming that would collect data from sensor nodes and transport them via wireless network to base computer, communication protocol configuration that would comply with respective usage target environment, as well as specific usage application programming development in the utmost simplified environment (in language C with possibly minimal specific knowledge about operating system and WSN physical realisation) that would ensure sensor control and communication between sensor nodes.

**Requirements for agriculture sensor networks**

From the analysis of projects there are the following requirements:

- The size of transmitted data (own measurements of sensors i.e. type sensor and the measured value) - approx 6B for one variable and one measurement;
- Measurement will always be started by the AP (gate);
- The data will be reading the AP point in periodic intervals, which can be configured. The shortest possible interval is 60 seconds;
- It is not necessary to encrypt the data due to their nature;
- Radio reach point must be at least 300 meters in open terrain and line of sight (the field, meadow);
- Estimated vastness of the network will be in the tens of elements. The elements must withstand work without charging at least 6 months (roughly agricultural season) provided data collection at least once every 2 hours;
- Easy operation of network elements (nodes);
- Mechanical resistance network elements.

As a minimum solution is required:

Communication in the range of about 500 m and 250 m in the field of forest;

- Two modes of operation - on-demand, event-driven;
- Support for multihoping;
- Integration of memory;

- The possibility of simple computational operations;
- Easy integration of sensors measuring;
- Life of at least 6 months;
- The ability to connect to existing mobile solutions;
- Integration into the Web environment;
- Operating frequency of 868 MHz;
- It uses a very sophisticated two-way communication protocol that allows bi-directional data transfer;
- The anti-collision protocol thus allowing to scan multiple nodes simultaneously in the field of an antenna (up to 256 tags);
- The sensor contains a function measuring the level of the incoming signal (RSSI), which will extend battery life;
- Communication protocol supports Point-to-Point, Point-to-MultiPoint and retranslation of the large distance across multiple devices.

**Sensor VLITnode implementation**

The development of thr second generation RFID offers the possibility to create a new generation communication nodes using RFID technology. Cominfo ltd. developed RFID technology with unique properties based on long-range communication and cost-effectiveness. The technology known as Very Long Range Identification Tag is characterized by a working frequency of 868 MHz and protocol that supports communication in Point-to-Point, Point-to-MultiPoint and retranslation of the large distance across multiple devices. In combination with the mobile unit and the software interface is generated by Research Centre CCSS. vLite NODE represents a completely new and unique solution for the construction of mobile sensor networks.

The node consists of two parts. The first is the host board for connecting the communication module and pour connectivity for data line sensors. The second part is the electronics that provides controlled power sensors and the module itself to achieve minimum energy consumption.

When are transferred the measured data with standard data packet the is also carried information about the signal strength (RSSI) and the voltage level of node. Receiving of data in the observation area provides network access point that deals with other communication and data transfer.

After consultation with representatives of the Agricultural and Forestry University, CCSS and Cominfo, we designed and manufactured first prototype of VLIT nodes. After the node testing, it became necessary to equip the sensor with temperature, humidity and radiation cover to avoid readings influenced by direct sunlight and wind. Currently we are working on a new version of VLITNode. One of the reasons for design changes were adding external antenna to the upper part of node to increase range of transmission. Another modification is the use of bottom circular connectors for connecting the soil moisture sensor.

**Mesh protocol**

One of the goals of the VLIT Node project is to build an extensive network of wireless sensors communicating with the MESH topology. MESH topology enables connection of nodes to any other node in the network. This connection can be established using one or more hops. As part of the MESH topology is provided automatic configuration of network structure, reliable routing between nodes and automatic access to new nodes in the network via the existing nodes. Hop identifies the network segment, where all participants can communicate to each other without the need for routing. Multi-hop network is a network composed of several such segments, where information could be routed among the nodes. In the area of wireless networks AH-HOC is used. AH-HOC is a network where actors do not require any pre-created infrastructure to be able to communicate with each other and it provides the necessary functionality for the network management.

The main benefit of using mesh topology is the possibility to form redundant links, due the nature of network topology guarantee transmission of information. Mesh topology is not restrictive in the network structure and therefore simplifies the automatic compilation of links and network recovery after failure. The connection between two points in a full mesh topology can be set up whenever they are able to communicate. Mesh topology can be set up almost always. Implementation of mesh networks in practice is highly dependent on the method of communication, the technical and application requirements.

Data collection in Agriculture is a major area of use of mesh networks and is often the only suitable solution for monitoring of large areas with a large number of sensors or as an additional source of data in places where fixed connections are difficult to install. Mesh networks are divided according to whether they are mobile or stationary, wireless or wired, occasional or defined (e.g. sensory). Each type of MESH network can solve a specific protocol, which is mainly different algorithm to find and build paths from the data source to the destination.

Firmware microprocessor module VLIT can generally be divided into several general programme blocks.

Mesh networks are a way to transmit data, voice and commands between nodes. They allow continuous connections and reconfiguration around the fallen or blocked paths by jumping from node to node until it is achieved. MESH network whose nodes are all interconnected with other nodes is fully connected network. Mesh networks differ from other networks in the fact that the parts can all connect to each other. Each node MESH network can be a router. Mesh network can be viewed as a type of temporary or occasional (ad hoc) network. Mobile ad-hoc network (MANET) and mesh networks are thus closely linked, but pose problems of MANET nodes mobility.

Mesh networks are self-healing. The network can remain in operation whenever any node fails or drops the connection. The result is large network reliability. This concept is applicable to wireless networks, cable networks, and software interaction.

Wireless Mesh Networks are the highest rank of MESH networks. They were originally developed for military applications but experienced great development. Design of Mesh networking nodes has become more modular - a single node can support multiple radio cards - each working on different frequencies.

Proactive algorithms require enough memory for routing tables. In these data are stored to reach any network node. The main problem of the algorithm is then given by constructing a routing table and its updates. Their main disadvantage is the memory consumption and slow reaction to changes in the network structure.

Reactive algorithms have low memory requirements, because they do not store routing algorithms for all network nodes, or even no routing information. Each connection is established just before the data transfer. Then the connection is terminated. The connection is omnidirectional. Their main disadvantages are large unexploited time in search of connection and network congestion at risk of broadcast queries.

Hybrid algorithms are used as the routing table establishing a connection before transferring data. These algorithms have been developed in a large amount of effort to optimize the memory requirements of nodes, the need for frequent updating of routing tables (optimization in time and space), minimizing broadcast queries to build path.

**Communicating with a Web interface**

The communication protocol used in networks was implemented by vLite to mobile units proposed under the project WINSOC. The communication protocol was modified for the project node and vLite and it ensured the communication between mobile sensors and web environment. Custom integration of measured data is ensured through the Open Geospatial Consortium Enabling Sensor Web (OGC SWE). In this area a cooperation is envisaged with running international projects in the GEOSS (EnviroGrid BlackSee), GMES (e.g. Humboldt) and 7FP ICT (SANY, GENESIS, etc.). The project compared two concepts, the implantation of this standard to ensure interoperable server to access data at the server level and in the second variant, directly implementing this standard in mobile units. In the first solution is the implementation easier and ensures achievement of desired functionality. The second solution would represent a significant shift in the integration of sensor data into information systems (plug and play) and a significant increase of independent sensor networks. It would also mean a generational shift in the possibilities of the mobile unit and its applicability usability.

**The area storing and accessing data**

For subsequent data processing a database was designed corresponding to the current standards and recommendations of OGC and SWE for collecting spatial data and sensor. It can optimally process data services through Sensor Observation Services, Sensor Alert Services, Notification Services, etc.

## Results

Currently, there were developed approx. 40 prototypes of sensor nodes and the deployment and filed testing started. Intensive field testing was provided in the Czech Republic. At this time (end of 2011), we installed the first testing node in Latvia for a customer. A testing in Italy in2012 is expected. Increasing the communication distance of sensors network from 50 meters to 500 meters will decrease the number of sensors necessary 100 times. It is an important aspect for agriculture application, where the first experiences demonstrate that use of sensors with communication distance between 500 meters and 1000 meters will be optimal distance to cover a farm. The density of sensors allows operational use of sensor network.


## Acknowledgement

*VLIT NODE* - the solution was achieved with financial support from state resources provided by the Ministry of Industry and Trade of the Czech Republic; the project of the program "TIP-2009" with registration number FR—TI1/523.


## References


Charvat *et al.* 2008, Spatial Data Infrastructure and Geovisualisation in Emergency Management, H. Pasman and I.A. Kirillov (eds.), Resilience of Cities to Terrorist and other Threats, Springer Science + Business Media B.V. 2008

Charvat *et al.*, 2009, INSPIRE, GMES and GEOSS Activities, Methods and Tools towards a Single Information Space in Europe for the Environment
Riga, Latvia

Charvat *et al.* 2010, enviroGRIDS sensor data use and integration guideline
www.envirogrids.net

Gnip *et al.* 2008, In situ sensors and Prefarm system(p.255-262), conference proceeding book, IAALD AFITA WCCA2008, Tokyo , Japan.

Wilson, 2005, Sensor Technology Handbook, Elsevier Inc, UK

# Selected Issues of Wireless Sensor Networks Geovisualization in Agriculture

Petr Kubíček[1], Vojtěch Lukas[2] and Jiří Kozel[1]
[1]*Geographic institute, Faculty of Science, Masaryk University,* [2]*Department of Agrosystems and Bioclimatology, Faculty of Agronomy, Mendel University in Brno*
E-mail: kubicek@geogr.muni.cz

## Introduction

The dynamics of plant growth is depended on the dynamics of environmental data as temperature and humidity of soil (or air) and solar radiation which can be measured stationery using local weather station or wireless sensor networks (WSN). WSN are a new technology that can provide processed real-time field data from sensors physically distributed in the field (Camilli *et al.* 2007). Akyildiz *et al.* (2002) described WSN as a networks of small sensor nodes, with limited processing capacity, including sensors and their specific conditioning circuitry, that communicate over short distances, normally using radio frequencies. The combination of high spatially data from remote or proximity sensing and real-time monitoring of dynamics of agro-meteorologically variables from WSN may provide a complex overview of spatial-temporary variability of crop related phenomena.

## Wireless sensor networks and agriculture

Wang *et al.* (2006) presented an overview on recent development of wireless sensor technologies and classified the applications in agriculture and food production into 6 categories:

- environmental monitoring - weather monitoring, geo referenced environmental monitoring;
- precision agriculture - spatial data collection, precision irrigation, variable-rate technology;
- machine and process control - vehicle guidance, machinery management, robotic control, process control;
- building and facility automation - greenhouse control, animal-feeding facilities;
- RFID-based traceability systems - animal identification and health monitoring, food packaging, transportation, food inspection.

Besides naming existing WSN applications and recognising the great potential of various applications they also named several obstacles to be passed. Among the most important were:

- Insufficient standardization
- Compatibility with legacy systems

- Security issues
- High costs for coverage in large plants
- Power supply insufficiencies
- Reliability of wireless systems

Some of these will be further commented in section 3.

Sensor Networks vary in their scale and function. Hart and Martinez (2006) described four principle types of Environmental Sensor Networks (ESN) automatically sensing the environment:

- Large Scale Single Function Networks - cover large geographic areas; take measurements for a single purpose and large, expensive nodes. These comprise (usually large) nodes that normally measure one or more variables and have been networked together. The simplest of these are weather stations, and more complex examples include the Global Seismographic Network.
- Localised Multifunction Sensor Networks - typically comprise smaller nodes to monitor the environment. These nodes normally measure relatively straightforward generic properties (temperature, humidity etc) which can be used in a variety of applications, however, these are more likely to be ad hoc systems and use 'intelligent' networking.
- Biosensor Networks - emerging set of systems which are distinguished mainly by their use of biotechnology. A biosensor comprises a biological sensing element attached to a physical transducer, which can be electrochemical, optical electronic, optical or acoustic. This type of ESN generally requires new sensor technology to test air, water and soil. Many pollutants are currently not monitored in situ but samples collected on site and later analysed in a laboratory.
- Heterogeneous Sensor Network - This type of network would include the data sources from the different types of ESN described above to monitor the environment at different scales.

Hart and Martinez (2006) further described the properties and examples of aforementioned WSN.

Environmental data are very important in agriculture, since crop yields depend on environmental conditions, and the response of plant growth to changing environmental conditions is extremely complicated (Lee *et al.* 2010). The measurements of agronomical phenomenon are mainly required for modern agricultural practices called precision agriculture. The aim of precision agriculture is an optimization of production inputs (fertilizers, pesticides, fuel, etc.) based on the local crop requirements and plants requirements. Crop management in this way can lead to the effective use of agrochemicals and avoid of environmental risks.

The aspects of precision agriculture are described by Pierce *et al.* (1999). They defined precision agriculture as “the application of technologies and principles to manage spatial and temporal variability associated with all aspects of agricultural production for the purpose of improving crop performance and environmental quality”. Assessing variability is the critical first step since because one cannot manage what one does not know. The factors and properties that regulate crop growth and yield vary in space and time. The higher is the spatial variability of a soil conditions (or crop properties), the higher is the potential for precision management and the greater its potential value. The degree of difficulty, however, increases with higher dynamics of temporal component.

Precision agriculture relies upon (Oerke and Gerhards, 2010):

- intensive sensing of environmental conditions in the crop
- extensive data handling and processing
- use of decision support systems (DSS)
- control of farm machinery in the field

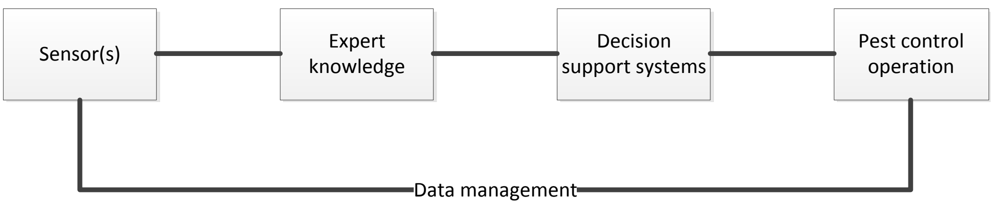



*Figure 1. Information technologies and data management for precision crop protection (Oerke and Gerhards, 2010)*

Techniques for assessing spatial variability are readily available and have been applied extensively in precision agriculture. Most common techniques are based on the spectral and thermal properties of plants and soil (remote or ground based optical sensing), geophysical and mechanical properties of soil (on-the-go measurement of soil electrical conductivity or soil resistance), crop yield mapping and many others. Techniques for assessing temporal variation also exist but the simultaneous reporting on spatial and temporal variation is rare and the theory of these types of processes is still in its infancy (Pierce *et al.* 1999).

Integration of sensors and geospatial infrastructure

Open Geospatial Consortium (OGC) has recently started an initiative called Sensor Web Enablement (SWE) for building a framework of open standards for exploiting Web-connected sensors and sensor systems of all types (Reichardt 2005). SWE presents many opportunities for adding a real-time sensor dimension to the Internet and integrate both the sensor and geospatial information in an interoperable way.

The initial focus of OGC's SWE has been to investigate standardized interfaces for live sensors operating in near-real-time, rather than the conventional static data stores. It addresses information gathering from distributed, heterogeneous, dynamic information sensors and sources of different structure, based on web services. It is the goal to develop common access, planning, and management interfaces and a descriptive mark-up language (SensorML) for managing sensor information and metadata in common consistent manners, independent of any application. The individual parts were initially designed to fulfil the following needs (Simonis *et al.*, 2003):

- Describe sensors in a standardized way
- Standardize the access to observed data
- Standardize the process of what is commonly known as sensor planning, but in fact is consisting of the different stages planning, scheduling, tasking, collection, and processing
- Building a framework and encoding for measurements and observations

Some sensors are already on the Web and able to return their location information as well as observations and measurements. The final missing element - a universal standard framework for describing and tasking sensors in XML - has already been built and prototyped by OGC members in specifications like Sensor Observation Services (SOS) or Sensor Planning Services (SNS) (Simonis 2008).

Since 2007, OGC members have been working on the second version of SWE. SWE 2.0 brings higher coherency of individual SWE specifications together with conceptual and formal clarity rather than new features. At the moment, only the SWE Common Data Model Encoding Standard (OGC 2011) is available in the version 2.0, nevertheless other SWE specifications are expected to be published during 2011. The most expected specifications are SensorML 2.0 and SOS 2.0, because the first one specifies XML encoding of SWE and the second one provides specifies access to measured sensor values. Observations and Measurement 2.0 as the first of the SWE standards family has become an ISO standard (ISO 19156).

## Visualization of sensor data

Visualizations are important when working with sensor data, it makes it more comfortable for a user to work with the data and the data can be understood faster and easier. With visualization it is possible to find patterns, connections or similarities in numerical data. That makes it a lot easier then to manually analyze the raw sensor data, which is sometimes impossible to understand for a person.

Sensor taxonomy or classification for the field of applications where sensors are used (see White 1987, Richter 2009, Zerger *et al.* 2010 for examples) can help a developer when implementing a visualization. With knowledge about the sensed value he or she can use fitting colour scales, create specific shapes.

## Conceptual background

Data like sensor data normally exists in numerical values so the process of understanding or analyzing it is not trivial, finding patterns, differences, and commonalities seems hardly possible. Visualizations take the numerical data, analyze it, use extraction and aggregation methods and then give the user a graphical interface and representation of the input data. This representation makes it easier to understand the data and to interact in the data set (Richter 2009). According to Plaisant (2004) the biggest problem when designing visualizations is the fitting of visualization to the wishes of the user, to the task and to real world problems. Trying to display the data there should be different views on the same data set. These features have to be implemented and the properties of sensor data have to be regarded to have a good visualization of sensor data. There is a dependency between the task and the visualization, so the *task of the sensor* is one starting point.

Another starting point can be found *in the topic of information visualization*. Several authors (Richter 2009, Andrienko and Andrienko 2006) are quoting Shneiderman (1996) and his "Visual Information Seeking Mantra". He says that this is a starting point to create a good visualization and proposes *type by task taxonomy* (TTT). *Types* are conceptually represented by seven data types (1-, 2-, 3-dimensional data, temporal and multi-dimensional data, and tree and network data) with associated attributes. These basic data types are searchable and selectable by attribute. On the other hand he enumerated seven *tasks* at a high level of abstraction, when creating a design, these are:

- Overview - gain an overview of the entire collection.
- Zoom – Zoom in on items of interest.
- Filter – filter out uninteresting item.
- Details-on-demand – select an item or group and get details when needed.
- Relate – View relationships among items.
- History – keep a history of actions to support undo, replay and progressive refinement.
- Extract – allow extraction of sub-collections and of the query parameters.

These steps are further explained in detail in Schneiderman (1996). Andrienko and Andrienko (2006) further developed the basic four principles (overview, zoom and filter, and details on demand) for exploratory data analysis. They however used a different naming for principles – "see the whole", "zoom and focus", and "attend to particulars". The main principle is to give an overview of the whole data set. To take a closer look at special data there are tools that allow zooming functions or which filter away uninteresting data. If the user wants to have details about the data, there should be a possibility to get these details (Richter 2009).

Multidimensionality is one of the biggest problems when visualizing sensor data (Cook 2007). A lot of sensors measure more than one kind of data at the same time which leads to the usage of multiview visualizations. While the first view shows temperature data of a single sensor, another view displays the sensor's humidity value and a third view could for example show a map of all sensors spatial location and their last measured temperature values. The target of this multiple view displays is to reduce the dimension of the data by showing it in more than one view.

**Agrisensor – from concept to pilot study**

The AgriSensor project (Kubicek *et al.* 2009) follows the above mentioned concept and aims to design and develop an integrated framework of dynamic cartographic visualisation for agricultural applications based on wireless sensor networks information. The terms "dynamic visualization" describes mainly the possibility to select required scale, interactive communication of users and map authors over the Internet, and also presence of simple modelling tools for spatial as well as temporal analysis. The suggested framework is based on following components (see fig. 2):

(1) Heterogeneous distributed network of hierarchical agricultural sensors
(2) Communication infrastructure and standardised interfaces between sensors and the Internet
(3) Web enabled geoinformation infrastructure for an effective visualization of sensor data
(4) Cartographic visualization rules and modelling tools for an effective support of agricultural decision making, and
(5) Agricultural knowledge base and data warehouse of legacy sensor data.

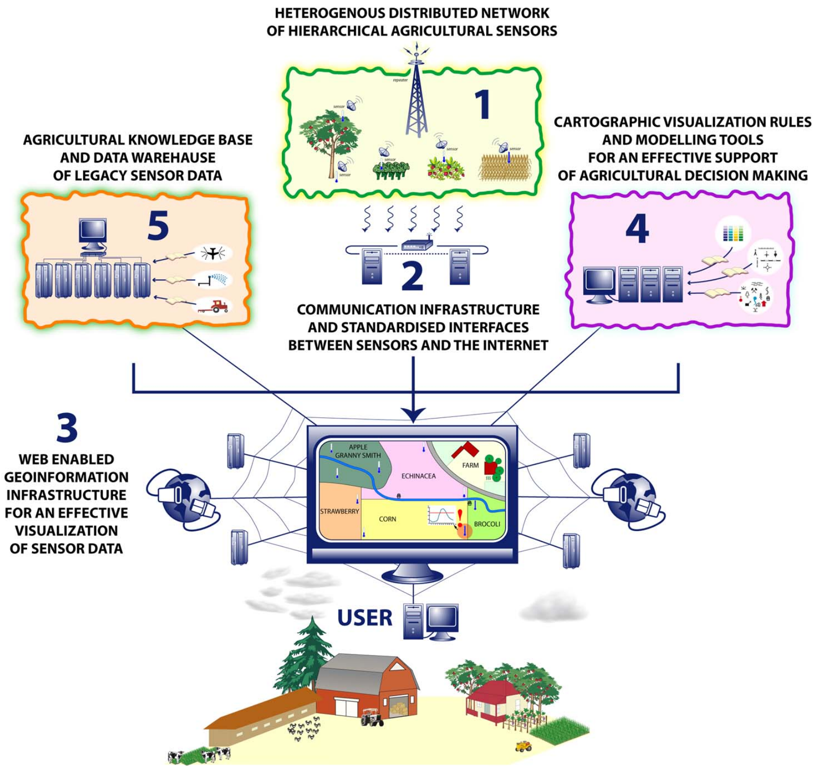



*Figure 2. Overall framework of Agrisensor project with main technological components.*

Currently all the components 1- 5 are partly developed, since the project is in the middle of framework. Following sections describe the reaction of Agrisensor project on aforementioned conceptual principles and overview the state-of-the-art as for January 2011.

## Material and methods

In Agrisensor research project, the agricultural monitoring of key weather and soil parameters using sensors is realized in three complementary levels with different spatial extent and temporal intensity:

A. Mapping of spatial variability of soil and field crops in large fields. Verification was carried out at two different localities in in South Moravia region (Czech Republic): Field "Pachty" (52.5 ha) with chernozem soil type and sandy clay loam texture, and Field "Haj" (37.8 ha; 49°15'N, 17°06'E) with haplic luvisol and silt loam texture. The measurement is done few times in year by ground based survey (soil moisture, temperature, soil compaction), on-the-go methods (electrical conductivity) and remote sensing (aerial multispectral, hyperspectral and thermal imaging).

B. Monitoring of temporal dynamics of weather parameters, which are important for crop growth (soil and air temperature, soil moisture, air humidity) is done by sensor nodes. Low count (5) of sensor nodes is distributed at mentioned localities over each field to cover areas with contrast soil condition. Other important meteorological parameters as precipitation, solar radiation, and wind speed and air flow direction are measured by meteorological station installed at the farm (3 km from the experimental fields). Their changes over field area are considered as constant. Combination of both surveys can model the spatio-temporal changes of observed meteorological characteristics for selected field and crop and can implement it into decision support systems used in site specific crop management.

*A* *B*

*Figure 3. Meteorological station installed at the farm near to "Haj" experimental field (a) and sensor node installed in maize small-plot field at Field Experiment Station of Mendel University in Brno (b).*

C. Monitoring of dynamics of key agrometeorological parameters in different crops species (maize, winter wheat, spring barley, potatoes) and soil tillage practices (traditional, conservation). The measurements are carried out by sensor nodes under controlled conditions in small-plot experiments at the Field Experiment Station of Mendel University in Brno. The obtained results will provide information about the differences in the parameters among individual crops and variants of their crop management practices.

Data flow and maintenance

Hardware of experimental wireless sensor network is composed from two hierarchical components:

- L1 – series of peripheral wireless sensors communicating with ZigBee protocol. Each sensor node includes a master processor and digital input/output with units measuring temperature and humidity for both soil and air. Besides meteorological variables also technological information are measured (signal strength, battery voltage) for better network maintenance.
- L2 – mobile communication unit based on ARM processor and GPRS communication unit (for specific occasion also WiFi communication module is ready to be used). This unit has a twofold function within the WSN – it is equipped with independent meteorological sensors (for data calibration) and also includes radio module for communication with L1 level nodes. L2 node is responsible for data pre-processing.

In many instances this hierarchical approach is also recommended and implemented by other pilot studies within agricultural area (De Filippis *et al.* 2010).

Each sensor of L1 component measures its phenomenon in preconfigured time period which is typically few seconds or minutes. When new value is measured, it is instantly transferred using local wireless network to the L2 component. From the L2 component the value is automatically sent to the remote main server using GPRS connection. Main server then performs insertions of data to the database. The whole process is done automatically and does not take more than a few seconds.

Central sensor data repository is implemented as Java web application and PostgreSQL database with PostGIS extension for its robustness and spatial geometry support. The database schema was designed taking performance and scalability into the account. Couple of function and trigger functions has been implemented in PL/pgSQL to process and sort the data during insertion for improving the performance. The database schema is also designed for the case of moving sensors.

When the measured data stored in the central database is accessible to the user by another Java web service which is based on Sensor Observation Service (SOS). At the moment, the service is not fully SOS 1.0 compliant because we are expecting SOS 2.0 which is nearly finished. However the Java web service follows all major concepts of SOS 1.0. It supports request describing units, i.e. set of sensors with the same location, request describing sensors and observed phenomena, and request for obtaining measure values. Java web service brings the measured data from database to the client side in JSON format (Javascript Object Notation). This format can be then easily used for AJAX based client application and near real time updating of visualization.

Both above described methods of measurements are combined in order to achieve synergy and reveal potential spatio-temporal dependencies (fig.4).

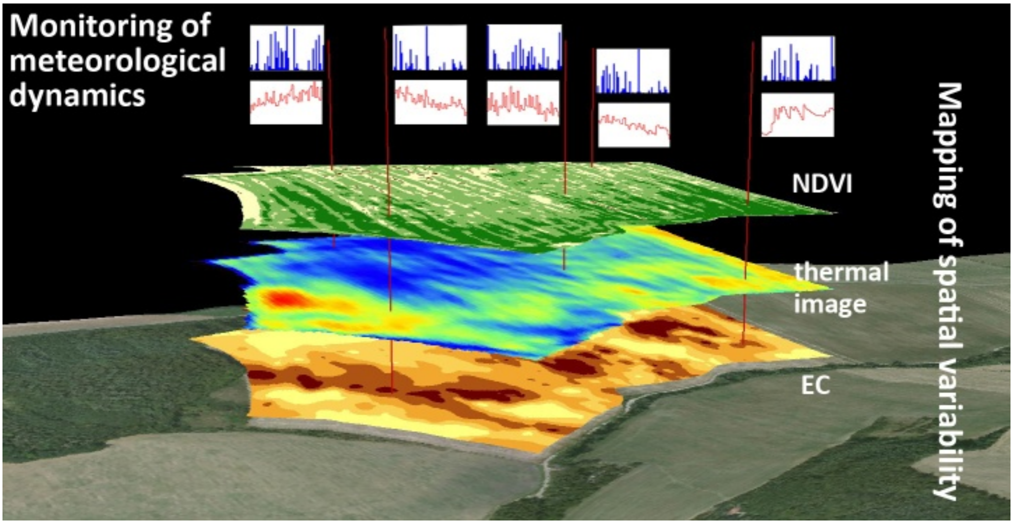



*Figure 4. Combination of spatial variability mapping (in low temporal resolution) and grid monitoring of dynamics of meteorological characteristics (in low spatial resolution).*

Preliminary results

For exploring and visualization of measured sensor data, an interactive web client was developed. Technologically it is based on HTML, CSS, and JavaScript libraries (ExtJS, OpenLayers, and Google Maps API JavaScript), so it can be accessed by most of modern web browsers including Internet Explorer, Firefox, Chrome, Opera, and Safari.

We tried to implement the basic conceptual principles described above. Namely the multiview and tasks (overview, zoom and filter, and details on demand) principles were followed. Client has two main views: map view and chart view (fig.5 and 6). The map enables the hierarchical predefined views from a general overview of all studied area, through the particular agricultural experimental plots to the detail view of sensors location. In the most detail view, units (i.e. poles with set of sensors) are displayed. By clicking on unit symbol, new window with latest measured values appears. In the chart view, an interactive chart is drawn based on user specified sensor and time range. It is also possible to choose all sensors measuring the same phenomenon, so the difference between measurement places can be seen. Below the chart, user can see statistical information about currently displayed dataset. Finally user is able to export or print the chart.

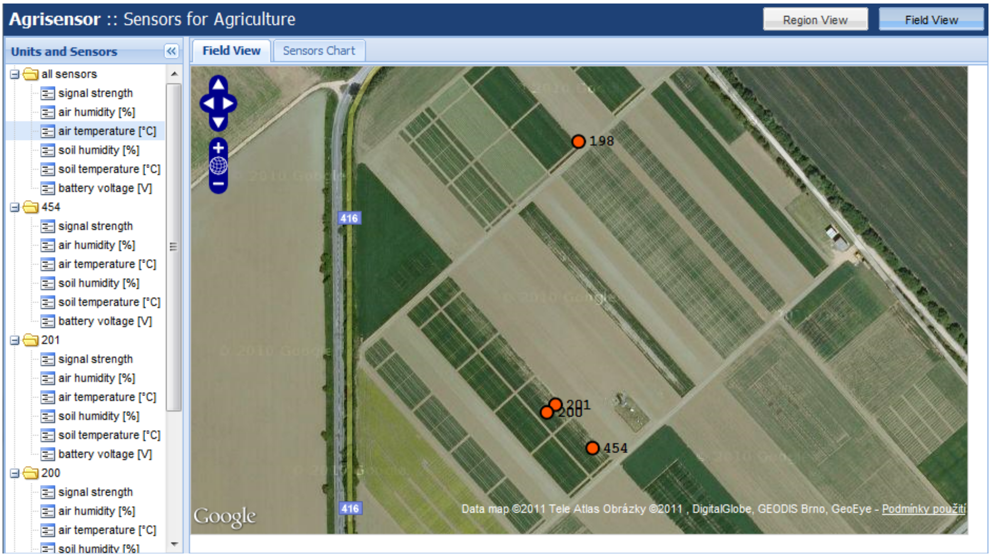



*Figure 5. Map view within the client*

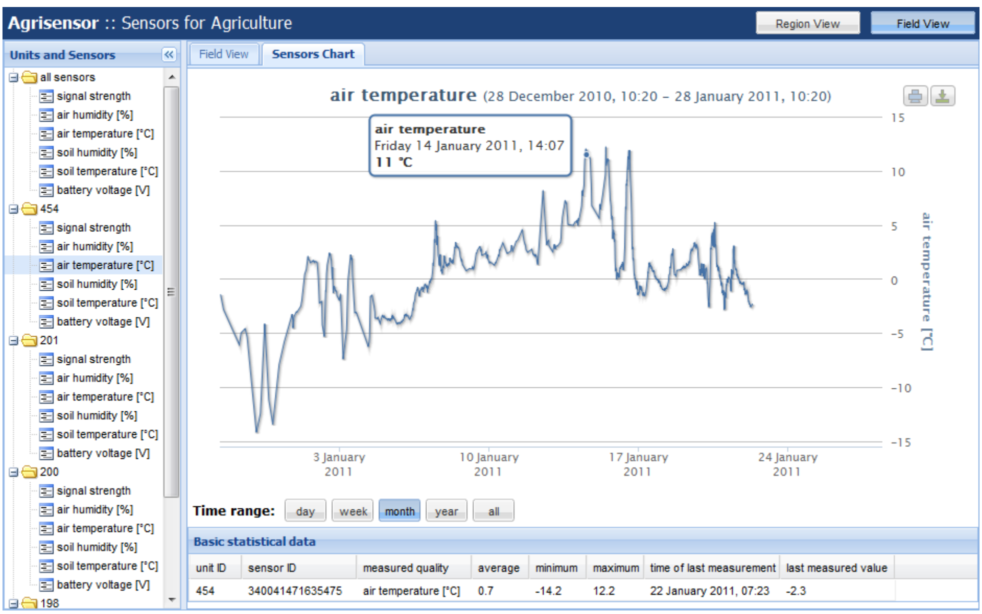



*Figure 6. Chart view within the visualization client*

Catalogue of available geospatial data both local and remote Web Map Services (WMS) was developed for selected experimental fields. This set of geodata sources is intended for the general overview and global views of experimental plots. It also serves as a knowledge base for visual exploration of spatio-temporal dependencies between measured sensor data and environment. Local geographic data were measured and collected for the most detailed level dealing with individual sensors location including the history of particular crops location and their alternative cultivation.

Second part of preliminary results deals with the acquisition of local information about the agronomic spatial variability of experimental plots. At both experimental localities were carried out the ground based mapping and remote sensing of within-field spatial variability of soil using different methods. Information about soil conditions to the depth of 0.75 m were obtained by on-the-go measurement of soil electrical conductivity, which enables rapid and non-invasive identifications in soil substrate. Methodology and results of this survey were described by Lukas *et al.* (2009). As remote sensing, an aerial survey was carried out in visible, NIR and thermal part of electromagnetic spectrum. Results of both indirect methods (EC and remote sensing) and their spatial patterns were visualized (fig. 7) and compared to the soil properties (see Table 1.).

| Sensor measurement | Agrochemical soil properties | | | | | |
|---|---|---|---|---|---|---|
| | pH | P | K | Mg | Ca | humus |
| Electrical conductivity | 0.565 | -0.575 | -0.500 | 0.577 | 0.268 | 0.469 |
| Visible bands (first PCA component) | -0.371 | 0.560 | 0.501 | -0.225 | -0.007 | -0.428 |
| NIR band | -0.379 | 0.638 | 0.549 | -0.326 | -0.015 | -0.547 |
| Termography | 0.424 | -0.534 | -0.569 | 0.363 | 0.079 | 0.276 |

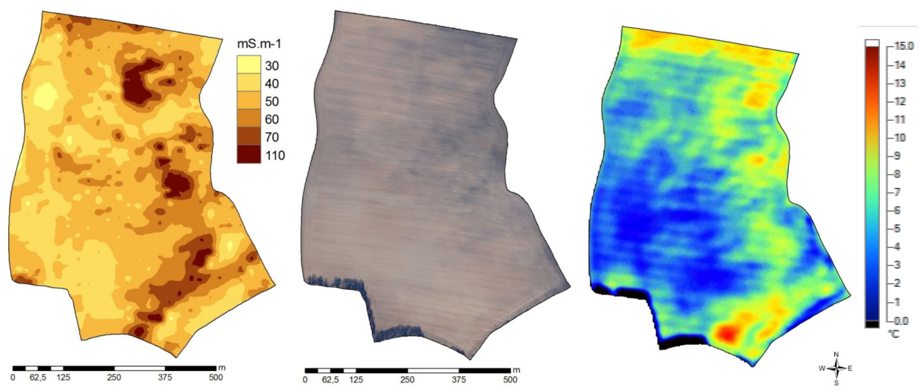



*Table 1. Correlation coefficients between agrochemical soil properties and indirect methods calculated from ground and aerial survey of 52ha field.*

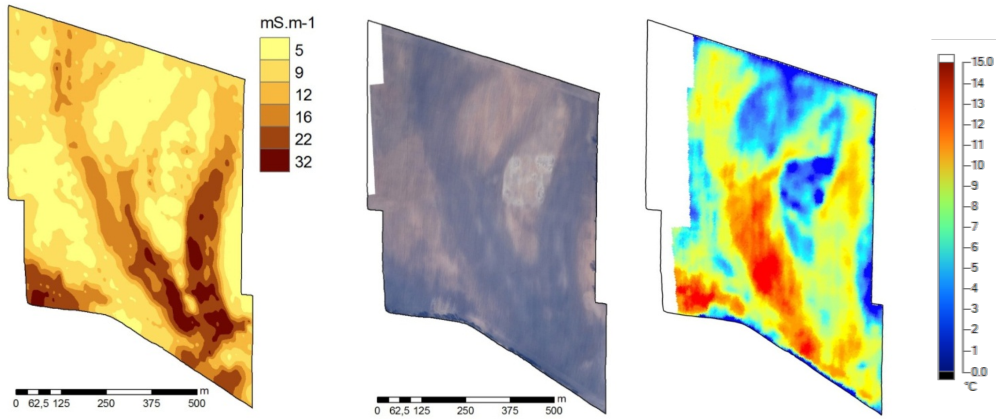



*Figure 7. Spatial patterns on the 38 ha (upper) and 52 ha (lower) field obtained from soil EC measurement (left), aerial imaging in visible spectrum (middle) and image from aerial thermography (right).*

**Conclusion and future development**

Wireless sensors networks geovisualization is here presented on two different levels of abstraction – the higher level deals with the conceptual approach and presents the theoretical background of wireless sensor networks in agriculture and sensor geovisualization issues. The lower is focused on application level and describes selected problems prototyped within the framework of Agrisensor project.

Zerger *et al.* (2010) stated that the applied operational use of sensor networks remains in its infancy and pointed following areas to be important for future directions. Project Agrisensor reacts on particular statements in the following way (*statement* – reaction):

- *A greater focus on validating results from sensor technologies to determine how effectively they replicate manual monitoring methods* – local knowledge database is developed within the project in order to enable the comparison of sensor node measures with calibrated meteorological sensor and also with historical measurements. Another important issue solved by project is the proximal soil sensing used with combination of WSN. In order to avoid sensors malfunction and prevent data loss also battery voltage and signal strength is monitored.
- *A shift from single-node deployments to multi-node deployments and wireless sensor networks* – project is prototyping a hierarchical network with 2 levels of nodes.
- *Coupling multiple domain sensors to examine causal relationships* – we are coupling meteorological and soil sensors together within the project WSN enabling to examine responses of soil variables (temperature and humidity) to atmospheric changes (rainfall, heat). There are ongoing talks with the Czech hydro meteorological institute

about the possibility to incorporate selected climatic stations within the experimental network.

- *In situ data processing and analysis* – all measurements are pre-processed on the second level of WSN and can be stored there in case of communication infrastructure failure.

Further development of visualization client is planned. The main focus is on so called context aware or adaptive visualization (Friedmannova *et al.* 2006, Kubicek and Kozel 2010, Kozel and Stampach 2010). A context influences map compositions and chart appearance for specific agriculture situations (e.g. maize harvesting or wheat sowing).

**Acknowledgements**

The Agrisensor project is supported by the grant no. 205/09/1437 of the Czech Science Foundation.

**References:**

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y.Cayirci, E. 2002. Wireless sensor networks: A survey. Computer Networks. 38 4), 393-422.

Andrienko, N., Andrienko, G. 2006): Exploratory Analysis of Spatial and Temporal Data. Springer – Verlag, Berlin Heidelberg, 703 p.

Camilli, A., Cugnasca, C. E., Saraiva, A. M., Hirakawa, A. R. and Corrêa, P. L. P. 2007. From wireless sensors to field mapping: Anatomy of an application for precision agriculture. Computers and Electronics in Agriculture. 58 1), 25-36.

Cook. D. 2007. Making Sense of Sensor Data. IEEE PERVASIVE COMPUTING, p. 105–108.

De Filippis, T., Rocchi, L., Fiorillo, E., Genesio, L. 2010. A webgis application for precision viticulture: from research to operative practices. In: Proceedings of 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services, WebMGS, Milano, 6 p.

Friedmannova, L., Konecny, M. and Stanek, K. 2006. An Adaptive Cartographic Visualization for Support of the Crisis Management. Autocarto 2006, Vancouver, Canada.

Grinstein, G., Trutschl, M., Cvek U. 2001). High-dimensional visualizations. In Proceedings of the 7th Data Mining Conference-KDD, 2001.

Hart, J., Martinez. K. 2006. Environmental Sensor Networks: A revolution in the earth system science? Earth-Science Reviews 78 2006) 177–191.

Kozel, J., Stampach, R. 2010. Practical Experience with a Contextual Map Service. In Geographic Information and Cartography for Risk and Crisis Management - Towards Better Solutions. 1st Edition. Heidelberg Dordrecht London New York : Springer -Verlag, 2010. Lecture Notes in Geoinformation and Cartography, ISBN 978 -3 -642 -03441 -1. p. 305 -316.

Kubicek, P., Charvat,K., Lukas, V. 2009. Cartographic visualization of agricultural sensor. In Sborník ze společné mezinárodní konference ISZL/CEE-SDI/ ICT for Rural 2009. Praha. 1. vyd. Praha : Česká zemědělská univerzita,. 9 p. ISBN 978-80-213-1932-5.

Kubicek, P., Kozel, J. 2010. Cartographic techniques for adaptive emergency mapping. In Kremers, H., Susini, A. eds.. Risk Models and Applications. Lecture Notes in Information Science Berlin : CODATA-Germany, p. 49 – 62.

Lee, W. S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D. and Li, C. 2010. Sensing technologies for precision specialty crop production. Computers and Electronics in Agriculture. 74 1), 2-33.

Lukas, V., Neudert, L. and Kren, J. 2009. Mapping of soil conditions in precision agriculture. Acta Agrophysica 13(2), 393-405.

Oerke, E.-C. and Gerhards, R. 2010. Precision crop protection - the challenge and use of heterogeneity. Springer, Dordrecht; Heidelberg [u.a.], 441 pp.

Open Geospatial Consortium (2011): SWE Common Data Model Encoding Standard 2.0. OGC Encoding Standard 08-094r1, http://www.opengeospatial.org/standards/swecommon.

Pierce, F. J. and Elliott, T. V. 2008. Regional and on-farm wireless sensor networks for agricultural systems in Eastern Washington. Computers and Electronics in Agriculture. 61 1), 32-43.

Pierce, F. J., Nowak, P. and Donald, L. S. 1999. Aspects of Precision Agriculture. Advances in Agronomy. 67, 1-85.

Plaisant, C. 2004. The challenge of information visualization evaluation. In Proceedings of the working conference on Advanced visual interfaces, p. 109–116. ACM New York, NY, USA.

Reichardt, M. 2005. Sensor Web Enablement: An OGC White Paper, Open Geospatial Consortium OCG), Inc.,.

Richter, R. 2009. Visualizing sensor data. Media Informatics Advanced Seminar on Information Visulaization.

Shneiderman, B. 1996. The eyes have it: a task by data type taxonomy for information visualizations. In. Proceedings IEEE Symposium on Visual Languages, p. 336–343,.

Simonis, I., Wytzisk , A., Raape , U., Schulze, T. 2003. Interoperability of Simulation and Geoinformation Standards. Proceedings of: Spring-SIW, Orlando, Florida.

Simonis, I., Ed.) 2008. Sensor Web Enablement Architecture, OGC Engineering report 06-021r2, http://portal.opengeospatial.org/files/ ?artifact_id=27775&version=1.

Wang, N., Zhang, N., Wang, M. 2006. Wireless sensor in agriculture and food industry – Recent development and future perspectives. Computers and Electronics in Agriculture, 50, Elsevier, p.1 – 14.

White, R. 1987. A Sensor Classification Scheme. Ultrasonic, Ferroelectrics and Frequency Control, IEEE Transactions on, 34 2), p. 124 – 126.

Zerger, A., Viscarra Rossel, R. A., Swain, D. L., Wark, T., Handcock, R. N., Doerr, V. A. J., Bishop-Hurley, G. J., Doerr, E. D., Gibbons, P. G. and Lobsey, C. 2010. Environmental sensor networks for vegetation, animal and soil sciences. International Journal of Applied Earth Observation and Geoinformation. 12 5), 303-316.

# Server-side solution for sensor data

Jan Ježek, Michal Kepka,
*University of West Bohemia in Pilsen*
E-mail: jezekjan@kma.zcu.cz

## Introduction

To better understand to basic concept of the solution we should at first describe in more details what sort of sensor network we have to deal with. In the sensor network sensors are constantly measuring the data and transmitting them into nodes gateway.

These gateways have usually Internet connection (e.g. GPRS) and send the data to the Internet cloud. The data are received by a web based service that inserts them into the database. This leads to approximately hundreds of millions of observations in one day. There are approximately thousands of insertions in one second. The basic elements of sensor network are:

- Mobile units – e.g. gateway with Internet connection that receives data by sensors and transmits it on main server.
- Sensor – simple devices that are measuring particular phenomenon. These devices send the observation to gateway using different method (bluetooth, RFID, cable). The most important property of sensor is low energy consummation and no need of maintenance.
- Main server – the server or server cluster that receives, stores, analyses and publishes the data.

Mobile unit together with sensors can be statically put on some place (e.g. field) or it can be attached to moving object (car). The propose of SensLog installed on main server is to provide various kinds of interfaces that enable final client application to easily access the data. SensLog also provide capabilities to define alerts, authentication, overall network maintenance and configuration.

## Database

We have designed particular data model shown at diagram on figure 1. for the propose of main observation storage. Our implementation is utilizing relational database management system PostgreSQL with spatial database extension PostGIS. Because we have to deal with big amount of data that are growing continuously, the database contains also triggers that helps to prepare data during the insertion for particular queries. This data model is suitable not only for static sensors but also for mobile sensors too. System of groups made a hierarchy which allows users to get data only from permitted units. The units and groups are in many-to-many relation.

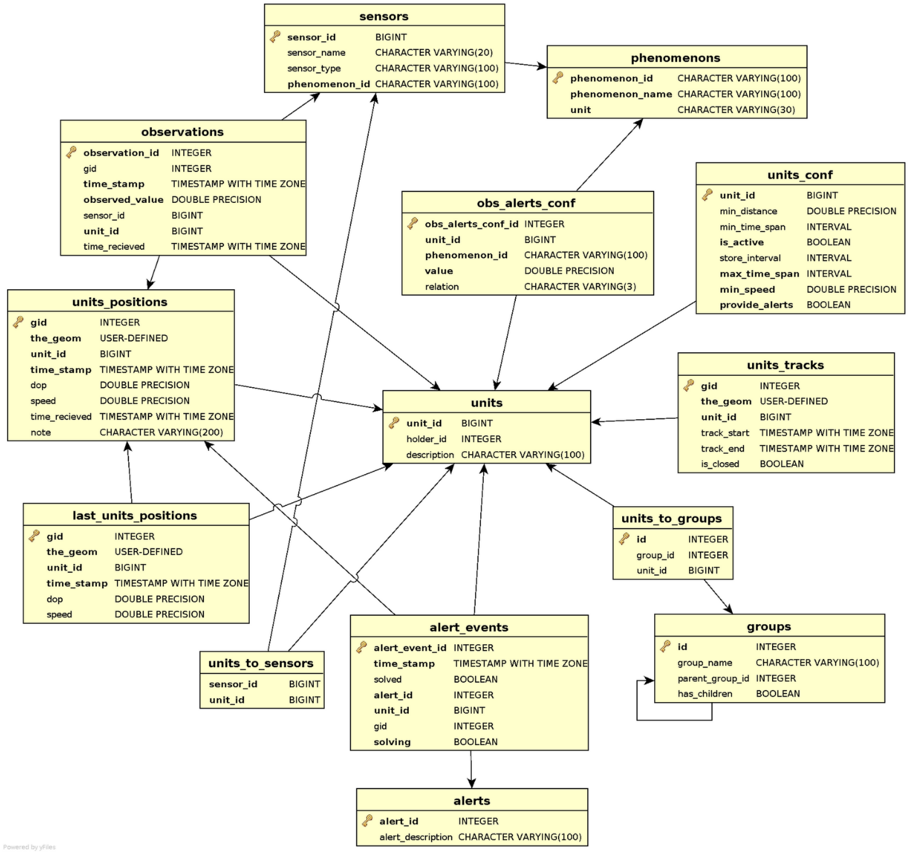



*Figure 1. Data model for sensor data*

As shown on figure 2.1 the most important tables in model are:

- units

List of units located in sensor network. Unit is hardware component that conects sensors to the internet.

- sensors

List of sensor types located in sensor network, sensors are attached to units in relation many-to-many, decomposition table is units_to_sensors.

- phenomenons

List of measured phenomena with units of measurement.

- observations

List of measured values of certain sensor and unit for certain timestamps.

- units_positions

List of unit positions for certain timestamps.

- alerts

List of potential exceptional states in sensor network.

- alert_events

List of exceptional situations for certain timestamps.

- groups

List of user groups which can access sensor network in own competent part. Units and groups are in relation many-to-many, decomposition table is units_to_groups.

As every software system, a DBMS for sensor data operates in a faulty computing environment. A failure can corrupt the respective database unless special measures are taken to prevent this.

We have to deal with scalability and high availability of data. For this we are using proper replication and clustering mechanism. We are using pgpool-II for this propose. Pgpool-II is a middleware that works between PostgreSQL servers and a PostgreSQL database client. It provides the following features:

- Connection Pooling

Pgpool-II saves connections to the PostgreSQL servers, and reuses them whenever a new connection with the same properties (i.e. username, database, protocol version) comes in. It reduces connection overhead, and improves system's overall throughput.

- Replication

Pgpool-II can manage multiple PostgreSQL servers. Using the replication function enables creating a real-time backup on 2 or more physical disks, so that the service can continue without stopping servers in case of a disk failure.

- Load Balance

If a database is replicated, executing a SELECT query on any server will return the same result. Pgpool-II takes an advantage of the replication feature to reduce the load on each PostgreSQL server by distributing SELECT queries among multiple servers, improving system's overall throughput. At best, performance improves proportionally to the number of PostgreSQL servers. Load balance works best in a situation where there are a lot of users executing many queries at the same time.

- Limiting Exceeding Connections

There is a limit on the maximum number of concurrent connections with PostgreSQL, and connections are rejected after this many connections. Setting the maximum number of connections, however, increases resource consumption and affect system performance. Pgpool-II also has a limit on the maximum number of connections, but extra connections will be queued instead of returning an error immediately.

- Parallel Query

Using the parallel query function, data can be divided among the multiple servers, so that a query can be executed on all the servers concurrently to reduce the overall execution time. Parallel query works the best when searching large-scale data.

**Saving data**

Hypertext Transfer Protocol (HTTP) API for inserting measured data was made in Java programming-language. Therefore common users will not need to deal with the database itself. There are HTTP GET methods for inserting observations, positions and alert events. These services are based on simple but proprietary protocol. Another possibility of insertion is to utilise more complicated SOS transactional profile that is based on HTTP Post and XML based communication.

**Publishing data - proprietary interface**

Main interface for client side application is designed as a set of web services written in Java that provides HTTP GET interface to retrieve data from database. These services provide users with particular data in JSON format. The services are using authentication so user can get just data that they have privileges to see. After authentication you can start to query the database using HTTP GET requests.

Usual procedure is:

1. Ask for units:

   *Service?Operation=GetUnits*

   - This request gives client basic data about available units, their last positions and sensors that are attached to units. There is also information about sensor measurement like first and last time of observation.

2. Ask for observed values:

   *Service?Operation=GetObservations*
   *&unit_id=101*
   *&sensor_id=360041*
   *&from=2011-04-05+00:00*
   *&to=2011-04-06+00:00*

   - This request gives client observations from defined unit and sensor and for specified time period.

3. Ask for last positions:

   *Service?Operation=GetLastPositions*
   *&user=test*

   - This request gives client last positions of available units.

Output in JSON format is useful for programming advanced graphical user interface based on AJAX technology. Example of some JSON output is shown on figure 2.

[{"gid":0,"time":"2011-11-02 20:56:35+01","value":23.9375},
{"gid":0,"time":"2011-11-02 21:06:02+01","value":24.1875},
{"gid":0,"time":"2011-11-02 21:16:02+01","value":24.25},
{"gid":0,"time":"2011-11-02 21:26:02+01","value":24.1875},
{"gid":0,"time":"2011-11-02 21:36:02+01","value":24.1875},
{"gid":0,"time":"2011-11-02 21:46:02+01","value":24.1875},
{"gid":0,"time":"2011-11-02 21:56:02+01","value":24.125},
{"gid":0,"time":"2011-11-02 22:06:02+01","value":24.125},
{"gid":0,"time":"2011-11-02 22:16:02+01","value":24},
{"gid":0,"time":"2011-11-02 22:20:00+01","value":24.125},
{"gid":0,"time":"2011-11-02 22:29:37+01","value":24.9375},
{"gid":0,"time":"2011-11-02 22:39:37+01","value":25},
{"gid":0,"time":"2011-11-02 22:49:37+01","value":24.9375},
{"gid":0,"time":"2011-11-02 22:59:37+01","value":24.9375},
{"gid":0,"time":"2011-11-02 23:09:37+01","value":25},
{"gid":0,"time":"2011-11-03 07:18:53+01","value":26.0625},
{"gid":0,"time":"2011-11-08 19:03:42+01","value":24.5}]

*Figure 2. Output in JSON format*

**Publishing data - standardized interface**

The interoperability is one of the most important thing today. Therefore, it is crucial to enable the access of sensor for another applications and services over networks (as the Internet) and with same technologies and protocols which enable the Web. The idea of Sensor Webs was established just for these cases. OGC's Sensor Web Enablement (SWE) activities have established the interfaces and protocols that will enable Sensor Webs. These activities have defined, prototyped and tested several foundational components needed for a Sensor Web [SOS1].

Four of the components are the most relevant for this paper:

- Observations & Measurements (O&M) – The general models and XML encodings for sensor observations and measurements.
- Sensor Model Language (SensorML) – The general models and XML schema for describing sensors and processes associated with measurement.
- Transducer Markup Language (TML) – General characterizations of transducers, their data, the phenomenon, transporting the data, and any and all support data (metadata) necessary for later processing and understanding of the transducer data.
- Sensor Observation Service (SOS) – The service provides an API for managing deployed sensors and retrieving sensor data (observations).

The goal of SOS [SOS1] is to provide access to observations from sensors and sensor systems in a standard way that is consistent for all sensor systems including remote, in-situ, fixed and mobile sensors. SOS leverages the O&M specification for modeling observations and the TML and SensorML specifications for modeling sensors and sensor systems. SOS is primarily designed to provide access to observations. A model for O&M is formalized in an OGC Best

Practice paper. But our prototype of SOS was implemented above our data model (see Fig. 2.1) used in SensLog.

The concept of SOS operations can be divided into two sections. The first section is oriented on a consumer of sensor observations; the second is focused on a producer of observations. Our implementation [Kep11] is concentrated on publishing observations for consumers.

Every SOS has to contain [SOS1] three mandatory (core) operations:

- GetCapabilities – provides the means to access SOS service metadata.
- GetObservation – provides access to sensor observations and measurement data, a spatio-temporal query filtered by phenomena can be used.
- DescribeSensor – retrieves detailed information about the sensors and processes generating those measurements.

Some non-mandatory operations have been also defined. There are two operations to support transactions – RegisterSensor and InsertObservation. There are six enhanced operations – GetResult, GetFeatureOfInterest, GetFeatureOfInterestTime, DescribeFeatureOfInterest, DescribeObservationType, and DescribeResultModel.

SOS communicates through requests and responses in form of XML encoded documents sent by HTTP POST method. The forms of requests and responses are provided in W3C XML Schema (XSD) language. These schemes are part of SOS Implementation standard. Because SOS is one of the SWE technologies builds the Sensor Webs, uses partially components from other SWE standards. On figure 3. is shown the SOS dependency on other OGC standards, made by [Tam10a].

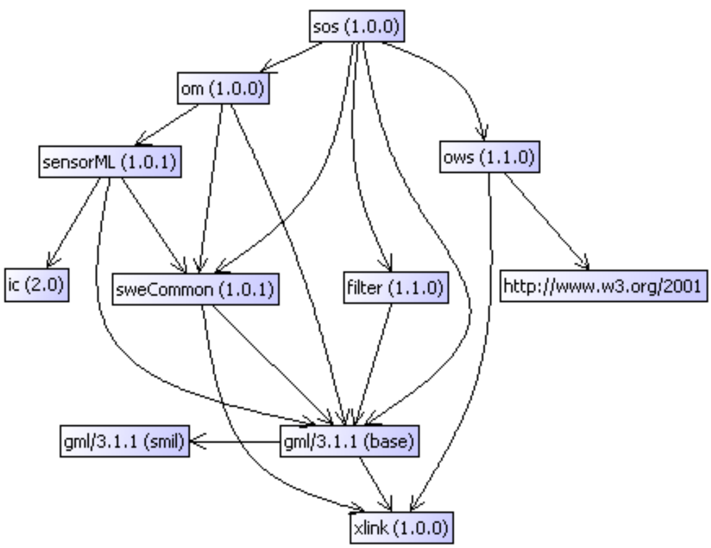



*Figure 3. SOS dependencies*

As figure 3. shows SOS uses lot of other OGC standards. If we remind SOS want to be complex service for all possible observations, it indicates that SOS schemes will be very complicated.

An implementation of SOS is a server-side application which accepts request, collects data from database and compiles responses. The response sent by SOS can be a standard document with requested data or an exception when SOS encounters an error while performing the operation.

Because schemes are part of implementation standard [SOS1], there are at least two ways how to create an SOS implementation. Firstly, it is possible to transform XML documents by Extensible Stylesheet Language Transformations (XSLT). Secondly, it is possible to use technology of data binding. Data binding provides an automated translation between a XML document and programming-language objects.

**Implemented SOS service**

Implemented SOS service [Kep11] is based on the technology of data binding. It means that classes and interfaces are derived from XML schemes by binding compiler. Java Architecture for XML Binding (JAXB) was used for our purposes [Lau08]. They are used compiled schemes from OGC Schemas and Tools Project at this time. OGC Schemas project provides JAXB 2.x bindings for XML Schemas defined by the OGC. These compiled schemes are developed by Aleksei Valikov under 3-clause BSD license [Val11]. Binding a schema means generating a set of Java classes that represents the schema. All JAXB implementations provide a tool called a binding compiler to bind a schema (XJC) [Lau08]. After binding we do not have to deal with XML documents ourselves, we deal with programming-language objects, in this case with Java classes. Reading a XML document in JAXB terminology is called unmarshalling. Unmarshalling an XML document means creating a tree of content objects that represents the content and organization of the document. Writing a XML document is called marshalling, it creates a XML document from a content tree.

Implementation is a server-side application that includes core operations of SOS with mandatory parameters at this time.

Implemented operations and their parameters are:

- *GetCapabilities*
  - Request – GetCapabilities document with mandatory parameters *service* and *request*
  - Response – Capabilities document with elements: *ServiceIdentification*, *ServiceProvider*, *OperationsMetadata*, *Filter_Capabilities* and *Contents*.
- *DescribeSensor*
  - Request – DescribeSensor document with mandatory parameters *service*, *procedure*, *outputFormat*, *version*
  - Response – SensorML document with elements: *identification*, *capabilities*, *outputs* and *positions*

- *GetObservations*
  - Request – GetObservation document with mandatory parameters and supported values:
    - service: SOS
    - version: 1.0.0
    - srsName: urn:ogc:def:crs:EPSG:4326
    - offering: according to values in Capabilites document
    - eventTime: TM_During
    - procedure: according to values in Capabilites document
    - observedProperty: according to values in Capabilites document
    - responseFormat: text/xml; subtype="om/1.0.0"
    - resultModel: om:Observation
    - responseMode: inline
      - Response – ObservationCollection document with one element Observation with elements: *samplingTime*, *procedure*, *observedProperty*, *result*

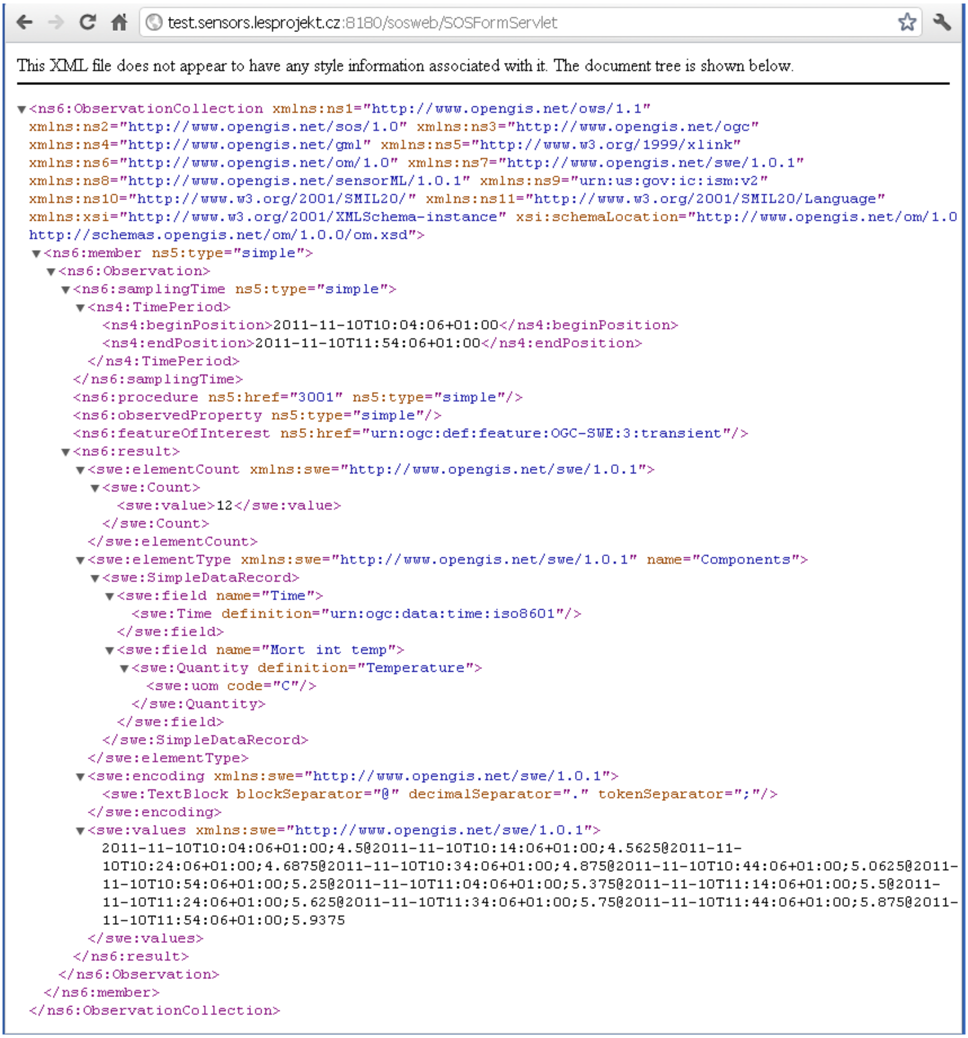
There are are few features that are not yet implemented like scalar or spatial filtering of observations. It might be worth to mention what term is in our data model (see Fig. 1.) equivalent to term in SOS standard (see Table 1).

| **SOS standard** | **Data model** |
|---|---|
| Observation offering | Group |
| Procedure | Unit |
| Observed property | Phenomenon |
| FeatureOfInterest | not in use yet |

*Table 1. Equivalent terms*

There must be set a procedure and an observed property in GetObservations request for getting observations from particular sensor. Affiliation sensors with units can be found in DescribeSensor response.

There are examples of SOS GetObservations responses on next two figures. Raw ObservationCollection document is shown on Figure 4. Figure 5. shows screenshot of graphical SOS client from HSLayers which draws graph whose a few values are from document on figure 4.


test.sensors.lesprojekt.cz:8180/sosweb/SOSFormServlet

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
▼<ns6:ObservationCollection xmlns:ns1="http://www.opengis.net/ows/1.1"
 xmlns:ns2="http://www.opengis.net/sos/1.0" xmlns:ns3="http://www.opengis.net/ogc"
 xmlns:ns4="http://www.opengis.net/gml" xmlns:ns5="http://www.w3.org/1999/xlink"
 xmlns:ns6="http://www.opengis.net/om/1.0" xmlns:ns7="http://www.opengis.net/swe/1.0.1"
 xmlns:ns8="http://www.opengis.net/sensorML/1.0.1" xmlns:ns9="urn:us:gov:ic:ism:v2"
 xmlns:ns10="http://www.w3.org/2001/SMIL20/" xmlns:ns11="http://www.w3.org/2001/SMIL20/Language"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.opengis.net/om/1.0
 http://schemas.opengis.net/om/1.0.0/om.xsd">
  ▼<ns6:member ns5:type="simple">
    ▼<ns6:Observation>
      ▼<ns6:samplingTime ns5:type="simple">
        ▼<ns4:TimePeriod>
           <ns4:beginPosition>2011-11-10T10:04:06+01:00</ns4:beginPosition>
           <ns4:endPosition>2011-11-10T11:54:06+01:00</ns4:endPosition>
         </ns4:TimePeriod>
       </ns6:samplingTime>
       <ns6:procedure ns5:href="3001" ns5:type="simple"/>
       <ns6:observedProperty ns5:type="simple"/>
       <ns6:featureOfInterest ns5:href="urn:ogc:def:feature:OGC-SWE:3:transient"/>
      ▼<ns6:result>
        ▼<swe:elementCount xmlns:swe="http://www.opengis.net/swe/1.0.1">
          ▼<swe:Count>
             <swe:value>12</swe:value>
           </swe:Count>
         </swe:elementCount>
        ▼<swe:elementType xmlns:swe="http://www.opengis.net/swe/1.0.1" name="Components">
          ▼<swe:SimpleDataRecord>
            ▼<swe:field name="Time">
               <swe:Time definition="urn:ogc:data:time:iso8601"/>
             </swe:field>
            ▼<swe:field name="Mort int temp">
              ▼<swe:Quantity definition="Temperature">
                 <swe:uom code="C"/>
               </swe:Quantity>
             </swe:field>
           </swe:SimpleDataRecord>
         </swe:elementType>
        ▼<swe:encoding xmlns:swe="http://www.opengis.net/swe/1.0.1">
           <swe:TextBlock blockSeparator="@" decimalSeparator="." tokenSeparator=";"/>
         </swe:encoding>
        ▼<swe:values xmlns:swe="http://www.opengis.net/swe/1.0.1">
           2011-11-10T10:04:06+01:00;4.5@2011-11-10T10:14:06+01:00;4.5625@2011-11-
           10T10:24:06+01:00;4.6875@2011-11-10T10:34:06+01:00;4.875@2011-11-10T10:44:06+01:00;5.0625@2011-
           11-10T10:54:06+01:00;5.25@2011-11-10T11:04:06+01:00;5.375@2011-11-10T11:14:06+01:00;5.5@2011-
           11-10T11:24:06+01:00;5.625@2011-11-10T11:34:06+01:00;5.75@2011-11-10T11:44:06+01:00;5.875@2011-
           11-10T11:54:06+01:00;5.9375
         </swe:values>
       </ns6:result>
     </ns6:Observation>
   </ns6:member>
 </ns6:ObservationCollection>
```


*Figure 4. ObservationCollection document*

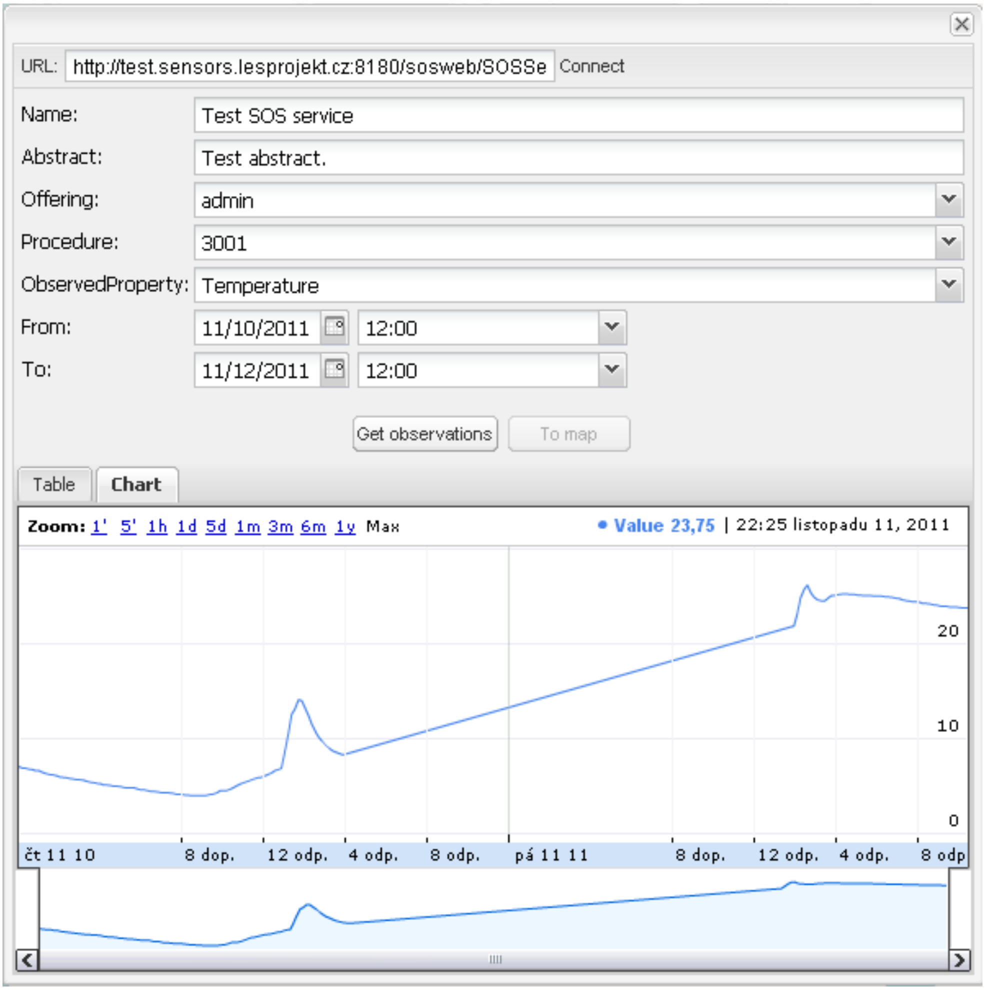



*Figure 5. HSLayers SOS client*

Figure 5. shows that implemented SOS interface can be connected to external graphical SOS client successfully, e.g. HSLayers SOS client.

**Conclusion**

The aim of this paper was to present created data model for sensor data. We also introduce server-side application SensLog that is capable to provide access to this data by using standardized interface called SOS.

However this data model is different from standard OGC model for observation, it brings more functionality for our purpose (e. g. in alerts or groups). Created server-side Java application SensLog can be used as interface for another clients based on AJAX technology by using simplified API based on JSON. This proprietary service based on JSON can fully utilize advantages of created model.

A SOS interface implementation was made for reaching the interoperability. Core profile of methods and their basic functionality has been implemented for now. Applicability of created SOS was tested successfully on HSLayers SOS client.

## References


[Lau08] LAUN, Wolfgang. A JAXB Tutorial [online]. 2008, 14. 4. 2011 [cit. 2011-04-30]. WWW: <http://jaxb.java.net/tutorial/>.

[SOS1] OGC 06-009r6. Sensor Observation Service: OpenGIS® Implementation Standard. Wayland (Mass.): Open Geospatial Consortium, Inc., 26. 10. 2007. xiv, 90 s. WWW: <http://www.opengeospatial.org/standards/sos>.

[Tam10a] TAMAYO, Alain; GRANELL, Carlos; HUERTA, Joaquín. OGC Schemas Browser: Visualizing OWS' XML Schemas. In Geospatial Thinking: Proceedings of AGILE 2010 [online]. [Guimarães]: [s.n.], 2010 [cit. 2011-05-01]. WWW: <http://plone.itc.nl/agile_old/Conference /2010-guimaraes/ShortPapers_PDF/82_DOC.pdf>. ISBN 978-989-20-1953-6.

[Kep11] KEPKA, Michal. *Implementation of Sensor Observation Service standard* [online]. Pilsen: University of West Bohemia in Pilsen, 2011. 77 p. Diploma thesis. WWW: <https://portal.zcu.cz/wps/PA_StagPortletsJSR168/KvalifPraceDownloadServlet?typ=1&adipidno=42585>.

[Val11] VALIKOV, Aleksei. *OGC Schemas and Tools Project*. 2006 - 2011. Karlsruhe. WWW: http://confluence.highsource.org/display/OGCS/Home

# Advanced monitoring systems for processing sensor network data

Filip Prochazka, Michal Oskera, Martin Svehla
*Mycroft Mind, a.s.*
E-mail: pro@mycroftmind.com

## Introduction

Maturity of sensor networks is still growing and sensor networks becoming quite usual as a base for environment condition monitoring. A lot of effort is invested to perpetual collection of the data from monitored environment. Data collection process has to be cheap, robust and reliable. It is important task, but to collect the data is only first step. Good solution has to offer more to users than just access to measured values. Good solution has to offer permanent pro-active monitoring.

Monitoring systems maturity classification

Let's introduce model for monitoring system maturity classification.

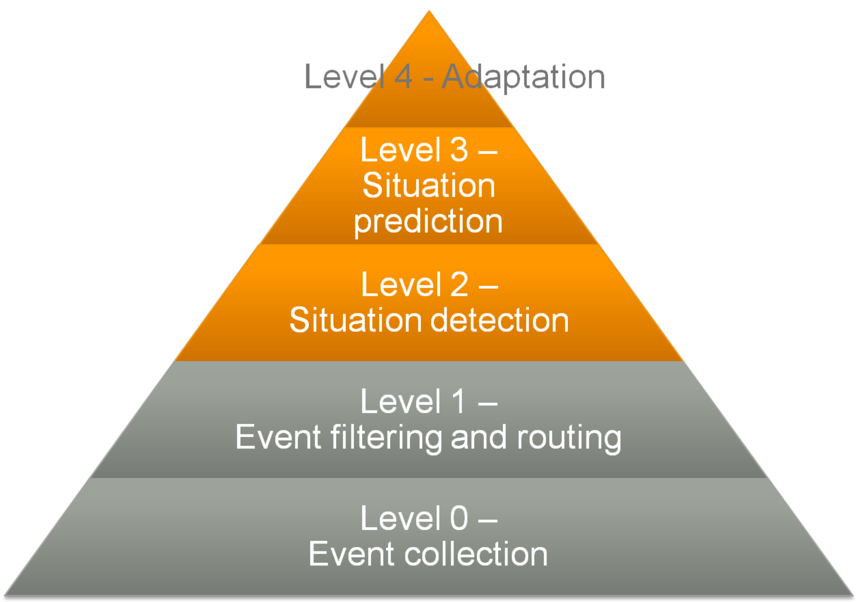



*Figure 1. Monitoring solution maturity pyramid*

Level 0: Event collection

Monitoring system at level 0 is just event or measurement collector. It is up to user to browse collected data systematically and to understand, what low level events or measurements mean and to decide, if monitored situation needs some reaction. It doesn't allow proactive monitoring.

Level 1: Event filtering and routing

Lot of todays monitoring systems are at level 1. They allow to filter collected measurements according to simple conditions (e.g. if temperature on sensor X is greater than 20 degrees) and to generate low level alerts. Alerts can be delivered to specific users according to users roles and type of alerts. Monitoring systems on that level have one key problems – it is very hard to tune the systems. It oscillates between two extrems: system is generating to much low level alerts which are hard to interpret - what does it mean if the temperature on sensor X is greater than 20 degrees? To decide which users have to be informed about individual low-level alerts is not easy task. Typical case is that users are overhelmed with low-level alerts. Typical user reaction on that kind of system is to ignore all alerts, because more than 90% are useles.

Level 2: Situation detection

To move monitoring system from level 1 to level 2 is crucial improvement. System at level 2 is able to detect complex event – also called situation. It is processing low level data and alerts and recognize event patterns. Event patterns are typicaly detecting sequences of events and correlate several event sources (e.g. sensors or low-lovel monitoring systems). It allows to recognize situation meaningful for users which are also much more easy to distribute to appropriate addresates. Every detected situation can be drill down to explain what low-level data causes situation detection.

Level 3: Situation prediction

Next maturity level allows to predict occurence of situations. Prediction is typically based on event pattern detection and ability to create partially instantiated patterns. We can define metric how much pattern is instantiated. If pattern instantion reach some threshold, we can alert user that given situation is nearly to come.

Level 4: Adaptation

Monitoring system has to be adapted to monitored environment as well as to users. This adaptation can be made during monitoring system implementation or there it can be supported by some advanced techniques from domain of machine learning. Adaptation to specific environment is typically done by computing profiles of monitored objects typical behavior. Advanced systems are able to update profiles perpetually. It allows to smoothly adapt to changing conditions in monitored environment (e.g. adapt from winter conditions in winter to spring conditions). Adaptation to specific users is mainly about setting relevance of alerts for given user. It helps to eliminate false alarms and harmonize alerting with users perspective.

We can summarize requirements on technology which can allow us to develop mature monitoring solution

- ability to process event based data (measuremets, logs, alerts, ...) in real-time
- ability to detect event patterns
- ability to adapt the system to actual environment

## Complex Event Processing Approach

Complex Event Processing (CEP) is technology designed for processing huge event streams in real-time [1]. It offers lot of useful techniques to fulfil mentioned requirements.

Main principle is illustrated in following figure:

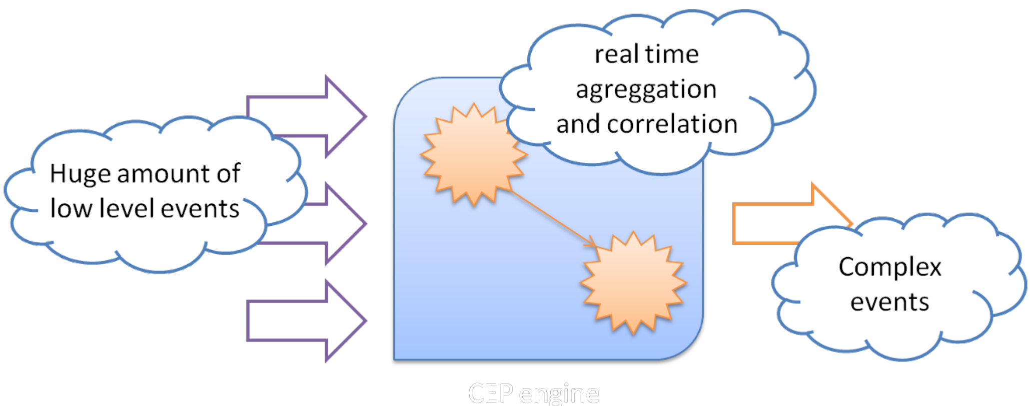



*Figure 2. CEP main principle*

CEP engine is typically offering tools for operating with event streams and event pattern specification and detection. There are forming standards for Event Processing Language (EPL), which allows to define event patterns quite easily. See example:

```
INSERT INTO
 NotFollowedEvent
SELECT
 *
FROM PATTERN
 [every A -> (timer:interval(10 sec)
 and not B)]
```

This example is detecting situation, where event A occur but it is not followed by event B within 10 seconds.

**Chimney platform**

We designed platform for developing advanced monitoring systems called Chimney. It is based on Java and uses open source CEP engine Esper [2] for processing EPL statements. Architecture of Chimney platform is described in following figure.

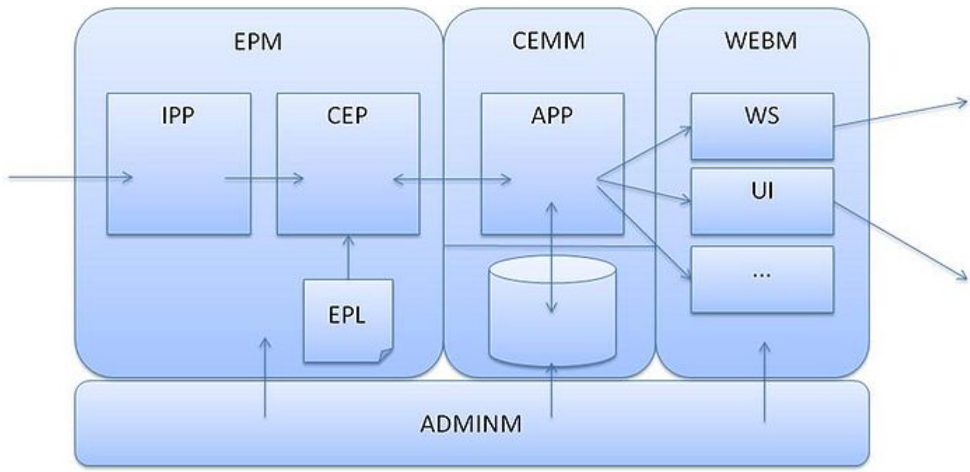



*Figure 3. Chimney architecture*

EPM – Event Processing Module Event Processing Module (EPM) manages data processing and pre-processing.

IPP – Input Pre-Processing manages inputs from various data sources and input preparation for complex event processing. It's typical task is format unification (e.g. timestamps), time synchronization (sort incoming events according to time of origin).

CEP – Complex Event Processing (CEP) interprets Event Processing Language statements. Chimney platform uses CEP engine Esper for that task.

EPL – Event Processing Language (EPL) is language designed for processing huge event streams.

CEMM – Complex Event Management Module manages persistent storege of detected situations by EPM module. There can be used various database engines. Chimney platform is now able to work with PostreSQL and Apache Derby. Communication with Event Store is done using JPA2.0 standard.

APP – specific application modules which handles results from EPM. Module implements two main functions: 1) storing complex events in Event Store 2) providing stored complex events and relevant context data to moduls which publish results to users or other systems.

WEBM – Web Module manages result publishing. There can be used standardized web interface (UI). Web service (WS) module is responsible for publishing information to other systems.

ADMINM – Administrator Management helps to configure, execute and tune event processing functions.
Chimney platform provides useful tools for developing advanced monitoring solutions of maturity level 4. Several monitoring application was build on Chimney.

## Conclusion

We described how maturity of monitoring solutions can be measured. We presented Complex Event Processing technology as a suitable approach for building advanced monitoring systems. We described specific platform Chimney for developing Complex Event Processing applications.


## Acknowledgment

This work was supported by Czech Ministry of Industry and Trade, TIP project FR-TI1/332


## References


Event Processing Technical Society 2011, www.ep-ts.com
Esper-tech 2011, www.espertech.com

# An evolutionary approach to optimally distributing a fleet of robots for weed control in arable crops based on economical costs

Jesus Conesa-Muñoz, Angela Ribeiro
*Center for Automation and Robotics (CSIC-UPM)*
E-mail: jesus.conesa@car.upm-csic.es

**Introduction**

The linear disposition of a crop strongly influences the preferred method of executing many of the operations applied when farming a field. For most of these operations, it is necessary to cover the field following the crop rows because not doing so would be harmful to the crop. This is a very important point to keep in mind to automate ground treatments for arable crops.

One of the most important operations that requires crop lines to be followed closely is weed control. After sowing, weed control is a critical step; depending on the crop, more than one application may be necessary if post-emergency treatment is required. Treatments can be mechanical or chemical, and the selected mode of delivery depends on the weed species to be removed and the way the weeds are distributed in the crop.

If the weed grows in patches, it can be removed by spraying herbicides. Sprayed herbicide treatments can be selective or non-selective. In a non-selective case, the farmer sprays the entire crop with the same herbicide dosage. In a selective case, the treatment is only applied where it is needed. The crop's weed density is approximated to ensure an efficient and complete treatment, with the lowest amount of agrochemicals, which lowers the cost of the treatment and helps to minimise environmental damage.

Information about a crop's weed density can be obtained from discrete sampling, followed by the use of interpolation techniques or by continuous sampling using an aircraft or satellite. In addition to giving us a rough idea of the level of infestation of the crop, any of these procedures may also provide the approximate distribution of weed patches. Once this information is well known and represented as a map of weeds, it can be used to improve the performance of the sprayed herbicide treatment, minimising the total amount of herbicide required.

A vehicle performing a weed treatment must cover the entire field, which requires close attention by the driver, who must receive frequent support from visible references and the operators in charge of changing the location of the references in each pass. The total crop cover in the treatment operation is so important and complex that commercial guidance systems have recently been introduced to help farmers follow crop rows. The driver performs a trace that the guidance system uses as a reference, determining the position of the parallel routes to cover the area to be treated. However, the driver has to manually drive the vehicle to the beginning of each line and execute all turns.

Many studies have aimed to find the best way to cover an entire crop without driver intervention in the headers; the best path is chosen according to cost criteria. In determining the most appropriate set of cost criteria, many of the factors that play an important role in agricultural operations must be taken into account. Thus, in Stoll (2003), the strategy of operation, the surrounding areas, the geometry of the field, field-specific data, specific data on the machinery and the terrain are all considered when evaluating the cost of the path. In this case, areas adjacent to the field are checked to determine whether they can be used in the rotation. The direction of the operation is set at the longest side of the field, and turns in the header are set according to the direction and the surrounding areas. Previous authors have ignored the issue of relief because of the complexity of including it. Taïx *et al.* (2006) have taken factors such as slope, operation direction and the obstacles size in the field (i.e., ditches, tension towers, and wells) into account when calculating the cost. Additional aspects that must be taken into consideration are soil erosion and the possible crop damage that occurs as a result of the way the vehicle's wheels move between the crop rows.

In the previous literature, many planning methods of obtaining routes that allow efficient distribution of the crop can be found. Visal and Oksanen (2007) propose a path planning method that uses two strategies: top-down and bottom-up. In the top-down strategy, complex fields are divided into smaller, less complex, trapezoidal zones that are then merged together in larger areas when certain conditions are reached. Some heuristics are used with the aim of determining the best driving directions. In the bottom-up strategy, the problem is solved recursively by following the contour of the edges of the field, going straight ahead until the field is completely covered. All possible routes are verified and the most efficient is selected.

Taïx *et al.* (2006) define different areas (work, turning, entry, and exit) and a direction for crossing the field. The route-planning method is based on covering all areas with less cost, starting at an entry area and ending at an exit. The best Hamiltonian path and geometric reasoning are used to find the best route that minimises a cost criterion.

Jin and Tang (2006) propose an algorithm for finding the minimum decomposition and the coverage direction. In it, both the field and the obstacles are described by polygons. The algorithm decomposes the field into multiple regions, and each region is covered with a Boustrophedon route (Choset 1997), applying the algorithm recursively to obtain the best cover. Finally, the subregions are matched optimally through an approach based on the travelling salesman problem, where the constraints are the entry and exit areas of the field.

Sorensen (2004) optimises the driving pattern based on prior information about the field, the vehicle and the treatment tools, determining optimal routes for headings, the main field and the sequence of the operational tasks. The problem is approached as a search into the space of possible solutions using optimisation techniques to find the best route.

As noted above, the change of direction in the headers is an important issue because of the long time required for this operation. Thus, Stoll (2003) calculates the rotation, keeping the effective width, the minimum turning radius, the conduction velocity and the rotation acceleration of the vehicle in mind. It also includes additional time to consider the change of direction in the rotation. Bochtis and Vougiakas (2008) distinguish three types of rotation and minimise distances in the non-crop zones by applying these different types of rotations.

The previous works reviewed show the complexity of path planning in the agricultural context. This problem becomes even more complex when a fleet of robots is used to perform the weed treatment, especially when they are small-size robots with insufficient capacity to individually handle an entire field.

Systems of multiple robots engaged in a collective behaviour to carry out an overall task are an important and current research challenge (Cao *et al.* 1997). In many situations, multiple robots operating in the same environment provide a good strategy for handling complex problems that are difficult (or impossible) for a single robot to accomplish. Moreover, performance benefits can be gained from the labour division among the robots. In some contexts, using several simple robots can be easier, cheaper, more flexible and more fault-tolerant than using a single powerful robot for the overall task.

Without a doubt, the agricultural world is an area in which robotic automation fits well. As we have explained, many agriculture operations can be automated and performed by a robot (vehicle); these tasks could also be accomplished by a fleet of robots. However, scheduling the paths of multiple robots is a very complex task because of the sheer number of factors that must be kept in mind when there is more than one mobile platform involved and even more when the members of the fleet are not homogeneous, that is, when they have different features for moving, spraying and turning so that different costs per task can be used and combined.

One of the few works that considers the problem of the management of a robot fleet in an agricultural context is Sørensen and Bochtis (2010), but this study only outlines a conceptual model, and the work is not verified in a practical way with economic and real costs.

In this chapter, an evolutionary approach to solving the management of a fleet with heterogeneous robots for weed control in arable crops is described. The vehicle operations are modelled in terms of costs, based on economic data presented by Pedersen and Fountas (2006). In this manner, the economic feasibility of the management solutions obtained by our approach are verified in real terms.

In summary, we propose a new approach based on a Genetic Algorithm (GA) (Goldberg 1989) (Michalewicz 1999) to effectively distribute a fleet of heterogeneous and small robots working simultaneously to carry out the selective treatment of weeds over a whole field.

For the sake of clarity, the chapter has been organised in the following manner. Section 2 presents a brief description of the problem and our objective. Section 3 describes the proposed approach. Section 4 shows and analyses the results achieved from some synthetic cases, taking the features and constraints of real cases into account. Lastly, section 5 presents the main conclusions and points out future research issues.

**Problem definition**

The problem can be expressed as follows: given a set of robots (*R*) with certain features (i.e., herbicide capacity, motion characteristics, and width of the treatment bar), a field with determined dimensions, a crop organised in rows and a map of the weed patches, the aim is to find the subset of robots and their associated paths that ensure complete coverage of the weed at the minimum cost.

$$\begin{gathered}(r_1, c_1, h_1), \quad \ldots, \quad (r_i, c_i, h_i), \quad \ldots, \quad (r_n, c_n, h_n) \\ \Big| r_i \in R \quad \wedge \quad \bigcup_{i=1}^{n} c_i = C \quad \wedge \quad \sum_{i=1}^{n} h_i = H \\ \forall n \in \mathrm{N}\end{gathered} \tag{1}$$

where *C* is the minimum field cover and *H* is the minimum quantity of herbicide that ensures the treatment of all weed patches existing in the crop.

That is, the problem can be understood as an optimisation issue of finding the multipath plan for the robot fleet that minimises the total cost.

To obtain the minimum cost, more factors than the price of the herbicide used must be considered. To achieve a realistic study, every relevant operation for every robot on the fleet must be modelled in terms of money and time according to its fuel consumption and the time required to complete its task. The relevant operations will be booting, moving, turning and spraying, and they will be computed for every robot in the fleet. In this manner, an accurate estimation of the real cost of a treatment will be performed.

Furthermore, four assumptions must be made to handle the defined problem appropriately: 1) Robots always travel parallel to the crop row, 2) Changes in the direction of movement of the robots leading to movement into the crop are forbidden and can only be made in the headers, 3) Fields have rectangular shapes, and 4) The locations of the weed patches are well known and mapped.

The first and second assumptions are compatible with real situations; to do otherwise would be harmful to the crop, due to the destructive effect of the vehicles on the crop rows. The third is a simplification that makes the development of the first version of the method easier, but does not limit further developments in the method to include more complex field shapes. In fact, in a real agricultural context, although fields can have any irregular geometrical shape, it is quite common to break them down into smaller, trapezoidal regions (Choset 1997) (Oksanen and Visal 2007) and to build a path plan from those regions. The fourth statement assumes the existence of a starting infestation map (locations of weed patches) as input, performed, for example, with aerial or satellite images.

Finally, we require that the solution to the described problem covers the whole field. This is a closer situation to reality, and it is more complex than considering the treatment of the infested part of the crop alone.

Multipath planning approach based on a GA

The proposed approach will internally represent the crop using a matrix with dimensions that depend on the treatment width: the matrix will have a number of columns equal to the field width/treatment width and the number of rows equal to the field length/treatment width. Therefore, each matrix component will represent a herbicide treatment unit, referred to as a cell in the rest of this chapter. Each cell will have a state corresponding to its weed occurrences that specifies whether the cell should treated or not.

According to previous assumptions, finding a path plan for the field is equal to finding the set of matrix columns through which each robot must pass to perform the treatment task. The order in which the columns are treated defines the transitions between them and therefore directly affects the quality of the plan. Furthermore, the optimal plan should minimise distances, turns, herbicide, spraying applications and so on. To understand the complexity of building the optimal path plan, let us suppose a 4-ha field and a field with a width twice its length, a common field size for maize crops. If the robots are small enough that they can only treat the space between two crop rows (75 cm), a plan that covers the whole field will have approximately 400 parallel routes, and the matrix that represents the crop will need 400 columns to store information about weed occurrences. Moreover, if the fleet is composed of $n$ units, and each of these can deal with any of the columns, the number of possible different plans will be $n^{400}$, if the order in which columns are linked is not taken into account. If the order is considered, the number of possible different plans will be considerably higher. The previous simple estimation allows us to demonstrate the high complexity of the problem and the way the solutions are immersed in a very large space of possible solutions and closely resembles the Travelling Salesman Problem (TSP), an NP-hard (Non-deterministic Polynomial-time hard) problem in which given a list of cities and their pairwise distances, the task is to find a shortest possible tour that visits each city exactly once. The TSP can be modelled as an undirected weighted graph in which cities are the graph's vertices, paths are the graph's edges and a path's distance is the edge's length. Similarly, in our path-planning problem, the columns or row crops of our field routes can be seen as the graph's vertices and the transitions between the columns as the edges.

From a practical standpoint, the TSP problem has not yet been solved, and from a theoretical point of view, the techniques used are only approximations that provide acceptable solutions. The classical optimisation algorithms are unable to solve the general problem because of the combinatorial explosion of possible solutions. Various acceptable, yet nonoptimal, solutions have been obtained through various computational techniques, among them GAs (Chatterjee *et al.*, 1996) (Bektas 2006). GAs do not find exact solutions, but they can find sufficiently good approximations (97% optimisation), and they can be applied over very large sets of cities (networks with millions of nodes) within a reasonable computational time on a supercomputer (weeks or months). Indeed, GA approaches match well with problems for which there is no known analytical method of finding the solution and that have a very large search space, both features of the path-planning problem.

Having described the problem mathematically, the next step is to formulate it in such way that GA can be used to resolve it. GAs imitate biological phenomena in which an individual’s genetic information determines its traits and fitness in a particular environment. In a GA, a population of strings, which encode candidate solutions (called individuals) to an optimisation problem, evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and occurs in generations. In each generation, the fitness of every individual in the population is evaluated, and multiple individuals are stochastically selected from the current population (based on their fitness) and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. The algorithm commonly terminates when either the maximum number of generations has been produced or a satisfactory fitness level for the population has been reached. If the algorithm has terminated due to the maximum number of generations, a satisfactory solution may or may not have been reached. In general, a GA has five basic components (Rawlins, 1991). The algorithm must include a representation scheme that provides a genetic structure representative of all solutions to the problem. An initial population of solutions must then be created. An evaluation function to rate the fitness of the solutions must also be developed. Genetic operators must be defined and used to alter the genetic composition for producing new solutions during reproduction; in this way, the evolutionary approach explores the solutions space. Finally, values for all parameters of the GAs must be defined. Next, the representation scheme, the genetic operators and the fitness function will be detailed for the path-planning problem. The initial population and the genetic configuration will be explained in the results section.

Representation scheme

The previous section explained that the path plan will be formed by the columns that each robot in the fleet must treat and the order in which each robot must travel the set of columns assigned to it.

To express the solutions in genetic terms, solutions are represented as vectors with as many components as columns (parallel routes) in the crop. Each component of the vector will contain the identification of the robot in charge of that column's treatment. Expression 2 shows the vector for a crop of *n* columns and a fleet of *m* robots:

$$Individual = \langle r_1, r_2, \ldots, r_n \rangle \,|\, 1 < \mathrm{i} < \mathrm{m} \tag{2}$$

This representation indicates which robot should treat each column in the field but does not explicitly define the order that each vehicle should follow to complete all of the columns assigned to it. To build a complete solution from an individual, the order is determined as follows: 1) Given a robot, look for the column closest to its original position on the two most external columns of the set of which it is in charge, 2) The robot moves from its initial position to the closest column, 3) Run the robot to treat the column, 4) Once the robot has left the column at the opposite end from which it entered, advance by the header to the nearest of the remaining columns, and 5) Repeat steps 3 and 4 until there are no more columns left to deal with.

This process will obtain the optimal path for a set of columns in terms of distance, turns and time. Therefore, enacting the previous process for each of the robots obtains the overall path plan of the entire fleet.

For example, Figure 1 shows the encoded path plan of the individual $\langle r_4, r_2, r_1, r_4, r_4, r_1, r_1, r_1, r_3, r_4 \rangle$ over a field of 10 columns (10 treatment parallel routes). The first and last matrix rows represent the headers of the crop, where a robot can move freely in a horizontal direction without damaging the crop rows. Cells containing weeds are labelled in green. All robots start at the same initial position, the component (1, 5) of the matrix. The robot r1 (Fig. 1a) is in charge of four columns: c3, c6, c7 and c8. The robot r2 (Fig. 1b) must treat c2, r3 (fig 1c) is devoted to the c9 column and r4 (Fig. 1d) treats columns c1, c4, c5 and c10. Using the previously explained strategy, robot r1 first moves from its initial position (1, 5) to the nearest column (c3) of the most external two columns assigned to it (c3 and c8). After treating column c3, the robot will move to the nearest of the assigned columns that remain untreated. This step is repeated until there are no more untreated columns. Plans for robots r2, r3 and r4 are shown in Figure 1.

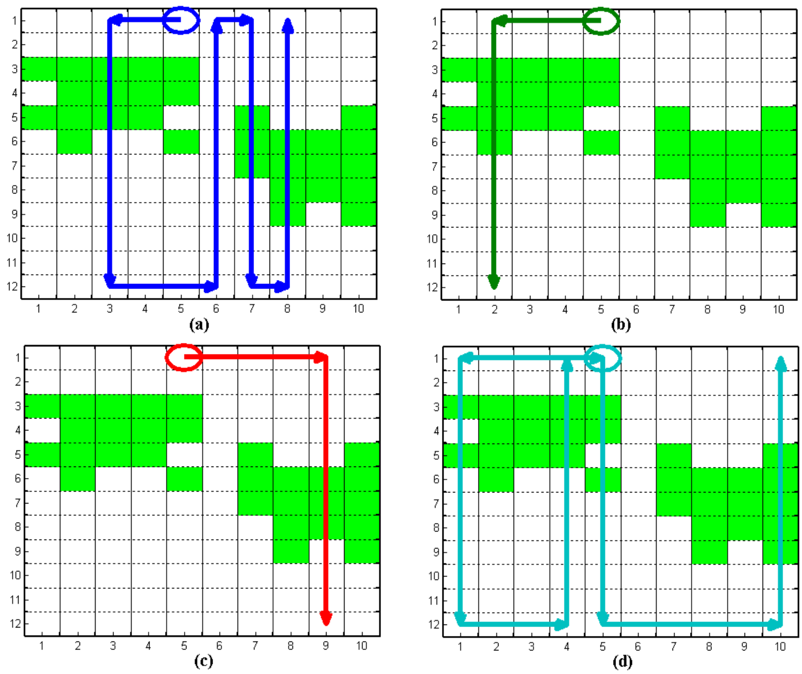


*Figure 1. Plan encoded in the individual* $\langle r_4, r_2, r_1, r_4, r_4, r_1, r_1, r_1, r_3, r_4 \rangle$ *for a field with 10 treatment parallel routes. (a), (b), (c) and (d) represent the paths of the robots r1, r2, r3 and r4. Green cells indicate the presence of weeds. Circles indicate the starting positions.*

This representation avoids, by construction, the collision of robots or vehicles in the middle of the crop, i.e., within a column. Collision may be possible in the headers, but it is assumed that the headers are large enough for the robots to manoeuvre around themselves. Within these regions, the problem of avoiding other elements is not critical and is easy to solve.

**Genetic operators**

In the selection stage, where potentially useful solutions for the next stage (reproduction) are selected, roulette-wheel selection, also known as fitness-proportionate selection, has been used. Usually, a proportion of the wheel is assigned to each individual (solution) based on its fitness value. The proportion assigned is achieved by dividing the fitness of the individual by the total fitness of all individuals in the population, thereby normalising them to 1. A random selection is then made, similar to the rotation of a roulette wheel. With fitness-proportionate selection, there is a chance that some weaker solutions may survive the selection process. However, this is may be an advantage because although a solution may be weak, it may include a component that could prove useful when following the reproduction process.

Regarding the reproduction stage (next generation of solutions), the genetic operators used are quite common. Two-point crossover has been selected for the recombination step, so two points on the individuals are randomly chosen (strings of robot identifiers) to be crossed; everything between the two points is swapped between the individuals, generating two new solutions. Regarding the mutation, it has been taken into account that solutions (individuals) are codified as vectors of integers instead of vectors of bits. Hence, in the mutation, each vector component value has a probability of being reached of 1/m (where m is the number of robots) instead of 1/2, as in the simple GA. The configuration parameters related to these operators will be explained in the results section.

**Fitness function**

The fitness function is a particular type of objective function that prescribes the optimality of a solution (that is, an individual) so that a particular individual can be ranked against all of the other individuals. In the approach proposed in this chapter, the fitness is a cost function that evaluates the appropriateness of a solution encoded in an individual. Because our goal is to measure the economic cost of a weed-selective treatment, the main factors to take into account are the fuel consumed, the herbicide used, and the time spent. In particular, time is considered because most of the agricultural services are hired and paid by the hour, so minimising the time involved minimises the cost as well. To properly calculate the overall and accurate cost of the fuel, time and herbicide used for each individual, the following parameters must be considered.

- Distance: the aim is to find the shortest path that covers the entire field to reduce the fuel consumption and the time of the treatment.
- Turns: turns are expensive operations for vehicles, especially in terms of time. Therefore, the paths obtained should reduce the number of turns.

- Number of robots used: in principle, solutions that involve fewer robots are preferred because no-used vehicles can be employed to simultaneously treat other crop. But fewer robots also mean that the total time of the treatment will be longer.
- Herbicide application: the crop must be completely treated with a minimum amount of herbicide to reduce cost. Each robot can save more or less money spraying herbicide according to its capability for this operation.
- Herbicide tank capacity: it is assumed that each robot has a tank with a specific capacity. When a robot sprays a cell, the quantity of herbicide in its tank decreases; therefore, it is possible that a robot expends all of the herbicide in its tank before finishing the treatment of a row. In this case, the robot has to reach a header to refill its tank before proceeding with the weed treatment; in addition, the column will remain partially treated. This factor is introduced to properly penalise individuals that represent incomplete solutions because not all weeds in the cells are treated.

Considering the above parameters, and supposing a fleet of *n* heterogeneous robots, the fitness function, without taking the time into account, can be expressed as follows:

$$
\begin{aligned}
Cost = & \sum_{i=1}^{n} bootCostRobot_i + \\
& + \sum_{i=1}^{n} distanceRobot_i \cdot moveCostRobot_i + \\
& + \sum_{i=1}^{n} turnsRobot_i \cdot turnCostRobot_i + \\
& \sum_{i=1}^{n} herbicideRobot_i \cdot sprayCostRobot_i
\end{aligned}
\tag{3}
$$

where the costs in euros for the operations booting, moving, turning and applying herbicide applied by robot i to one cell of the matrix are $bootCostRobot_i$, $moveCostRobot_i$, $turnCostRobot_i$ and $sprayCostRobot_i$, respectively. These costs are calculated according to economical and real values in the results section. The $distanceRobot_i$ and $turnsRobot_i$ variables refer, respectively, to the meters and turns carried out by the robot *i*. The $herbicideRobot_i$ variable is the number of crop matrix cells with weeds treated by the robot *i*.

For this fitness, the fuel and the herbicide used determine the definitive cost of the treatment; however, the time is not considered. To include time costs in our study, the total time spent to accomplish the overall task must be calculated. Because the robots work in parallel, this time is the maximum among the times of each robot. Consequently, the total time can be expressed as (4):

$$TotalTime = max(TimeRobot(i))$$

$$TimeRobot(i) = bootTimeRobot_i + distanceRobot_i \cdot moveTimeRobot_i + \\ + turnsRobot_i \cdot turnTimeRobot_i + herbicideRobot_i \cdot sprayTimeRobot \tag{4}$$

As in the case of monetary cost, the costs in time to boot, move one cell, make one turn and apply one unit of herbicide for robot *i* are $bootTimeRobot_i$, $moveTimetRobot_i$, $turnCostRobot_i$ and $sprayCostRobot_i$, respectively. These times are calculated in the results section.

The $distanceRobot_i$ and $turnsRobot_i$ variables refer, respectively, to the meters moved and turns carried out by the robot *i*. The $herbicideRobot_i$ variable is the number of crop matrix cells with weeds treated by the robot *i*.

Once the total time is obtained, the new fitness considering both cost of fuel, herbicide and times can be expressed as (5):

$$NewCost = \cos t + TotalTime \cdot salaryWor\text{ker} \tag{5}$$

where *salaryWorker* is the hourly wage of the worker in charge of supervising the task performed by the fleet of robots.

To precisely fix the aim of the genetic search and improve the activity of the operators as in a roulette-wheel selection, the fitness function is normalised to range from 0 to 1. For this normalisation, it is necessary to estimate the minimum and maximum values of the function.

The minimum value can be assessed as the total coverage path that meets the following conditions: 1) It passes only once for each cell in the field (minimum distance), 2) It makes the fewest turns, that is, 2 per column (one to enter in the column and another to exit of it), 3) It uses the minimum number of robots, that is, 1, and 4) It applies the minimum cost for spraying herbicide with the maximum possible tank size (expression 6).

$$MinimumCost = \\ = \min imumBootCost + \min imumDis\tan ceCost + \min imumTurnsCost + \min imumHerbicideCost = \\ = \min(bootCostRobot_i) + columns \cdot rows \cdot \min(moveCostRobot_i) + \\ + (2 \cdot columns - nRobots) \cdot \min(turnCostRobot_i) + weeds \cdot \min(sprayCostRobot_i) \tag{6}$$

Similarly, the maximum value of the function can be approximated by the path that covers the whole crop field, providing the following conditions: 1) The maximum distances in the transitions between columns, 2) The maximum number of turns, 3) The maximum number of robots, and 4) The highest cost of the herbicide with the minimum possible tank size (expression 7).

$$\begin{aligned}&MaximumCost = \\&= AllBootCost + maximumDistanceCost + maximumTurnsCost + maximumHerbicideCost = \\&= \sum_{i=1}^{nRobots} bootCostRobot_i + (columns \cdot rows + maximumTransitionsCost) \cdot max(moveCostRobot_i) + \\&+ 2 \cdot columns \cdot max(turnCostRobot_i) + weeds \cdot max(sprayCostRobot_i)\end{aligned} \tag{7}$$

Finally, the normalised value of the fitness is calculated by expression 8.

$$fitness = \frac{(Cost - MinimumCost)}{(MaximumCost - MinimumCost)} \tag{8}$$

**Results and discussion**

This section presents a set of examples to demonstrate the performance of the proposed approach.

The values of the parameters of the GA for all tests have been established as follows. The crossover probability has been fixed to 0.8. The mutation has been applied using a random binary mask with a 0.2 probability. The population is composed of a set of 100 individuals or solutions. Furthermore, an elitism operator has been added to the conventional genetic operators to facilitate the convergence of the algorithm. In this case, the best individual of a generation replaces the worst one of the next generation.

The characteristics of the crops used in the trials were fields of approximately 4.5 ha with a width twice their length. According to these dimensions, a 200x400 matrix is required to represent the weed map if small robots with the capacity to treat only the space between two rows (bar size) are used. Consequently, the vector that encodes the solution (individual) has 400 components. In every trial, it has been supposed that 30% of the field is infested by weed. This means that approximately 24,000 components of the matrix (cells) have a weed occurrence. To obtain weed-heterogeneous distributions similar to real situations, the central positions of the weed patches are randomly generated. Additionally, the starting position for the robots of the fleet is fixed on the left side of the upper header, that is, on element (1, 1) of the matrix.

To test realistic experiments and obtain realistic results, the main operations for every robot of the fleet have been modelled in terms of money and time, in particular, in euros and seconds. Therefore, any robot operation to boot, move, turn or spray has an associated financial and time. The financial costs have been calculated considering the fuel and the herbicide used to apply a weed treatment for 1 ha. Pedersen and Fountas (2006) assume 776 € for the fuel (and other additional costs) for executing 3 weed treatments (1-year cycle) over 80 ha with a bar width of 4 rows. Consequently, the fuel used to treat 1 ha with only 1 weed treatment is 776 €/80 ha/3 treatments, that is, 3.23 €/ha/ treatment. Because Pedersen and Fountas (2006) consider a 4-row spraying bar, and we assume (for these experiments) a fleet of small robots that only deal with 1 row (to simplify), the definitive fuel cost could be estimated as 3.23*4 = 12.93 €/ha/treatment. Furthermore, because our field is 4.5 ha, to treat the entire field costs

12.93*4.5 = 58.2 €/field/treatment. Because our fields are discretised into cells, with each cell having an area of 0.75m*0.75m = 0.5625 $m^2$, covering the field for one of the robots involves covering 4.5 ha/0.5625 = 80,000 cells (therefore, the field is represented by a 200x400 matrix), or 80,000 motion operations between cells, 1 booting operation, and about 400 turns (changing from a row crop to the next one implies 1 turn). Therefore, 1 booting, 80,000 motion operations and 400 turns cost 58.2 €. Because a motion operation covers 0.75 m and turns with agricultural machinery are always expensive, we have decided to model every turning operation to 10 m (turning radius = 3.2 m, which fits the real conditions). Finally, considering the price of the fuel consumed in booting as 0.2 € and the distances and number of executed operations, 58.2 – 0.2 = 58 € supposes 80,000*0.75 m (distance for motion operations) + 400*10 m (turning distances). Consequently, moving 1 m costs $0.91 \cdot 10^{-3}$ €, the cost of moving per cell is $0.75 * 0.91 \cdot 10^{-3} = 0.68 \cdot 10^{-3}$ €, and the cost per turn is $10 * 0.91 \cdot 10^{-3} = 0.91 \cdot 10^{-2}$ €.

On the other hand, Pedersen and Fountas (2006) estimate 90.2 € as the cost of the herbicide to treat 1 ha, so 4.5*90.2 = 405.9 € is the price for the entire field. Because we assume a 30% infestation, the definitive value is 0.3*405.9 = 121.77 €. Therefore, 30% of the field is infested, and 0.3*80,000 = 24,000 cells present weeds. There are 24,000 spraying herbicide operations, and each one of them costs 121.77/24000 = $0.51 \cdot 10^{-2}$ €/cell.

To calculate the costs in time, the speed of the robots must be taken into account. A typical speed for spraying herbicide is 6 km/h. Considering this fact and knowing the cell length, the motion time per cell can be calculated as 0.75 m * 3,600 s/6,000 m = 0.45 s. Because turns are expensive operations for agricultural machinery and because they are carried out in the headers where the speed can be increased, we use a typical speed for manoeuvring, 8.5 km/h: the time per turning is 10 m * 3,600 s/8,500 m = 4.24 s. Considering these two parameters, the total time to completely treat a field with only one robot is at least 80,000*0.45 + 400*4.24 = 37696 s = 10.47 hours (remember that only one robot is covering a distance of 60 km at 6 km/h). Therefore, the small amounts of time used to boot and spray can be ignored; that is, they are equal to 0.

These values and their slight variations will be used in the following experiments. They could be considered canonical prices or reference costs for building different and coherent economic conditions to test different but realistic situations in this agricultural context. They are summarised in table 1.

| Criterion \ Operation | Boot | Move (1 cell) | Turn (1) | Spray (1 cell) |
|---|---|---|---|---|
| Money (€) | 0.2 | $0.68 \cdot 10^{-3}$ | $0.91 \cdot 10^{-2}$ | $0.51 \cdot 10^{-2}$ |
| Time (s) | 0 | 0.45 | 4.24 | 0 |

*Table 1. Reference values in euros and seconds for the main operations in a weed treatment*

Finally, every robot will have a tank with a capacity modelled according to the number of infested cells it is able to treat without stopping to refuel. Treating each of these infested cells implies one spraying operation.

In the following tests, two definitive costs will be calculated and compared. The first is the associated cost of planning without considering the time. On the other hand, the second takes the time spent into account, using the typical hourly salary for a medium worker according to Pedersen and Fountas (2006). This last parameter is set to 27 €/hour.

Test 1

In the first experiment, a large fleet of 10 robots is considered. All mobile platforms have equal characteristics except motion cost (in money) per cell, as shown in table 2. To justify these different motion costs, we assume different performances or fuel consumption levels between the vehicles in the fleet, due, for example, to their age, tire inflation, or even the skill of the driver.

Moreover, all the robots have the same tank capacity for completing the same number of infested cells: 8,000 (8,000 spraying operations).

| Operations / Robot | Boot | | Move (1 cell) | | Turn (1) | | Spray (1 cell) | | Tank (cells) |
|---|---|---|---|---|---|---|---|---|---|
| | € | s | € | S | € | s | € | s | |
| 1 | 0.2 | 0 | $0.43{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 2 | 0.2 | 0 | $0.46{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 3 | 0.2 | 0 | $0.49{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 4 | 0.2 | 0 | $0.52{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 5 | 0.2 | 0 | $0.55{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 6 | 0.2 | 0 | $0.58{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 7 | 0.2 | 0 | $0.61{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 8 | 0.2 | 0 | $0.64{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 9 | 0.2 | 0 | $0.67{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |
| 10 | 0.2 | 0 | $0.7{\cdot}10^{-3}$ | 0.45 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.51{\cdot}10^{-2}$ | 0 | 8,000 |

*Table 2. The estimated costs for the robots in the fleet in test 1*

The best solution (without considering time) is achieved with the cheapest 3 robots (R1, R2 and R39) covering the entire crop, as shown in Figure 2. A first consideration could lead to a solution involving more than one robot, but because there are penalties for booting, and time is not considered, the best solution is that which uses the cheapest robots and has enough tank capacity to treat the entire crop. Because 24,000 spraying operations are required, and each robot has a tank capacity of 8,000, three robots are used.

For the time solution, the rows appear to be distributed among the fleet. However, as shown in Figure 2, the robots accomplish their tasks in proportion to their features. Respectively, they carry out 8, 15, 23, 30, 37, 44, 51, 58, 64 and 70 rows.

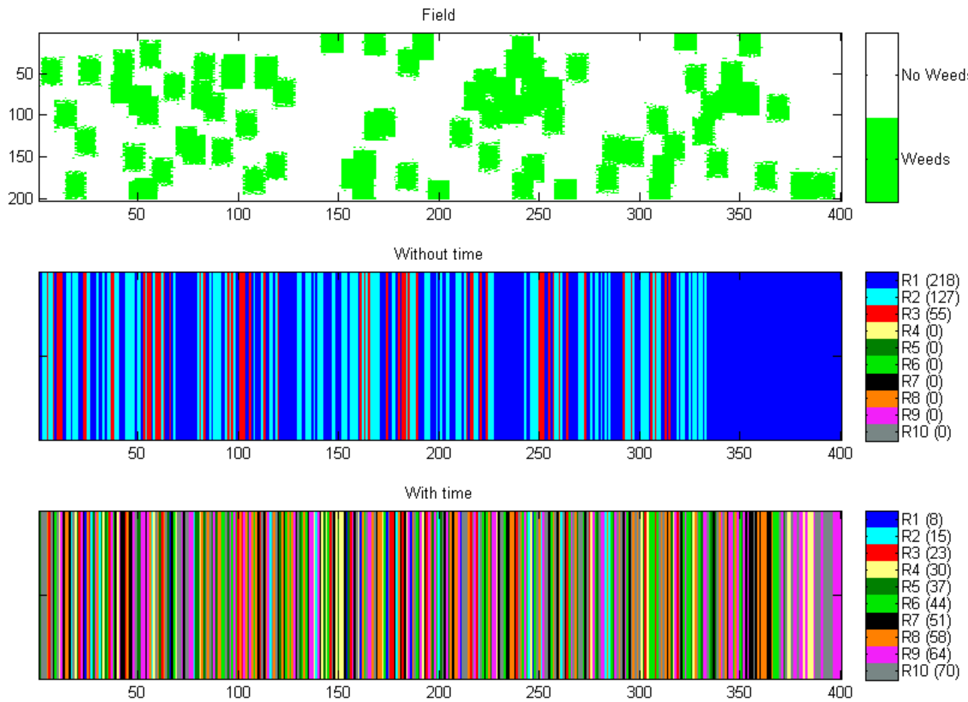



*Figure 2. Weed distribution in the crop used in test 1. The best path plan uses only the 3 cheapest robots without considering the time of treating the entire crop, and the entire fleet is used in proportion to its motion cost for the time simulation.*

Similar trials have been carried out with fleets of homogeneous robots, except for one of their characteristics: the different costs for the turns and for the herbicides. The results obtained are always the same: the lowest cost robots appear in the solution covering all of the columns in the plan without considering time, and in the cases considering time, the total task is shared in proportion to the prices of the operations. Indeed, the solutions found match the expected behaviour of the algorithm.

Test 2

According to the results of test 1, it can be concluded that if a robot is cheaper to operate than the rest of the robots in the fleet, it will be selected to cover the main part of the field as far as its tank capacity allows it without stopping to refuel (additional costs). On the other hand, for time simulations, the entire fleet is launched in order to properly parallelise the total task. However, it could be interesting to analyse what happens if the tank capacities are increased and no robot is any better than others; that is, if all robots have the same costs for any operation. Test 2 has been designed to check the behaviour of these cases of a small homogenous fleet with high tank capacities (table 3). Notice that in this experiment, the fleet is composed of four robots.

| Operations / Robot | Boot | | Move (1 cell) | | Turn (1) | | Spray (1 cell) | | Tank (cells) |
|---|---|---|---|---|---|---|---|---|---|
| | € | s | € | S | € | s | € | s | |
| 1 | 0.2 | 0 | $0.68 \cdot 10^{-3}$ | 0.45 | $0.91 \cdot 10^{-2}$ | 4.24 | $0.51 \cdot 10^{-2}$ | 0 | 20,000 |
| 2 | 0.2 | 0 | $0.68 \cdot 10^{-3}$ | 0.45 | $0.91 \cdot 10^{-2}$ | 4.24 | $0.51 \cdot 10^{-2}$ | 0 | 20,000 |
| 3 | 0.2 | 0 | $0.68 \cdot 10^{-3}$ | 0.45 | $0.91 \cdot 10^{-2}$ | 4.24 | $0.51 \cdot 10^{-2}$ | 0 | 20,000 |
| 4 | 0.2 | 0 | $0.68 \cdot 10^{-3}$ | 0.45 | $0.91 \cdot 10^{-2}$ | 4.24 | $0.51 \cdot 10^{-2}$ | 0 | 20,000 |

*Table 3. Costs for each robot of the fleet in test 2*

As shown in Figure 3 for the simulation without time, only 2 robots are booted (because the additional tank capacity covers the entire treatment). In the obtained solution, the robots involved are R2 and R3, but in additional simulations it will be established that because all the robots present the same features, any 2 of them can be selected to accomplish the planning.
In the time simulation, because we are working with a homogenous fleet, all the robots do the same work: a treatment of 100 rows of crops.

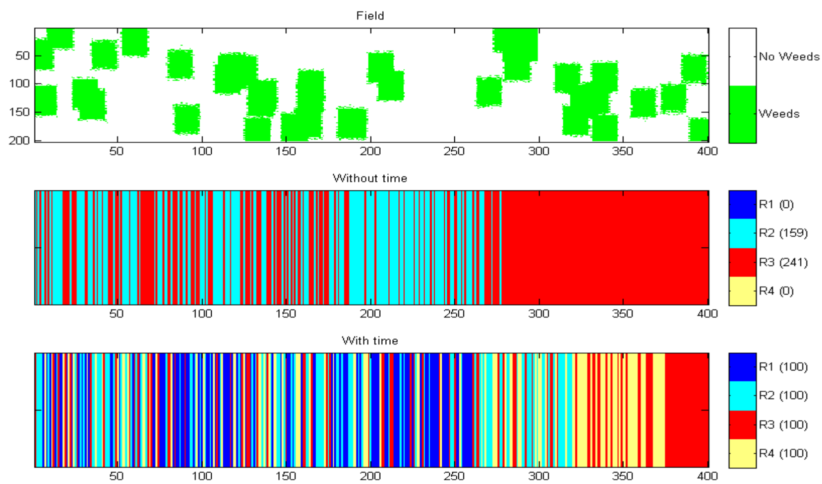



*Figure 3. Weed distribution in the crop used in test 2. In this case, any 2 robots in the fleet are booted for the simulation without considering time. If time is considered, the entire fleet is launched and carries out the same total work: 100 rows.*

Test 3
According to the results of test 1, it can be concluded that if the criteria without time are considered, the cheapest and minimum number of robots will be selected to cover the whole field. This is a very clear case when a robot has features better than or equal to the rest of the fleet. If the time criterion is evaluated, the algorithm distributes the task equally among all the robots. However, this is not the desired behaviour if the robots are not homogeneous in their time operations. In this case, the faster robots can accomplish more work in the same amount of time. Test 3 has been designed to check the behaviour when the costs in money and time for moving are different among the robots (Table 4). This test fits with situations in which, for

example, we deal with 2 robots with a smaller working speed (R1 and R2). Due to this slow speed, the spraying operations are more accurate; therefore, more herbicide is saved.
Notice that in this experiment, the fleet is composed of four robots, and the costs and tanks for both criteria are the same except for moving and spraying.

| Robot \ Operations | Boot | | Move (1 cell) | | Turn (1) | | Spray (1 cell) | | Tank (cells) |
|---|---|---|---|---|---|---|---|---|---|
| | € | s | € | s | € | s | € | s | |
| 1 | 0.2 | 0 | $0.5\cdot10^{-3}$ | 0.9 | $0.91\cdot10^{-2}$ | 4.24 | $0.42\cdot10^{-2}$ | 0 | 20,000 |
| 2 | 0.2 | 0 | $0.5\cdot10^{-3}$ | 0.9 | $0.91\cdot10^{-2}$ | 4.24 | $0.42\cdot10^{-2}$ | 0 | 20,000 |
| 3 | 0.2 | 0 | $0.68\cdot10^{-3}$ | 0.45 | $0.91\cdot10^{-2}$ | 4.24 | $0.51\cdot10^{-2}$ | 0 | 20,000 |
| 4 | 0.2 | 0 | $0.68\cdot10^{-3}$ | 0.45 | $0.91\cdot10^{-2}$ | 4.24 | $0.51\cdot10^{-2}$ | 0 | 20,000 |

*Table 4. Costs for each robot of the fleet in test 3*

As shown in Figure 4, two robots are booted when the time is not considered and are distributed to minimise the number of herbicide refills as well as the total distance travelled. The two robots used are R1 and R2, which are the least expensive for moving and spraying. For the time simulation, four robots are booted. However, R3 and R4, which are faster (time) than R1 and R2, treat many more columns, specifically 68, 68, 132 and 132 columns for robots R1, R2, R3 and R4, respectively. This result makes sense according to the proportions of the operation times.

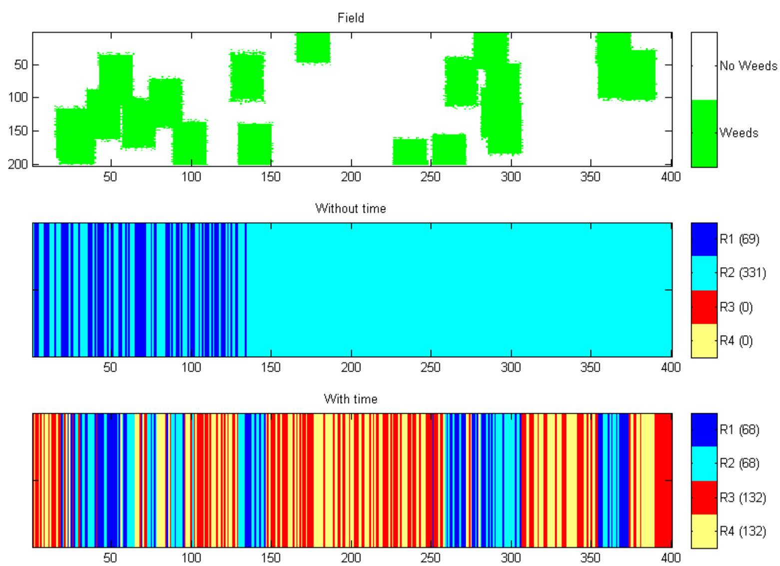



*Figure 4. Weed distribution in the crop used in test 3. The optimal solution is a path plan with two robots if the time is not considered. Four robots are used if the time is considered, and the two fastest robots complete more rows.*

Test 4

Regarding both criteria, with and without time, the previous tests have shown that the algorithm is able to distinguish among the features of the robots and adapt them to their most appropriate tasks. Test 4 aims to verify this fact with a very interesting situation: when robots have inverse time and money costs for moving and spraying. In this case, the best plan is to assign a robot with a small moving cost and a high spraying cost to the treatment of the columns with fewer weeds, whereas a robot with a small spraying cost and a high moving cost is placed in charge of the columns that are most infested. The established values for the operation costs are shown in Table 5.

| Operations / Robot | Boot | | Move (1 cell) | | Turn (1) | | Spray (1 cell) | | Tank |
|---|---|---|---|---|---|---|---|---|---|
| | € | s | € | s | € | S | € | s | (cells) |
| 1 | 0.2 | 0 | $0.7{\cdot}10^{-3}$ | 0.6 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.2{\cdot}10^{-2}$ | 0 | 20,000 |
| 2 | 0.2 | 0 | $0.6{\cdot}10^{-3}$ | 0.5 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.4{\cdot}10^{-2}$ | 0 | 20,000 |
| 3 | 0.2 | 0 | $0.5{\cdot}10^{-3}$ | 0.4 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.6{\cdot}10^{-2}$ | 0 | 20,000 |
| 4 | 0.2 | 0 | $0.4{\cdot}10^{-3}$ | 0.3 | $0.91{\cdot}10^{-2}$ | 4.24 | $0.8{\cdot}10^{-2}$ | 0 | 20,000 |

*Table 5. Costs for each robot of the fleet in test 4*

As shown in Figure 5, the best solution obtained for test 4 when the time is not considered involves only 3 robots. Robot 1 is the robot with the lowest moving cost and the highest spraying cost. For that reason, it is in charge of treating the columns with a very high number of weed cells. Robot 4, which has the lowest spraying cost but the highest moving cost, is devoted to covering empty or almost-empty columns. Finally, robot 2 treats the columns with a small number of weeds because it has a medium cost for both moving and spraying. For the time simulation, to parallelise the overall task, the rows are more evenly distributed but keep the desired behaviour: the robots with the best spraying performance head to more infested columns, whereas the robots with better motion time lead to fewer infested zones.

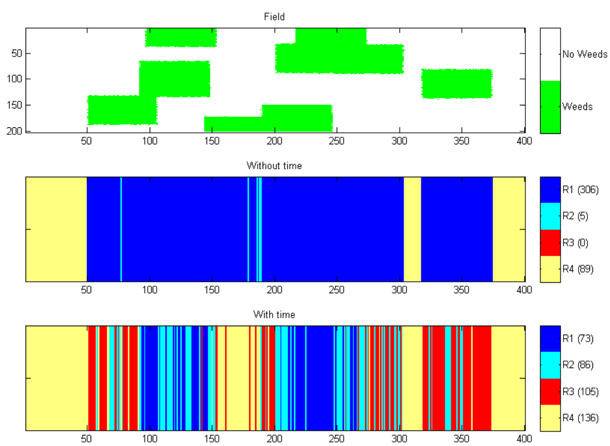



*Figure 5. When time is not considered, the solution involves 3 robots that cover, respectively, columns with a low, medium and high number of weed cells. In the time criterion case, the solution shows a more distributed situation according to the features of each robot.*

Finally, a set of randomly generated trials has been conducted to measure the mean fitness obtained by the proposed approach. In all cases, the initial fleet has been composed of 5 robots. The costs for these robots are the reference prices shown in Table 2. The level of infestation ranged from 10% to 30% and the number of patches from 1 to 15. Table 6 summarises the most important data obtained from these experiments. The best solution for each trial always involves two robots, and the mean fitness is approximately 0.06, which is very close to the optimal solution that would only involve one robot. The situation will change if the time is taken into account; in this case, to reduce the time of the treatment, the best solutions would probably involve more than two robots.

| Infestation level (%) | Number of patches | Size of lowest patch (number of cells) | Size of highest patch (number of cells) | Best Fitness (without time) | Final Cost (without time) | Best Fitness (with time) | Final Cost (with time) |
|---|---|---|---|---|---|---|---|
| 26% | 12 | 1,136 | 10,256 | 0.0751 | 180.96 | 0.0036 | 239.4 |
| 11% | 14 | 400 | 1,240 | 0.0443 | 180.66 | 0.0072 | 240.9 |
| 27% | 3 | 4,812 | 10,065 | 0.0726 | 180.93 | 0.0032 | 239.2 |
| 29% | 14 | 1,167 | 6,572 | 0.0699 | 180.9 | 0.0046 | 239.9 |
| 24% | 10 | 880 | 4,905 | 0.0751 | 180.96 | 0.0061 | 240.7 |
| 25% | 2 | 10,000 | 10,000 | 0.0824 | 181.16 | 0.0031 | 239.1 |
| 13% | 9 | 550 | 1,394 | 0.0544 | 180.72 | 0.004 | 239.6 |
| 16% | 7 | 1,826 | 2,634 | 0.0445 | 180.68 | 0.0039 | 239.4 |
| 11% | 15 | 369 | 1,228 | 0.0401 | 180.59 | 0.0062 | 240.7 |
| 26% | 11 | 325 | 1,002 | 0.0737 | 180.94 | 0.0086 | 241.2 |

*Table 6. Summary of the most important data from the experiments carried out. The trials have been generated randomly.*

**Conclusions**

Path planning management for optimally economical schedules in an agricultural context involves obtaining the best path for a vehicle in terms of any desired criterion (in our case, cost in euros). The proposed approach is based on a genetic algorithm and offers a method of obtaining the best solution for not only a single vehicle but also a fleet of several mobile platforms. Furthermore, this approach is able to take a wide range of parameters for each vehicle into account, such as the herbicide storage capacity and the minimisation of both turns and distances. It can even determine the best number of vehicles to be used in the treatment in a number of real situations.

A set of varying experiments under heterogeneous conditions have been tested, and the path plans obtained have taken the conditions into account and been properly adapted to minimise the global cost. The solutions reached are strong enough to be considered optimal path plans. Indeed, the mean fitness for the best solutions obtained in the experiments is approximately 0.06 without considering time and 0.003 with time, which are very close to optimal solutions that would involve only one robot with cost 0.

In summary, a reasonable approach to a complex problem has been designed and implemented. The results presented are part of the first simplified approach but do not alter the effectiveness of the proposed method.

The next developments of the proposed approaches must include the following aspects:

- Solving for irregular fields. The same strategy can be used, but variable instead of fixed-length columns must be considered. The new approach must also correctly deal with the transitions between columns.
- Combining the two criteria, time and money, as a multi-objective optimisation instead of using a linear combination.

### Acknowledgements

The Spanish Ministry of Science and Innovation and the European Union have provided full and continuing support for this research work through projects PLAN NACIONAL - AGL2008 - 04670 - C03 - 02/AGR (PERICIA) and UE - CP - IP245986 - 2(RHEA).

### References

Bochtis, D.D. and Vougioukas, S.G. 2008. Minimising the Non-working Distance Travelled by Machines Operating in a Headland Field Pattern. Biosystems Engineering, 101(1): 1-12.

Cao, Y. U., Fukunaga, A. S., Kahng, A. B., and Meng, F. 1997. Cooperative mobile robotics: antecedents and directions. In Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots. 27:226-234.

Choset H. and Pignon P. 1997. Coverage Path Planning: The Boustrophedon Decomposition. In Proceedings of the FSR'97: International Conference on Field and Service Robotics, edited by Alexander Zelinsky.

Goldberg, D. E. 1989. Genetic Algorithms in Search, Optimization and Machine Learning, Kluwer Academic Publishers, Boston, MA

Jin, J., and Tang L. 2010. Optimal Coverage Path Planning for Arable Farming on 2D Surfaces Transactions of the ASABE. 53(1) 283-295.

Michalewicz, Z. 1999. Genetic Algorithms + Data Structures = Evolution Programs, Springer Verlag.

Oksanen, T. and Visala, A. 2007. Path Planning Algorithms for Agricultural Machines. Agricultural Engineering International: the CIGR Ejournal. Manuscript ATOE Vol. IX.

Pedersen, S. M., Fountas, S. and Have, H. (2006) Agricultural robots—system analysis and economic feasibility. Precision agriculture Vol 7: 295-308

Rawlins, G. J. E. 1991. Foundations of Genetic Algorithms. Morgan Kaufman Publishers.

Chatterjee, S., Carrera, C. and Lynch, L. 1996. Genetic Algorithms and Traveling Salesman Problems, European Journal of Operational Research 93(3) 490-510.

Sørensen, C.G., Bak., T. and Jørgensen, R.N. 2004. Mission Planner for Agricultural Robotics AgEng 2004, Leuven, Belgium, 12-16 September 2004. Abstract in Book of Abstracts, p. 894-895. Full paper on CD. 8 pp.

Sørensen, C.G. and Bochtis, D.D. 2010. Conceptual model of fleet management in agriculture. Biosystems Engineering, Elsevier, 105, 41-50.

Stoll, A. 2003. Automatic operation planning for GPS-guided machinery. In: Precision Agriculture '03, Proceedings of the Fourth European Conference on Precision Agriculture, Berlin, 15–19 June, edited by J.V. Stafford & A. Werner, pp. 657-664.

Taïx, M., Souères, P., Frayssinet, H. and Cordesses, L. 2006. Path Planning for Complete Coverage with Agricultural Machines Field and Service Robotics. Springer Tracts in Advanced Robotics, edited by S. Yuta, H. Asama, S. Thrun, E. Prassler, and T. Tsubouchi. Springer Berlin / Heidelberg. Vol. 24, pp. 549-558.

Bektas, T. 2006. The multiple traveling salesman problem: an overview of formulations and solution procedures, Omega, 34(3) 209-219.

# Using linear programming for tactical planning in agriculture in the frame of the COIN project

Karel Charvat junior[1], Pavel Gnip[2], Karel Charvat[2]
[1]*Department of Econometrics, University of Economics, Prague, [2]WirelessInfo*
E-mail: xchak07@vse.cz

## Introduction

The main objective of the COIN Open Agriculture Service (OAS) is to improve collaboration within the farming sector. COIN OAS is developing innovative solutions based on currently used SaaS services platform Prefarm. This solution has certain limitation. Prefarm platform is currently used mainly for preparing recommendation for variable fertilisation. The solution currently supports only static collaboration relation. It is not possible to involve new users.

## Tactical production planning

The COIN tactical planning module is focused on introducing new methods oriented on tactical planning for the next seasons. The main goal of tactical planning is to recommend optimal production and land use for the next seasons to maximize expected profits. The goal for tactical planning is to recommend production for the next season, eventually also some vision for the next (two till five) seasons. The decision problem could be described in the way described below.

A farmer has a number of parcels with different properties available for his production. It is possible to grow different kinds of products on these lands.
On each of these lands it is possible to grow just one product using one of these two methods:

- Production without using variable fertilizer application.
- Production with variable fertilizer application.

Due to many reasons, for example crop rotation and field localisation, it is not possible to grow each of the products on each of the lands in the next season. In this approach, optimal production plan and profit for each scenario has to be computed. The key problem for the decision is the existence of data. In case of farmers, who were using precision agriculture in the past, it is possible to estimate many data values with quite high accuracy. Obviously most of the information does not exist in the farm information systems in the form requested by the model. The interoperability of data is missing; we need tools transforming data into the required form. To get values of some data it is necessary to create procedures, which will calculate the values from available data. However, some of the data values have probabilistic character. Selling prices of each product kind is very difficult to predict. The yields per hectare of some products and costs also bring some uncertainty into the decision problem. The approach which is used in case of parameters, which are difficult to predict, is creating several scenarios for factors, which are affecting values of these parameters. For generating scenarios a lot of external information is needed. Mainly in case of prices, we are probably not able to do without services of agricultural analysts. We will consider two approaches to the scenarios.

In the first approach the optimal production plan and profit for each scenario has to be computed. For each of these production plans, the value of profit in case of all other scenarios is expressed. Then a suitable method for choosing one of the production plans is used. The second approach is based on multiple criteria mathematical programming. In multiple criteria mathematical programming problems, there is more than one objective function. In case of our problem, the individual objective function represents profits in different scenarios. Solutions found by solving multiple criteria mathematical programming are called compromise solutions, because most of the problems don't have a solution providing the best values for all objective functions at the same time. For these purposes, it is necessary to use some of the available modelling systems and optimisation solvers. These tools have to be integrated with the current Prefarm SaaS information system using COIN MDIS. Purchasing commercial software belonging to these categories is usually very costly. In this moment we are using the modelling system PuLP with a modelling library created in Python.

The basic system architecture is divided into three layers

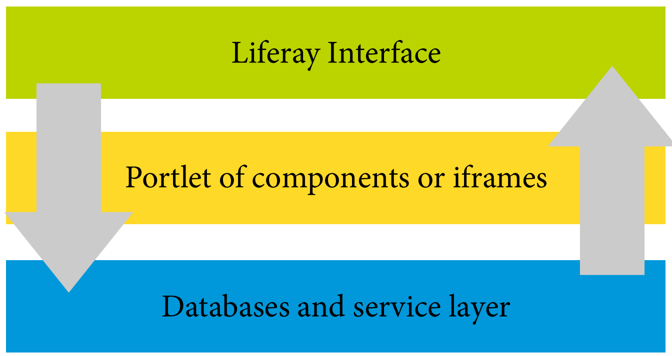



*Figure 1.*

Liferay interface guarantees communication with end users, but also authorisation and authentication. Layers of *portlets* or *iFrames* guarantee communication with single components or databases. Database and service layers are layers, where single components really interact. From component point of view architecture could be characterised by next components:

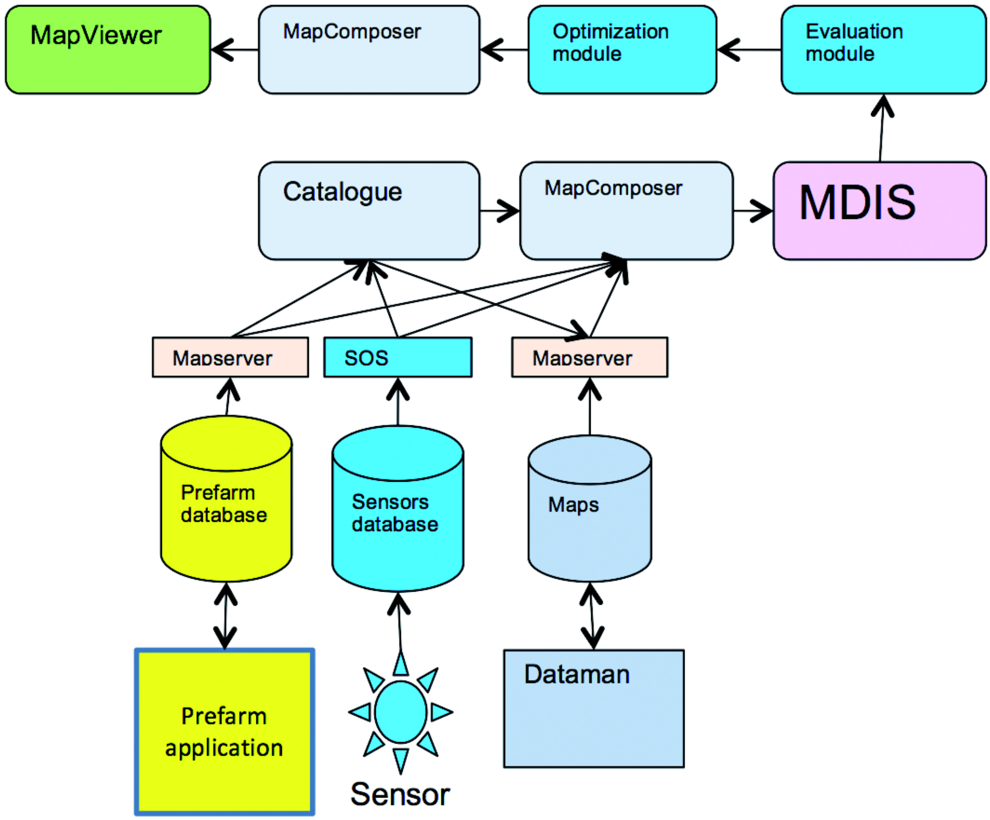



*Figure 2. System overview.*

The architecture allows discovering and selecting information sources from existing services - precision farming database, sensors and other existing databases and combining selected information into form of Web Feature Services (GML format – version of XML for Geographic Data. The geographical data are shared using services defined by Open Geospatial Consortium. For optimisation processes it is not necessary to use directly spatial representation of objects (areas, etc). For optimisations it is necessary to transform data into tabular form protecting only identifier of graphical object. This identifier is later used for merging optimisation results with graphical object for visualisation. In MDIS there is defined transformation model for transformation data into interoperable form. Own development is focused on Optimization module which is generating variants of production plan. This information is transferred into graphical form and visualised using visualisation client.

**Decision problem description**

A farmer has a number of lands with different properties available for his production. It is possible to grow different kinds of products on these lands.

On each of these lands it is possible to grow just one product using one of the following methods:

- standard production without using variable fertilizer application,
- standard production with variable fertilizer application,
- intensive production without using variable fertilizer application,
- intensive production with variable fertilizer application,
- extensive production without using variable fertilizer application.

Due to many reasons, for example crop rotation and field localisation, it is not possible to grow each of the products on each of the lands in next season. Also some of the methods of growing are not allowed on some lands. For formulation and solving the problem we need various input data.

Main factors affecting the profit are:

- prices at which the farmer will be able to sell his products,
- subsidies for plant production or different usage of land,
- yields / hectare of particular crops,
- costs including several items:
  - The costs of growing different crop kinds depending on land properties and land area. Lot of items like seed, fertilizer, operating costs and labor costs can be included into these costs.
  - Fixed costs connected with each kind of crop.
  - Fixed costs of using variable fertilizer application, which have to be included, if farmer uses variable application on any of his lands.

Some of the costs can be equal to zero in some cases, eventually they can significantly decrease during time. This includes for example fixed cost of variable fertilizer application, where farmer used this technology in the past.

We can suppose that use of variable fertilizer application in addition to decreasing cost will also influence crop yields and subsidies.

The goal of this problem is to create recommendation saying which crop should be sowed on each land and which way of their growing should be chosen, so as all constraints were satisfied and maximum value of profit was achieved.
To start we will suppose, that we are able to estimate all data values with sufficient accuracy. Further we will sketch one possibility, how to take in account stochastic character of same input data.

**Mathematical model**

Now we will describe the problem by mathematical model. The result of this description will be an integer linear programming model.

Variables definition

In our model formulation the following variables are used:

- $y_i$ – binary variable, which equals 1, if product *i* will be grown on at least one land, and 0 otherwise,
- $x_{ij}^{St}$ – binary variable, which tells, whether product *i* will be grown on land *j* using standard intensity,
- $x_{ij}^{StVA}$ – binary variable, which tells, whether product *i* will be grown on land *j* using standard intensity and variable fertilizer application,
- $x_{ij}^{In}$ – binary variable, which tells, whether product *i* will be intensively grown on land *j*,
- $x_{ij}^{InVA}$ – binary variable, which tells, whether product *i* will be intensively grown on land *j* using variable fertilizer application,
- $x_{ij}^{Ex}$ – binary variable, which tells, whether product *i* will be extensively grown on land *j*,
- $VA$ – binary variable which equals 1 if variable fertilizer application will be used, and 0 otherwise.

**Data description**

For constructing model and solving the problem following data are required:

- $m$ – number of crop kinds, from which farmer is choosing his production mix,
- $n$ – number of farmer's lands,
- $a_j$ – area of land *j* (in hectares),
- $c_{ij}$ – costs per one hectare, if product *i* is grown on land *j*,
- $C^{VA}$ – variable application fixed cost,
- $f_i^{prod}$ – product *i* fixed cost,
- $s_{ij}$ – subsidies per one hectare, if product *i* is grown on land *j*,
- $h_{ij}$ – expected yields of product (tons), when product *i* is grown on land *j*,
- $p_i$ – expected sell price of one ton of product *i*,
- $Q^{up}$ – maximum number of different crop kinds, which farmer plan to grow in next season,
- $\pi_i^{up}$ – maximum ratio of area sown by crop *i* to total area of all lands,

- $\pi_i^{down}$ – minimum ratio of area sown by crop *i* to total area of all lands,
- $w_{ij}$ – equals 1 if it is possible to grow product *i* on land *j*, and 0 otherwise.

Superscript $VA$ means, that the data item refers to variable rate of fertilizer application. Superscript *St* refers to standard intensity of growing, *Ex* to extensive growing and *In* to intensive growing. For example $c_{ij}^{StVA}$ denotes costs per one hectare, if product *i* is grown on land *j* with standard intensity and variable fertilizer application is used on land *j*

**Model formulation**

Using previously defined data and variables, we will create a model in the following form: maximize

$$z = \sum_{i=1}^{m}\sum_{j=1}^{n}\left[a_j\left(h_{ij}^{St}p_i^{St} + s_{ij}^{St} - c_{ij}^{St}\right)x_{ij}^{St} + a_j\left(h_{ij}^{StVa}p_i^{St} + s_{ij}^{St} - c_{ij}^{StVA}\right)x_{ij}^{StVA}\right]$$
$$+\sum_{i=1}^{m}\sum_{j=1}^{n}\left[a_j\left(h_{ij}^{In}p_i^{In} + s_{ij}^{In} - c_{ij}^{In}\right)x_{ij}^{In} + a_j\left(h_{ij}^{InVa}p_i^{In} + s_{ij}^{In} - c_{ij}^{InVA}\right)x_{ij}^{InVA}\right] \quad (1)$$
$$+\sum_{i=1}^{m}\sum_{j=1}^{n} a_j\left(h_{ij}^{Ex}p_i^{Ex} + s_{ij}^{Ex} - c_{ij}^{Ex}\right)x_{ij}^{Ex} - \sum_{i=1}^{m} f_i^{prod}\, y_i - (VA)C^{VA}$$

*subject to*

$$\sum_{i=1}^{m}\left(x_{ij}^{St} + x_{ij}^{StVA} + x_{ij}^{In} + x_{ij}^{InVA} + x_{ij}^{StVA}\right) \le 1 \qquad j = 1, \dots, n, \qquad (2)$$

$$\sum_{j=1}^{n}\left(x_{ij}^{St} + x_{ij}^{StVA} + x_{ij}^{In} + x_{ij}^{InVA} + x_{ij}^{StVA}\right) + \le ny_i \qquad i = 1, \dots, m, \qquad (3)$$

$$\sum_{j=1}^{n} a_j\left(x_{ij}^{St} + x_{ij}^{StVA} + x_{ij}^{In} + x_{ij}^{InVA} + x_{ij}^{StVA}\right) \le \pi_i^{up}\sum_{j=1}^{n} a_j \qquad i = 1, \dots, m, \qquad (4)$$

$$\sum_{j=1}^{n} a_j\left(x_{ij}^{St} + x_{ij}^{StVA} + x_{ij}^{In} + x_{ij}^{InVA} + x_{ij}^{StVA}\right) \ge \pi_i^{down}\sum_{j=1}^{n} a_j \qquad i = 1, \dots, m, \qquad (5)$$

$$\sum_{i=1}^{m} y_i \le Q^{up} \qquad (6)$$

$$x_{ij}^{St} + x_{ij}^{StVA} \le w_{ij}^{St} \qquad i = 1, \dots, m,\; j = 1, \dots, n, \qquad (7)$$

$$x_{ij}^{In} + x_{ij}^{StIn} \leq w_{ij}^{In} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (8)$$

$$x_{ij}^{Ex} \leq w_{ij}^{Ex} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (9)$$

$$\sum_{i=1}^{m} \left( x_{ij}^{StVA} + x_{ij}^{InVA} \right) \leq VA \qquad j = 1, \dots, n, \qquad (10)$$

$$y_i \in \{0,1\} \qquad i = 1, \dots, m, \qquad (11)$$

$$x_{ij}^{St} \in \{0,1\} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (12)$$

$$x_{ij}^{Ex} \in \{0,1\} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (13)$$

$$x_{ij}^{In} \in \{0,1\} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (14)$$

$$x_{ij}^{StVA} \in \{0,1\} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (15)$$

$$x_{ij}^{InVA} \in \{0,1\} \qquad i = 1, \dots, m \quad j = 1, \dots, n, \qquad (16)$$

$$VA \in \{0,1\}. \qquad (17)$$

Interpretation

Optimal production plan for the given input data is described by optimal solution of this problem. In terms of mathematical programming solving this problem means finding optimal values of all variables, which are used in this model.

In general, optimal solution has two basic features:

- All constraints are satisfied. This feature is called feasibility.
- In the set of feasible solution, there is no other solution with better value of objective function. In our case this means, that there is no other feasible solution promising higher value of profit.

The meaning of constraints in the model is as follows:

(1) Is objective function which expresses the value of profit.

(2) Ensures, that maximum one crop is grown on each land.

(3) Forces the variable $y_i$ to take value 1, if product $i$ will be grown on any land. This constraint is necessary to ensure that fixed cost of each product, which is grown, will be counted in the objective function (profit calculation). This variable is also necessary to counting different product kinds, which will be grown.

(4) Ensures respecting maximum ratio of area sown by each crop kind.

(5) Ensures respecting minimum ratio of area sown by each crop kind.

(6) Ensures respecting maximum number of different crop kinds

(7) Ensures, that each crop kind will be grown with standard intensity only on lands, where it can be grown this way.

(8) Ensures, that each crop kind will be grown intensively only on lands, where it can be grown this way.

(9) Ensures, that each crop kind will be grown extensively only on lands, where it can be grown this way.

(10) Together with objective function helps to ensure, that variable fertilizer application will be used only in case, it is convenient for the value of total profit.

(11) Ensure, that certain variables will be binary.

**Using scenarios in decision problem**

Especially in case of farmers, who were using precision agriculture in the past, it will be possible to estimate many data values with quite high accuracy. The most of information will not of course exist in the farm information system in form requested by model.
To get values of some data it is necessary to create procedures, which calculate the values from available data.
However some of the data values have probabilistic character. Especially selling prices of each product kind are very difficult to predict. The yields per hectare of some products and costs also bring some uncertainty into the optimization problem.

The approach which can be used in case of parameters, which are difficult to predict, is creating several scenarios for factors, which are affecting values of these parameters. For generating scenarios a lot of external information is needed.

Based on values of certain factors in individual scenarios, different sets of selling prices, yields per hectare and costs are computed. This means that different matrixes of yields per hectare $h_{ij}$, selling prices $p_i$ and costs $c_{ij}$ are used in each scenario.

According to the chosen approach, the result of the decision problem is one production plan, ore more different production plans, from which the farmer chose one. The approaches should be used in way which enables taking in account farmer's individual attitude to risk. In this time, we are working on implementation of these approaches into the optimization model
In following subchapters, we will describe two ways, which seem to be suitable for this purpose.

Choice from set of production plans

In this approach optimal production plan and profit for each scenario is computed. For each of these production plans, the values of profit in case of all other scenarios are expressed. The result is matrix containing values of expected profit for each combination of production plan and scenario. Than suitable method of choosing one of the production plans must be used. If it is possible to estimate probabilities of occurrence of each scenario, some of the methods of decision making under risk can be used. In the opposite case it is be necessary to use some method of decision making under uncertainty.

Multiple criteria mathematical programming using maximin principle

In multiple criteria mathematical programming problems, there is more then one objective function. In case of our problem, individual objective function represents profits in different scenarios.

Solutions found by solving multiple criteria mathematical programming are called compromise solutions, because the most of problems doesn't have solution providing the best values of all objective functions in the same time.

It the most of methods, some additional information like for example weights of each criteria is required from decision maker.

One of the methods of multiple criteria mathematical programming is finding solution according to the minimal component. This method is using maximin principle. In our case, the result is production plan, which in case of occurrence of the worst scenario for the production plan promises higher profit than any other production plan in combination with its worst case scenario.

This approach to scenarios is suitable for farmers with very negative attitude to risk.

**Software usable for solving of this decision problem**

The decision problem described above can hardly be solved without software, which enables creating and solving mathematical programming problems. For this purposes, it is necessary to use some of the available modelling systems and optimization solvers. These tools will be integrated with current Prefarm SaaS information system. Purchasing commercial software belonging to these categories is usually very costly.

In this moment we are using modelling system PuLP, which seems to be the most suitable tool for our purposes. PuLP is modelling system created in Python, which enables creating files in LP and MPS formats and which is able to interact with optimization solvers GLPK, CPL/CBC CPLEX and GUROBI. GLPK and CLP/CBC are open source tools, GUROBI and CPLEX are the top commercial products. From these possibilities we have chosen CLP/CBC, which has shown better computational performance then GLPK in our model. CLP/CBC is not as powerful tool as CPLEX, or GUROBI, but for our purposes the performance is satisfying and investment to the commercial solvers is not reasonable in this moment.

**Conclusion**

In this chapter we have described tactical production planning problem and presented mathematical model, which is supposed to help with solving the problem. However before using this approach in practice, it will be necessary to perform a lot of testing and development.

We suppose that the model formulation and description will further evolve as a result of testing on real data and adjusting according to the farmer's needs.

We also suppose that there won't be one universal problem formulation suitable for all potential users, so it will be necessary to add a possibility to customize the model.

Customization at basic level can be done quite simply by setting values of parameters, but in some situations adding, removing or changing constraints is the only way of taking in account individual requirements. For this reasons, mathematical programming is suitable approach, because it is quite flexible, and modified problems can be solved by the same algorithms as the original problem. Although mathematical programming is approach older than sixty years, the progressing in optimization algorithms and hardware in last ten years, and new possibilities of connecting models with data sources are dramatically increasing its practical usability.

**Acknowledgements**

The paper is prepared on the basis of outputs from the following project:
*COIN* - the solution was achieved with financial co-funding by the European Commission within the Seventh Framework Programme with registration number 216256 and name "COllaboration and INteroperability for networked enterprises"

**References**

CHARVÁT, KAREL, JR., GNIP, P., CHARVÁT, KAREL. COIN – Module for Tactical planning of agriculture production. Praha 11.07.2011 – 14.07.2011. In: EFITA/WCCA. Praha : Czech Centre for Science and Society, 2011, s. 166–174. ISBN 978-80-904830-0-2.

CHARVAT, K at all D1.2.3 Visions and recommendations for knowledge management, Report from FUTUREFARM project.

CHARVAT, K., GNIP, P. AND MAYER, W., 2009. FutureFarm vision, IPCA congress 2009, Wageningen.

CHARVAT, K., GNIP P.,VOHNOUT, P., CHARVAT, K. JR. VISION FOR A FARM OF TOMORROW, IST Africa 2011, Gaborone, Botswana

STUART, M., KEAN, A., MASON, A., O'SULLIVAN, M.,PHILPIS, A.: PuLP v 1.4.9 documentation [Online]. [cit. 2011-11-25]. Accessible from: http://packages.python.org/PuLP/

CBC home page [Online]. [cit. 2011-11-25] Accessible from : https://projects.coin-or.org/Cbc

FORREST, J.: CBC User Guide [Online]. [cit. 2011-11-25]. Accessible from: http://www.coin-or.org/Cbc/index.html

FIALA, P. 2008. Modely a metody rozhodování (Model and Methods of Decision Making). VŠE, Praha.

GNIP, P., CHARVAT, K. AND KROCAN, M., 2008. Analysis of External Drivers for Agriculture, IALD WCCA INFITA congress 2008 Tokyo.

JENSEN, P. A., BARD J. F. 2003. Operations Research - Models and Methods. John Wiley & Sons.

# Section IV
# Vision of the future

# What is the vision of the future?

Karel Charvat[1], Zbynek Krivanek[1], Pavel Gnip[2], Karel Charvat junior[2]
[1]*Czech Centre for Science and Society,* [2]*WirelessInfo*
E-mail: charvat@ccss.cz

What will be the future of rural ICT in the period untill 2030? What are the expected changes? How they will be driven? What are the ideas about the future ICT technologies? We will try to find some ideas on the basis of work of the Future Farm, Fi-Ware and COIN IP projects and authors experiences.

**Farm of Future vision**

This chapter is based on future farm studies. We define the vision in two periods, untill 2020 and until 2030. For every period the three-level approach is taken: firstly, the vision should address the drivers that will influence agriculture in next period; secondly, the vision of future farm is defined and finally, the vision addresses requirements on knowledge management systems. The vision development approach takes into account a number of existing roadmaps methodologies as well as other key reports and studies from other projects. The approach is based on information collected from studies, experts, but also from stakeholders.

<u>2020</u>

By 2020, the current food production methods will be unable to meet the worldwide food and energy demands of the growing world population. It will have the influence of the European farming sector. Food security will be a problem. Larger populations will start consuming at the level of present developed countries. We can't afford unsustainable production with the growing human population. Food demand increases requirements for better utilisation of research results and for new management methods. Combined with advanced bioprocess engineering the development of high performance crop plants is the key to this vision becoming reality. Crops will serve as factories for enzymes, amino acids, pharmaceuticals, polymers and fibres, and will be used as renewable industrial feedstock to produce bio-fuels, biopolymers and chemicals. Green biotechnology will be employed since conventional or smart breeding alone will probably not be able to provide the required increase in performance. It is anticipated that already by 2020, in addition to the then mature gasification technologies, the conversion of ligno-cellulosic biomass by enzymatic hydrolysis will be standard technology opening up access to large feedstock supplies for bioprocesses and the production of transport fuels. Research breakthrough in the second generation of bio-fuels derived from lignocellulosic material will make bio-fuels production more competitive and without using food material. The medium term influence will be to have food products with higher nutritional values, reduced chemical contamination and more advanced traceability systems. In this period, the average age of populations will continuously grow. This generation will be more active than previous senior generations and will require specific diets. It is expected, that the percentage of ethnic groups in Europe and US will increase. They will have an influence on specific requirements of agriculture and food production. Investments in high-value crops, high quality food products and new technologies in crop production will be the case in the medium term. In the medium term the need for more food and for energy from crops due to the high prices of fuel will also boost R&D in Europe and worldwide.

In the medium term, climate change could be a benefit for agriculture in higher latitudes by enabling the introduction of new crop varieties, increasing yields and expanding areas of land under cultivation. In certain lower altitudes, it will be probably necessary to focus on varieties that are more resistant to draught.

The new standards and regulations will require strong cooperation with WTO, but also with food producers and markets. New standards will combine market requirements for food, with requirements for energy, but also environmental protection. Currently, it is difficult to judge the result of the CAP reform. There is a demand, that subsidies should be made available for specific innovative investments and experiments aimed at increasing environmental performance of the farm, such as resource-saving technology (energy, water, etc.), renewable energy technologies, low emission stable systems, etc. With the CAP reform, national strategy will also be changed. It is expected that there will be a shift from direct payment to supporting environmentally friendly production and food safety.

The success of all measures to reduce the environmental problems and mainly to reduce the climatic changes effects will be revealed. Any success or failure will affect directly the farming practices and management of the farms leading to more agro-environmental measures taken by the states. By 2020 and further on, the effects on the climatic change will be clear and it is believed that more worldwide treaties will be enforced. The first environment valuation programmes will start and failures will lead to restructuring and changes; as the environment is not a local or regional or national problem, this question has to be taken on a higher platform and must be discussed on a worldwide level. The economical instruments will stronger reflect the issues including healthy production, food safety and environment protection. Agriculture will largely adopt new collaborative models, which will support not only sharing of resources, but also of knowledge. In the medium term, partnership agreements will be more widely used and will be specialized in different sectors, such as the production of energy crops, where local industries will produce oil and bio-fuels as well as in other agricultural sectors mainly for industrial crops.

There will be large diversification of farm and we can characterise three main types of farms: multifunctional and large scale industrial farms and farms with focus on specific production.

***Multifunctional farms***

The adaption to environmentally friendly production will continue, but this will strongly depend on valuation of ecological parameters. In order to guarantee the sustainability of multifunctional farms, it will be necessary, on the basis of public dialogue, to pay their non production role. The main paradigm will not be an extensive production with focus on greenhouse effect. The main task will be to produce food in a sustainable way which meets the consumer's demands. The target will be a balanced use of agricultural products (between food & biofeed stocks, bio-energies). Farms will produce its own fuels and will decrease their dependence on energy sources. Important parts will be of multifunctional farm tourism and consumptions of products directly on the farm.

***Large scale industrial farms***

Large scale industrial farms will guarantee food security. Since the subsidies for this type of farm will be dramatically reduced, it will be necessary to discus all standards of production, which will be required from this type of farm. It is clear, that there will be needs for such standards like water protection or $CO_2$ production. Tt is necessary to take into consideration, that too strong regulations applied only in Europe could destroy this important group of farms. In any case, there will be growing demand from consumers for guaranteed food quality. It can be expected, that in this period the energy cost will grow and there will be open questions about energy production.

***Farms with focus on specific production***

There will be two types of farms where specific production will grow. The first group of farms will be focused on delivery of high quality food, which will be produced by environmentally friendly methods e.g. vegetables and fruits or foods for people with specific cultural or deictic requirements. This group of farms with respect to their specificity will converge to industrial farms. This group of farms will have extremely high requirements on new results of research and will be a basis for knowledge based economy. There will be also extremely high requirements on ICT, mainly on monitoring, but also on very precise production management. Important will be marketing and real time delivery.

Second group of farms will be focused on consumption of products on the farm. So the requirements will be mainly management of quality of production.

<u>2030</u>

Global food production must grow by 50 per cent by 2030 to meet the increasing demand of the growing population. Massive efforts are required to maintain fertile cropland. Demand for animal protein may increase, triggering massive investments into genetically modified food, aquaculture, and stem cells for meat production without growing the animal. Seawater agriculture on desert coastlines could produce bio-fuels, pulp for the paper industry, and food for humans and animal bio-fuels, while absorbing carbon and reducing the drain on fresh water. In the biggest part of Europe urbanization and land abandonment will result in more concentrated production in the urbanised areas and reduction of the production in less favourite areas. In a long term strategy, it is necessary to solve the impacts of raising energy prices like increasing field areas or increasing production of bio-energy from agriculture.

The benefits of climatic changes in a long-term are less certain. Particularly in lower areas, droughts and desertification will create significant social challenges in some of the world's poorest economies. Areas including Siberia, Scandinavia and Canada will profit from global warming.

The key issue will be food quality. Long term influence will be on the intensive use of traceability systems in the food supply chain and this will be compulsory to all farmers producing food and to the retailers. The focus on aging population and health will be a major requirement for food production. There will be an important shift in composition of production in direction towards vegetables, fruits, fish, chickens, etc. The percentage of ethnic groups will grow further. Around 2030, ethnic groups could comprise a major portion of the European population. This will influence food and agriculture production. There will be complete change of economical instruments, which will influence production. The main focus will be on removing distortion of the market and to support healthy and environmental friendly production and support worldwide food security. In the long term, partnership agreements will be more ‘mainstream’, where local industries will be closely connected to the region and farmers will directly sell their products to them securing prices disposal of their production. Agricultural production will be horizontally and vertically integrated.

Biotechnology will be an important pillar of Europe’s economy by 2030, indispensable to sustainable economic growth, employment, energy supply and to maintaining the standard of living. It will be increasingly used in labour-intensive sectors, e.g. industrial processing, pharmaceuticals, agriculture and food. The increasing demand for energy will keep prices high and support the demand for bio-fuels. Therefore, the investment interest will continue. In this period, if the oil reserves estimation is correct, it is expected that some oil reserves will be depleted and this will worsen the supply of energy. Thus energy prices will increase and investment in renewable energy sources and biomass produced by the farms will be enhanced. New dimension of farms will also take place, such as pharmaceutical crops, industrial crops as well as high quality and safety food. Research in agriculture for new and advanced agricultural commodities will be needed to keep raw material supply at low cost. The trends of the previous period will be maintained and funds will be available for research in the sector. If the climatic change scenarios are verified, strict measures have to be adopted. This will be an important driver for farms to change practices and management to more environmental friendly direction. A worldwide valuation of ecological performances with rules like “who has to pay, how much and for whom”, taking into consideration the impact of environmental caretaking for local, regional, national, continental or worldwide influence will take place.

Efforts to enhance the environmental performance of agriculture will play an important role. Social and political pressures for increased environmental standards are expected. Resulting policy tools, whether positive (subsidy based) or negative (penalty based), if substantial enough, could play a major role in shaping future agriculture. On the other hand, the cost of dealing comprehensively with the above set of environmental issues would be many times greater than the public funds currently available through the main policy programmes. It may be that public funds continue to play a marginal role in protecting or enhancing the rural environment. No dramatic increase in environmental regulation governing agriculture is expected.

The important question, which is open until now, is if agriculture production will be subsidised, and what agriculture production will be subsidised. The final decision will depend on discussion of member countries. The final decision is not clear, but it is expected, that financial incentives should be available for farmers who produce according to sustainability criteria, because they bear additional costs (compared with unsustainable production) which will not directly be paid by consumers. In the transition period until 2030, the incentive should increase with the degree of implementation of such criteria.
It could be expected, that due the requirements on quality of production and also on environmental friendly production on one side and on the other side in increasing demand on high quality food, vegetables and fruits and also growing demand on special production, the convergence of two types of farms in two main groups will continue. In the final stage we will have two basic types of farms:

- Multifunctional farms;
- Industrial farms with focus on high efficiency and high quality of production.

***Multifunctional farms***

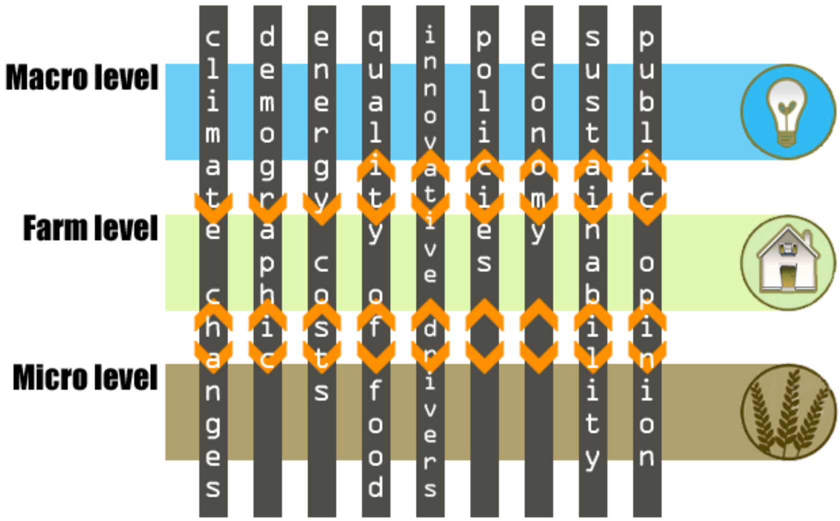
The focus will be on efficient agriculture from an environmental and socioeconomic point of view. The role of farm will depend on public dialogue and valuation of non production goods. The focus will be not only on production, but also on landscape and cultural heritage. Landscape is one of the most commonly cited term of the multifunctional characteristics of the agriculture sector. However, the impact of agriculture on landscape has not always been positive. This will be changed and agriculture will play an important role in forming of landscape. Rural viability via agricultural employment will be one of the multifunctional outputs of agriculture. A link between agricultural employment and agricultural production is expected. But part-time farming, diversification of income sources of the farm household, and the development of non-agricultural activities in rural areas, will mean that agricultural employment and rural development are much less inter-dependent than in the past. There will be also other value as historic buildings and associated cultural heritage values in rural areas.

***Industrial farms with focus on high efficiency and high quality production***

The main requirements will be to produce enough food and energy in a sustainable way which meets the consumer's demands. The farms will have to respect strong regulations and they will need to have an effective production. There will be necessary to find a balance between restrictive methods like regulation and standards on one side and economic stimulus like subsidies on the other side. The requirements to produce high amount, high quality and environmentally friendly will require new methods of management and also application of new scientific results. Research in agriculture for new and advanced agricultural commodities will be needed to keep raw material supply at low cost. The trends of the previous period will be maintained and funds will be available for research in the sector.

Demands for Future Farming knowledge system

It is necessary to take into account previous analysis for suggestion of future knowledge management system functionalities and interrelation. The basic principles of interrelation could be expressed by the following image.




By studying of this schema we can expect the following transfer of knowledge.

<u>Climate changes</u>

Macro level

There will be a need for a long time modelling of influences of global climatic changes on production in single regions. It is not only a focus on a selection of best crops and their varieties, but also analysis of new pests or insects. It will be important to analyse and predict the needs for irrigation. In short term, it will be necessary to predict extreme weather events.

Farm level

Prediction model for climatic changes in middle term period (minimally for one season, but better for three or five seasons) will be important aspect for changing of strategy on the farm level. There is no possibility to change arable production during the season. It is better to continuously move to a new production.

Micro level

With changing climatic conditions and with higher probability of extreme local events the importance of local monitoring and local weather forecast for short (weeks, month) time period will grow. Usee of new sensor technologies and local forecasts systems will play an important role.

<u>Demographic</u>

Macro level

On macro level it is necessary to forecast the amount of food that will be necessary worldwide and regionally only. This forecast should enable in middle term perspective from one till five years for farmers adapt their production for next season, but also to provide some changes for next years. It is also necessary to provide short time forecast of yield during season to be possible for farmers to optimise their marketing strategy in common season.

Farm level

The seasonal forecast of yield has to be used on farm level for optimising other costs of production during the season. Middle term forecast has to be used for optimisation of farm production (selection of crops, varieties, crop production focus). Other important aspect could be to decrease the demand on labour resource by better farm management and logistic with robotic introduction and size of robots implementation.

Micro level

The short and middle term information about yield demand will require adding used model of precision farming. It is not possible to implement one model. Tt will be necessary to offer different possibilities, like maximise yield, to keep good yield quality, to minimise costs (to keep soil sustainable development), to optimise farm profit for future farm development, etc. Robots' implementation helps to solve problems with time needed for data collection, accuracy, objectivity of data collection and labour resources.

Energy cost

Macro level

From the development of costs of energy in the last year, it seems that costs of energy are not predictable. But in long term we can expect that the cost of energy will grow. There will be a need for prediction of the development of energy costs. The cost of energy will be important from two different reasons:

- Energy costs will influence costs of production.
- Energy costs will be important for decision.

Farm level

The new knowledge management system on farm level has to be focus on three aspects:

- How to decrease cost consumption, which will be focused on farm management strategy and improving of logistic.
- To use such machinery that will guarantee lower energy consumption for the same amount of work.
- Decision about optimal crop composition.

Micro level

Development of energy cost will influence farm management on field level in next aspects:

- Optimisation of energy consumption using robotic technology.
- Use PF technologies in new direction to decrease energy costs as for example precision seeding or precision crop protection

- With potential energy production will be necessary to change the PF models like in previous case.

New demands on foo quality

Macro level
There will be necessary to collect and analyse information of consumer behaviour and consumer requirements in the direction of food quality or specific diet requirements. It will be necessary to qualify and also quantify the willingness of population to pay for certain quality of production or for certain products. Currently, for major population costs of production are the main criterion. This criterion is in relation with economical situation of single citizens and globally with economical situation in countries.

Farm level
The main decision on farm level will be whether it will be better to orient farm production of lower amount of high quality of yield with higher prizes or higher amount of lower quality of yield with lower prizes.

Micro level
The new food quality demands will have on field level next influences on knowledge management systems:

- Using new tools of traceability to give evidence about quality of production.
- To use new PF algorithms to guarantee higher quality of production.
- To use new technologies like robotics to guarantee more precise fertilisation and crop protection.

Innovative drivers

Macro level
From innovation point of view, it will be necessary to offer information to farmers about new crops that could be used for concrete climatic or geographic conditions, about hey resistance productivity etc. It will be important to give farmers an access to new possibilities of using different crops for energy production, different bio products and use of different crops for example in pharmacy.

Farm level
The information about new products has to be used for optimal selection of sorts in relation with geographical and climatic production. The information about needs for different crops for non food production and also their prizes has to be used for decision, which will be orientated of farm production.

Micro level
On micro level, PF methods should be able to adapt for specific requirements of new sorts.
It is important to introduce new methods of traceability, which will guarantee usability of crops for specific bio production and in pharmacy. In order to develop new technologies like

robotics, which will guarantee better fertilisation and protection, which will be required by new crops or for new kind of production.

Development of sustainable agriculture

Macro level
On macro level, there is a need to guarantee access of farmers to information about possible payment and the possibility of environmental valuation. It is important to take into account that if farmers will play role in environment protection, they will need to have a profit from it. It is necessary to offer them concrete values, which they will obtain using certain methods of production.

Farm level
The knowledge management system has to support the following tasks:

- To define optimal methods of production (selection of crops, crop rotation, etc.) to guarantee long term sustainability of the production and protection of quality of soil as the mean of production.
- To support decision. If the farmer will be more oriented on classical intensive production or on the methods of environment protection of production. It has to be done on the concrete level of environmental valuation.

Micro level
The knowledge management on micro level has to cover the following tasks:

- Monitoring of soil quality.
- Precision farming methods, which will guarantee protection of environment.
- Traceability tools giving evidence of used methods.

Policies

Macro level
The knowledge about changes in policies, standards, changes in subsidies has to be transferred into such forms that could be easily acceptable by farm management systems. Policies, subsidies, standards are usually not possible to change on lower level, but they represent some level of limitation. It is necessary to transfer this knowledge into decision supporting systems on farm level as a limitation. There is necessary to support subsidies management and controlling system.

Farm level
It is necessary to include limitation given by policies as components or limitations for decision support system on farm level. It is necessary to analyse and collect information that is required as evidence from governmental bodies.

Micro level
The knowledge management on field level has to guarantee traceability and all monitoring of all parameters that will be required by legislation.

Economy

Macro level
It is necessary to collect knowledge about costs of inputs on market and also about prizes of products on the market to give chance to farmers to adapt their behaviour. It is necessary to provide integration of information in vertical and horizontal agri-food chains and also integration of knowledge among farmers, advisory and service organisations using common workspace to guarantee effective cooperation.

Farm level
There is a need for decision support system, which will be able, on the basis of economical information from market, to analyse possible changes in production, selection of suppliers and to select to whom to sell the production. It is necessary to give the possibility to share a part of the information in horizontal and vertical chain.

Micro level
The PF tools have to guarantee effective exchange of information with suppliers.

Public opinion

Macro level
Public opinion could have influence on consumer behaviour, on the costs of products on the market, but also could influence policies (for example global warming problem). It is necessary to analyse different public opinion campaigns and analyse their possible influence on the market.

Farm level
On farm level, it is necessary to include as one from inputs of DDS possible influences of public opinion campaigns on market.
Good traceability on all farm levels has to be supported.

Micro level
Support for full traceability.

Decision process problem

In order to provide optimal decision, the system needs to have an access to as much knowledge from global level as possible (the different macro knowledge was mentioned before). Any missing knowledge could have influence on the final decision. There exits many limitations, but also many freedoms for decision. It is necessary to mentioned one risk. If all farmers will use the same input knowledge and the same deterministic algorithm, the usage of such decision could lead to distortion of the market. There exist two possible options, which could guarantee non uniform decision:

- To use suboptimal variants.

- To use non deterministic methods for decision.

**ICT vision**

In Chapter III the new emergency concept in ICT was introduced – Future Internet [3]. Such concept was already elaborated and tested in c@r project [4], [5].

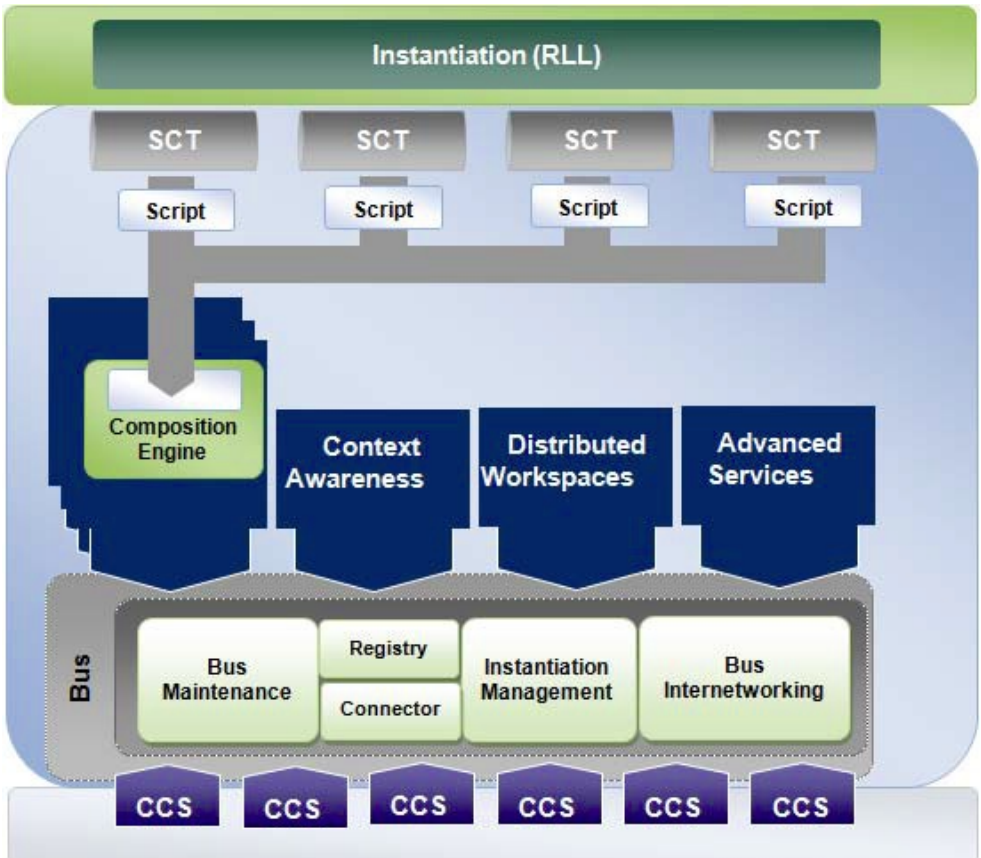



*Figure 1. The C@R architecture from Christian Merz et al. / Reference Architecture for Collaborative Working Environments [5].*

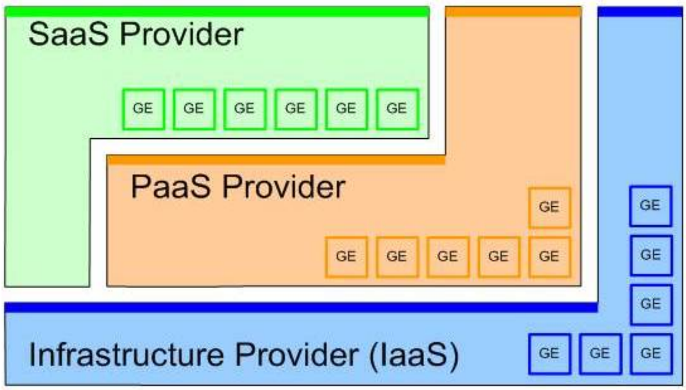
The architecture was described as:

- CCS – Collaborative Core Services implemented as reusable software
- SCT – Software Collaboration Tools
- OC - Orchestration Capabilities
- LLA – Living Lab Applications cover end user interactions

Currently, this concept is extended by more projects e.g. FI-WARE and COIN IP. FI-WARE is focused on architecture design, COIN IP on enterprise collaboration and interoperability services [6].

According to FI-WARE future internet includes:

- Cloud computing - There are three basic models of Cloud Computing
    - Infrastructure as a Service (IaaS): rating compute resources;
    - Platform as a Service (PaaS): use of a specific programming model and a standard set of technologies;
    - Software as a Service (SaaS): end-users are renting stacked Final Application; [3]




*The XaaS stacked model [3].*

- Internet of things – understand interconnection of physical object, living organism, person or concept interesting from the perspective of an application. [3].
- Data and content management - of data at large scale is going to be cornerstone in the development of intelligent, customized, personalized, context-aware and enriched application and services.
- Application Service Ecosystem includes reusable and commonly shared functional building blocks serving a multiplicity of usage areas across various sectors.
- Security - has to guarantee that the personal information provided by users will be processed in accordance with the user rights and requirements.
- Interface to networks and devices - has to guarantee access a variety of physical networks, contents, services, and information provided by a broad range of networks.

This technological concept is now in the stage of design, but is expected, that this concept will dominate in the future.
COIN IP is looking for a concept of Future Internet from the perspective of business collaboration and business interoperability. It defines three types of services:

- Enterprise Collaboration Services supporting collaborative processes in supply chains, collaborative networks or business ecosystems;
- Enterprise Interoperability Services reducing incompatibilities among enterprises;
- Service Platform as integrating services for enterprise collaboration and enterprise interoperability based on semantically-enable Service Oriented Architectures (SSOA) [4], [5].

COIN IP provided a study of business models. It is working with the concept of SaaS-U (Service Utility) as specific business model. It expects that utility paradigm will be used for offering services. Model expects that use-value and exchange-value are not identical and that in the future mainly added value services will be sold. The analysis also mentioned the possibility of domination of market by few organizations, like in real utilities.
The vision of cloud probably will become a reality to overcome the problems of monopoles, and also opening possibilities for small and medium developers. Two aspects are important:

1. Interoperability and service oriented architecture, which allows easy replacement of one components or service by other. This concept is already currently broadly used in geographic information systems.
2. Support for large scale utilisation of Open Sources [7], [8]. Currently Open Source generates business for companies, which customise solution into final applications. Such web based solutions could generate profit for SMEs. For the future growth of the Open Source market, it is necessary to look on such models, which will attract producers of software to publish their systems as Open Source. It is also important to find for them sustainable business model. The goal is in principle not only to build communities, but also to open the chance to generate profit for primary producers of such systems and components.

We believe that the future solution will be in some way convergence of current methods introduced by Future Internet, mainly interoperable cloud computing, with principles of Open Sources.

**References**

Karel Charvat, Pavel Gnip Future Farm for Farm of Future, IST Africa 2011, Gabarone Botswana

Karel Charvát, Pavel Gnip, Walter Mayer FutureFarm vision, AGRIS on-line Papers in Economics and Informatics

https://forge.fi-ware.eu/plugins/mediawiki/wiki/fiware/index.php/FI-WARE_Product_Vision

www.c-rural.eu

Hans Schaffers, Javier García Guzmán, Mariano Navarro de la Cruz, Christian Merz at all Living Labs for Rural Development Results from the C@R Integrated Project, Published by: TRAGSA and FAO, ISBN-13: 978-84-693-0040-4

http://www.coin-ip.eu/

Karel Charvát Future of Open Source systems, Geoinformatics FCE CTU 2009

Karel CHARVAT, Zbynek KRIVANEK Dual Licensing Policy – New Way for Software Development, IST Africa 2010

# The need for continuous foresight in agriculture and food research – a case study based on Teagasc 2030


Patrick.Crehan

*CKA*

E-mail: Patrick.Crehan@cka.be


## Background

Teagasc means 'teaching' or 'instruction' in Gaelic. It is the name of the food and agricultural research, education and advisory body in Ireland. By 2006, fundamental changes happening to the Common Agricultural Policy in Europe were already being felt throughout the Irish agri-food sector. New and emerging issues were gaining importance and looked likely to have an impact on the sector. It was necessary to ask how Teagasc could maintain its relevance to clients and stakeholders as it moved ahead. The study builds upon previous foresight exercises and long-term strategic studies undertaken in Ireland and the EU.

## Employing Knowledge for Developing a Positive Vision and Creating Opportunities

Teagasc 2030 was designed to establish a broadly-shared vision of what the Irish agri-food and rural economy would look like in 2030 and a vision of what Teagasc could become as the leading science-based knowledge organisation in the sector. It set out to develop the strategic capabilities of Teagasc, improve its ability to provide proactive leadership on complex issues, identify strategies and mechanisms to maximize the impact of its knowledge generation and procurement, technology transfer and education activities through innovation support and to develop an internal culture of continuous renewal.

The ***Steering Committee (SC)*** included key Teagasc managers, high-level representatives from relevant organisations, such as the university system and the Environmental Protection Agency, influential business leaders from both the farming and food sectors, as well as international experts. The members of the SC played a decisive role in the process in that they were fully engaged and provided constructive input each time the group convened. The ***Working Group (WG)***, consisting of Teagasc employees aided by two international consultants, was responsible for the detailed planning and execution of the exercise. The ***Foresight Panel (FP)*** consisted of experts from Teagasc, representatives of the farming and food sectors, as well as experts from the research community, including a commercial research service provider. FP members participated in and contributed to workshops and other activities organized by the WG.

Early consultations with the SC reinforced the need for a structural approach that went beyond the traditional sectoral view. The SC emphasized the need for new strategic capabilities that would enable the organisation to operate in a rapidly changing context. One of the first tasks of the WG was to review foresight exercises on food, agriculture and the rural economy that had been conducted previously, whether in Ireland or around the world, start a discussion on the scope of the exercise and get agreement on the nature of the results it should provide. The first observation of the WG was that previous foresight exercises on food, agriculture and the rural economy tended to focus on problems related to commodity markets and the Common

Agricultural Policy (CAP) system of payments. It was resolved at an early stage that Teagasc 2030 would have to do more than this by identifying how knowledge could help create opportunities for young people in the sector and by developing a positive and realistic vision of an innovation-led rural economy.

The work itself was organized in two phases. A ***Divergent Phase***, where the main purpose was to study issues relating to the organisation, the sector and the broader economy in a creative and exploratory fashion, brought in outside knowledge and expertise, as well as relevant case-studies from abroad. The second ***Convergent Phase*** focused on choices to be made about desired outcomes, long-term visions for the future of Teagasc and the context in which it would operate, as well as the practical immediate steps to be taken on the basis of an action plan. Just before the end of the Divergent Phase a ***Radical Thinkers Workshop*** was organized to challenge peoples' thinking and try to overcome any remaining inertia or scepticism as regards new ideas and the necessity for change.

**The Divergent Phase**

This consisted of paper writing on a number of key topics that provided important background to the members of the Foresight Panel. The papers were especially important as they allowed people who are not experts in a domain to get an overview of what is happening. The real action, however, was in a series of four workshops (WS).

Turning Towards a Knowledge Based Bio-Economy

***WS1*** consisted of a ***scoping and profiling*** activity to determine the boundaries of the Teagasc 2030 exercise and to verify that the FP included a sufficiently broad range of actors. Important discussions arose concerning how agriculture and food related to the use of land in Ireland, the relationship between this and both the rural and national economy, how both the theatre and the actors might be changing, and how there was a need to revisit ideas of who the typical Teagasc client was, is now or would be in the future. The immediate output of this workshop was strongly criticized by the SC as not being radical enough. It was thought too traditional or sentimental in its attachment to 'land'. The modern reality consists of urban agriculture, gardens on the sides of buildings, forests, marine and lake habitats, greenhouses and bio-reactors, as well as a food industry that has long outgrown a dependence on local production and that in some sectors relies almost entirely on imports for raw material inputs. This workshop started a process of reflection that lasted until the end of the exercise.

The feedback of the SC on the results of this first workshop was very important. Its intervention ensured that some of the issues addressed in the workshop did not conclude pre-maturely, but stayed open and continued to be debated for the best part of a year. New ideas need time to mature. The workshop started a process whereby traditional and ultimately limited thinking about farming and the rural economy were replaced with entirely new thinking about the knowledge-based bio-economy or ***KBBE***.

***WS2*** focused on trying to understand relevant ***drivers*** of change, the factors shaping the future of Teagasc and the environment in which it operates. The focus was on identifying the drivers and the ***impacts*** that they could have on the economy in 2030. The discussion included references to ***trends*** and ***trend breaks***. The exercise was intended to help people develop their 'intuition' about 2030.

***WS3*** focused on strategic ***issues*** and started the process of formulating the opportunities and challenges that the various sectors and stakeholders would face in 2030. By this stage the concept of the '***Sustainable KBBE***' had started to come into focus.

***WS4*** was about developing ***scenarios*** to further develop thinking ***about*** the 'Sustainable KBBE' in 2030, to further explore and define the issues and challenges, and to identify the ***big questions,*** whose answers would impact on the structures and programmes of Teagasc going forward.

A ***Radical Thinkers Workshop*** was timed to take place between WS3 and WS4 to provide new ideas to the ongoing foresight process. This consisted of a series of talks followed by discussions, involving speakers from a variety of areas who were capable of presenting challenging views on relevant topics. It involved scientists, geographers, venture capitalists and policy makers. For some participants it was an opportunity to hear for the first time about a renewable chemicals industry based on crops grown for their chemistry rather than for food, feed or fibre. For others, it was an opportunity to hear what foreign experts think. A venture capitalist provided his vision of where important opportunities for investment would arise in future. A Danish speaker raised important questions about the organisation of research and innovation when he explained that, while Denmark performs about 1% of all global research, Danish industry requires access to the other 99% of global research if it is to achieve or maintain global competitiveness.

**The Convergent Phase**

This consisted of a series of three workshops involving the FP and had to provide an actionable plan for the transformation of Teagasc. Such a plan would require the commitment of Teagasc senior managers. It had to be something they would own and act upon. To make sure that they were adequately prepared, a series of internal meetings was arranged involving senior managers and representatives of the WG to help them understand the implications of the exercise, identify the main axes of change for the organisation and anticipate the detailed requirements of the last workshop. Although the foresight workshops were usually animated by members of the WG with help from the external consultants, the goal was for key sessions of the final workshop to be led by members of senior management with support from the WG. At the same time, an internal dissemination or consultation process took place involving all parts of the organisation. The goal was to explain what was happening and gather feedback on the changes required for moving forward. An external consultation process separately involved farming and food industry industry representatives. It too explained the ideas that were emerging. It gathered feedback and inputs from Teagasc clients as inputs to the final stages of the foresight exercise.

***WS5*** was dedicated to the development of ***scenarios*** about the ***Sustainable KBBE***. In particular, the goal was to develop more specific thinking about the role of knowledge, learning, research, innovation, training and advice in the sector in 2030.

***WS6*** was used to finalize the ***scenarios*** and flesh out a ***vision*** for the sector in 2030 along with an identification of its knowledge requirements and the role that Teagasc would occupy in the system.

***WS7*** was devoted to the issue of organizational transformation and the ***directions of change*** for Teagasc. The senior **management** meetings played a significant role in determining the structure of this last meeting. Based on their discussions it was decided to focus on transformation under the major headings of ***leadership***, ***partnership*** and ***governance***.

The issue of leadership originally emerged in meetings of the SC and was echoed in discussions with industrial stakeholders. Leadership gaps emerged on long-term scientific and technological issues not only for small and medium-sized enterprises, but for larger companies as well.

**The Vision of a Sustainable Bio-Economy**

One of the most important results was the development of a vision for the Agri-Food and Rural Economy in 2030 as a '***knowledge intensive, innovative, internationally competitive and market-led bio-economy***'. This helped to place the sector at the centre of something big and positive, with significant opportunities that would play a role not only in the rural economy, but also in the general knowledge economy, via its contribution to climate change, energy security, sustainability and the transition to a post-petroleum era.

Recognizing that countries with excellence in agriculture have opportunities for moving up the value-chain by selling not only their products but their know-how, the final report speculated about a time when the most important export of the dairy sector in Ireland might no longer be its milk, cheese, yoghurt and functional foods, but its management expertise and its technical knowledge about the organisation of competitive dairy production systems.

<u>The Four Pillars of the KBBE</u>

From an Irish perspective it made sense to complete this vision by distinguishing '***four pillars of the KBBE***':

- ***Food Production and Processing***, which mainly represents mature industries where competition is relentless and global, where competitiveness often relies on efficiencies of scale, automation and process technologies, as well as scientific management and competitive sourcing.
- ***Value-Added Food Processing***, which includes advanced food processing and food service, functional foods, as well as food-additives and ingredients, bio-actives, nutraceuticals and cosmaceuticals. This sector is fast moving and innovative. There is continuous adoption and improvement of technologies for production, processing, distribution and preparation. Supply chains are constantly changing and considerable attention is given to intangibles such as patents, brands, provenance and traceability.
- ***Agri-Environmental Goods and Services*** includes food-safety and traceability, animal welfare, energy security, climate, clean air and water, fertile soils, bio-diversity, areas of public amenity, natural beauty and those of importance for cultural heritage. Although these are normally treated as spin-offs from other activities based on multi-functionality, they are given a separate identity in recognition of the overall role they will play in the quality of life of citizens, in energy and climate security as well as in the overall sustainability of society and the economy.
- ***Energy and Bio-Processing*** includes the production of feedstock for bio-fuels and bio-polymers. This sector makes substantial investments in harnessing knowledge. It places great importance on knowledge as a factor of production. It corresponds to new and

emerging areas of science and to entire new markets. It is characterized by a high level of risk and provides opportunities for government support to ***lead markets***. This sector is where high-value-added and commodity sectors of the future are being created.

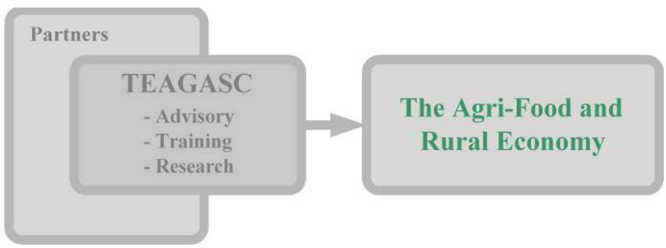
***Demographics Facilitating Change***

A key observation concerning the future of Irish agriculture was the observation that approximately 40% of farmers in Ireland would retire in the next 10 years and that almost all farms would change hands at least once by 2030. This pointed to an opportunity to use the unavoidable dynamic of retirement and property transfer to restructure the farming sector so that land as a natural resource could make the greatest possible contribution to the economy. This would include enabling successful farmers to increase the area they manage and less successful ones to move on perhaps using models based on leasing.

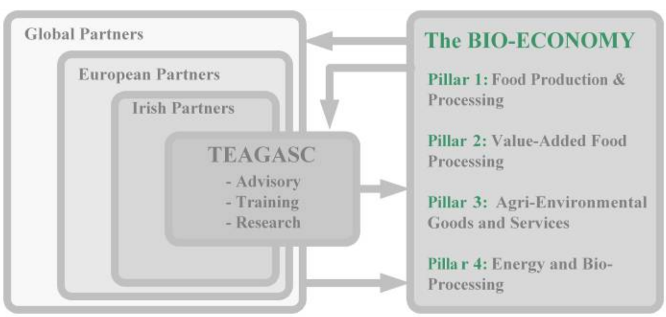
Discussions arose about ‘future farmers’ and ‘foresight farmers’. It is possible that the land transfers that will happen in the coming years will give rise to a younger, better educated and more international generation of farmers. Armed with agricultural MBAs, or degrees in bio-technology, many will approach farming as a business more than a family tradition or vocation. Their approach would be less sentimental and more scientific-entrepreneurial. Such farmers represent very different clients for Teagasc than those it has served before.

***Leadership, Partnership and Governance***

One of the most important currents of debate throughout this foresight exercise concerned the traditional push-approach to technology transfer, the so-called ‘linear model’. The old approach was summarized as follows




whereas Teagasc in 2030 would need to focus on innovation support that would resemble something more like this:



One challenge that emerged was the need to become more demand-led as an organisation. Another challenge emerged from the recognition that no organisation can meet all of its research or knowledge needs internally and that an increasing share of these would need to be sourced outside. This is something that traditional research organisations are not used to doing, and, in future, they will need to engage both private and public service providers, as well as cooperate with international knowledge networks.

The vision that emerged for Teagasc as an organisation in 2030 was that of an organisation suffused with a culture of support for innovation by its clients, capable of:

- providing ***leadership*** where necessary on innovation-related issues,
- developing and maintaining the ***partnerships*** required for research, innovation, technology transfer and education,
- employing ***governance*** mechanisms to assure relevance and accountability to its clients and stakeholders.

Creation of a Permanent Foresight Unit

In many ways, the implementation of the action plan started even before the exercise was finished. A part of the action plan is a natural continuation of consultations with major stakeholder groups that was started as part of the foresight process. The most immediate and tangible result was the creation of a permanent foresight unit within Teagasc to oversee the implementation of the Teagasc 2030 action plan and to support other foresight activities as needed within the organisation.

The action plan is outlined in the Teagasc 2030 report. It includes steps to create a broader culture of innovation within the organisation and to intensify systematic interaction with client groups and stakeholders. It addresses reform of personnel structures to enable greater mobility of staff within the organisation, facilitate transdisciplinary work and align incentives with the needs of clients. Other structural reforms include a focus on network-based activities, as well as time-limited project-network-like interventions such as technology platforms and commodity working groups that pool the resources of partners and involve stakeholders in management.

The general message of Teagasc 2030 is a positive one based on the opportunities offered by the KBBE, not only for actors currently involved in the agri-food and rural economy, but for a whole new generation of bio-entrepreneurs who may have no prior link to the land.

The key to success continues to be innovation. What is new is the pace of innovation and the need for organisations such as Teagasc to operate simultaneously on several fronts in a more international context and in shorter time frames. The challenge for Teagasc in the future will be to increasingly channel its efforts and resources towards support for innovation, in particular for the development of the knowledge-partnerships required by clients for innovation in the KBBE.

## References

All background papers, scenarios and proceedings as well as the final report are available from the Teagasc 2030 website at www.teagasc.ie/foresight/index.htm. The papers and presentations of the Radical Thinkers Workshop are available at http://www.teagasc.ie/publications/2007/20070725/index.htm.

Lance O'Brien is the head of the new Foresight Unit. He can be contacted at lance.obrien@teagsc.ie

# Section V
# Conclusion

The book is composed of contribution from many authors. It doesn't cover all the past and current research projects. It presents the vision of limited number of authors.

The book demonstrates the development of ICT for agriculture, rural development, food and environment mainly in the last decade. It showed rapid roll out of infrastructure on one side, but also slower take up on new applications on the other side. The adoption of new solutions by rural communities is a problem until now. We see important to make better link between different groups of scientists but also between scientists, policymakers and developers. If research results will be transferred into innovation, research will be successful. To guarantee the sustainability of research results is one of the challenges for the next period.

The book collects different visions of future for farming sector and rural regions. It is clear that the rural regions will be under strong pressure in the next period and ICT, knowledge management and research have to play the key role.

We would like to use the book as a basis for further discussion and intensive exchange of experiences and research results. The information collected in this book will be a basis for Strategic research Agenda of the agriXchange project.

ISBN: 978-80-905151-0-9