

Návrh datového formátu pro elektrofyziologické experimenty

Jiří Vaněk¹

1 Úvod

Na Západočeské univerzitě probíhá měření a zpracování elektrické aktivity mozku s využitím metod elektroencefalografie (EEG) a evokovaných potenciálů (ERP), tímto způsobem se zkoumá např. pozornost řidičů, poruchy vývojové koordinace u dětí nebo slepota u myši. V rámci výzkumu vzniká ročně několik desítek experimentů, jejichž parametry a výsledky jsou v současné době ukládány do tzv. EEG/ERP Portálu (EEGBase). Výsledky experimentů jsou uloženy v souborech vytvořených nahrávacím softwarem Brain Vision Recorder. Datový formát používaný tímto softwarem, ačkoliv je velmi dobře dokumentovaný, je proprietární.

Obecně je většina dat z elektrofyziologických experimentů po vyhodnocení a publikaci výsledků zahozena či uložena bez dalšího sdílení. Motivace pro sdílení experimentů je však značná. Sdílení naměřených dat může urychlit výzkum. Milham (2012), Podrack (2012), zlepšit kvalitu naměřených dat (sdílením dat se mohou odhalit chyby při měření, rušení,...) a také snížit časovou i finanční náročnost experimentů. Poline (2012). V současnosti však neexistuje žádný obecně přijatý standard pro sdílení elektroencefalografických dat.

2 Současný stav

Jelikož v současnosti neexistuje žádný standard pro reprezentaci a uchovávání dat a metadat z elektrofyziologických experimentů, vznikla pracovní skupina při International Neuroinformatics Coordinating Facility (INCF) s názvem „Program on Standards for Data Sharing“ Friedrich (2014), která se vytvořením takového standardu zabývá. Tato skupina má za úkol kromě standardu navrhnout nástroje, které mají usnadnit dlouhodobé ukládání a sdílení elektrofyziologických dat.

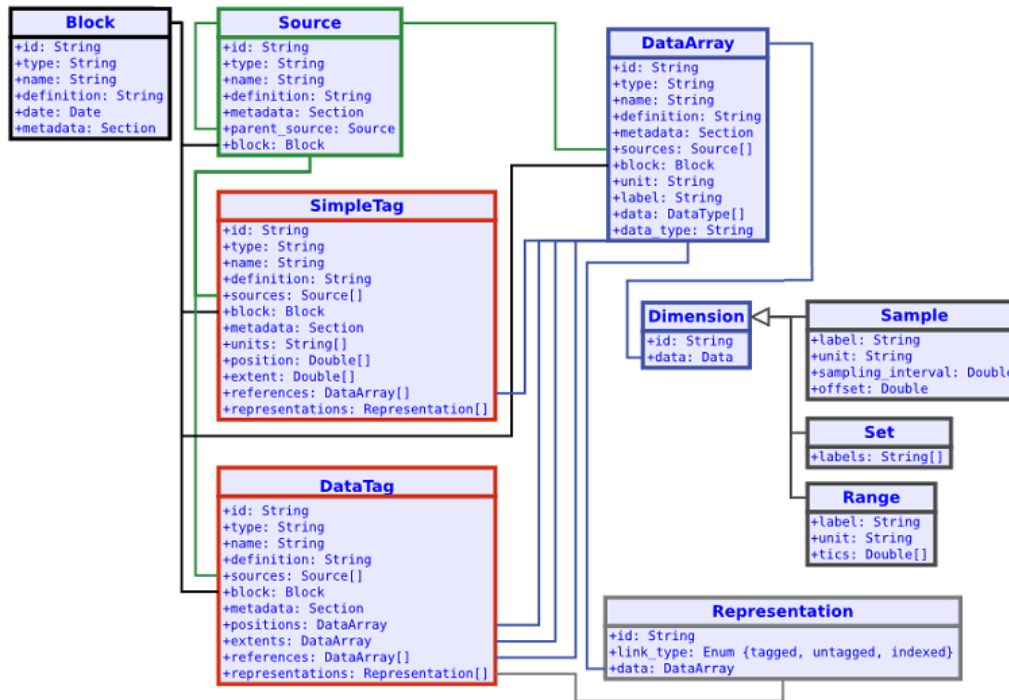
V rámci návrhu formátu pro ukládání elektrofyziologických dat jsem prozkoumal již existující formáty pro ukládání těchto dat, včetně příslušných terminologií, eventuálně ontologií. Zaměřil jsem se především na otevřené formáty a formáty využívající pro ukládání dat Hierarchical Data Format 5 (HDF5) kontejner, jelikož INCF doporučila využití právě tohoto formátu pro ukládání elektrofyziologických dat. Mezi zkoumané formáty patří Ovation, European Data Format, NeXus Format, NEO, NIX (Pandora) a epHDF.

3 Návrh datového formátu

Vzhledem k vznikajícímu standardu (požadavky na ukládaná data z elektrofyziologických experimentů jsou k dispozici v interním INCF dokumentu „Requirements for storing electrophysiology data“ ve verzi 0.72 z 3. 11. 2014) jsem vybral doporučovaný formát NIX, který posloužil jako základ pro vznikající EEGBase model. Výhody tohoto formátu jsou, že je open source, využívá kontejneru HDF5 a pro ukládání

¹ student navazujícího studijního programu Inženýrská informatika, obor Softwarové inženýrství, specializace Softwarový inženýr, e-mail: vanek2@students.zcu.cz

metadat (informací o prováděném experimentu a měření) využívá open metadata Markup Language (odML). Jelikož byl model NIX (zobrazen na obrázku 1) pro naše potřeby příliš obecný, vynechal jsem některé části, které nebyly povinné, čímž se EEGBase model značně zjednodušil, ale přesto je kompatibilní s modelem NIX. Samotné ukládání metadat není v modelu NIX definováno, proto jsem použil odML terminologii, kterou jsem rozšířil (rozšíření byla přijata INCF), a strukturu ukládání metadat jsem zvolil vlastní.



Obrázek 1: Datový model NIX

3 Program pro ukládání do nového formátu

V rámci diplomové práce byl také vytvořen program, jenž umožní export naměřených dat a experimentů ve formátu Brain Vision, jenž jsou uloženy v EEGBase portálu do nového formátu. Jelikož formát NIX nemá žádné Java API, bylo nutné použít knihovny HDF5 formátu jak pro ukládání datové, tak metadatové části souboru.

Literatura

- EEGBase, 2015. RRID:nif-0000-08190. Dostupný online z <http://eegdatabase.kiv.zcu.cz>
- Friedrich, S. et al., 2014. Mission and activities of the INCF Electrophysiology Data Sharing Task Force. *Frontiers in Neuroinformatics*. DOI: 8. 10.3389/conf.fninf.2014.08.00088. Dostupný online z: <http://dx.doi.org/10.3389/conf.fninf.2014.08.00088> .
- Milham, M. P., 2012. Open Neuroscience Solutions for the Connectome-wide Association Era. *Neuron*. jan 2012, 73, 2, s. 214-218. 10.1016/j.neuron.2011.11.004. Dostupný online z: <http://dx.doi.org/10.1016/j.neuron.2011.11.004>.
- [3] Poldrack, R. A. The future of fMRI in cognitive neuroscience. *NeuroImage*. aug 2012, 62, 2, s. 1216-1220. 10.1016/j.neuroimage.2011.08.007. Dostupný online z: <http://dx.doi.org/10.1016/j.neuroimage.2011.08.007>.
- Poline, J.-B. et al., 2012. Data sharing in neuroimaging research. *Frontiers in Neuroinformatics*. DIO: 6. 10.3389/fninf.2012.00009. Dostupný online z: <http://dx.doi.org/10.3389/fninf.2012.00009>.