NES: a free software to manage data from neuroscience experiments

Margarita Ruiz Olazar[1], Evandro S. Rocha[1], Sueli D. Rabaça[1], Carlos E. Ribas[1], Claudia D. Vargas[2], Amanda S. Nascimento[3] and Kelly R. Braghetto[4*]

1 Research, Innovation and Dissemination Center for Neuromathematics, Institute of Mathematics and Statistics, Brazil
2 Instituto de Biofisica Carlos Chagas Filho, Laboratório de Neurobiologia II, Brazil
3 Universidade Federal de Ouro Preto, Instituto de Ciências Exatas e Biológicas, Departamento de Computação, Brazil
4 University of São Paulo, Department of Computer Science, Institute of Mathematics and Statistics, Brazil

Neuroscientists perform complex experiments aiming to reach a more effective understanding about the functioning of the brain and the treatment of its pathologies. Each research laboratory uses different techniques and methodologies to produce and to analyze its findings. These experiments generate large volumes of data in diverse formats. Furthermore, with the increasing of the scale of research projects, many research laboratories are confronted with new technical (e.g. data size, quality control, analysis complexity, guaranteed reproducibility) and social (e.g. employee turnover, data sharing requirements) challenges [1]. This scenario imposes the use of computational tools to document each step of the experiments and to facilitate the electronic data capture. Addressing this problem, we present the Neuroscience Experiments System (NES), a software to manage data from neuroscience experiments.

NES is a web-based system that provides facilities to record data and metadata from each step of a neuroscience experiment in a secure and user friendly platform. It was developed to assist researchers in their data collecting routine throughout a neuroscience experiment, integrating data records from different types such as clinical, electrophysiological, imaging, and behavioral. Furthermore, it provides a unified repository (database) for the experimental data of an entire research laboratory, group, or project. Its web interface and modular format provide an intuitive use of its data management functionalities. Its use does not depend on any specific knowledge on informatics. NES is a free software tool, developed using open technologies and tools which can be easily installed and used in any research laboratory.

NES was developed to keep together experimental data and its fundamental provenance information, defined by the seven W's (Who, What, Where, Why, When, Which, (W) how) [2]. Examples of provenance information maintained by NES are: information about the scientists responsible for the experiment and collection of data and the description of the subject groups (who); the details about the recording protocol or behavioral data collection (e.g. the types of data collection performed) (what); the details of the experimental protocol used in the collection of primary data (how); the start/end date-time for data collection (when); the purpose of the experiment (why); the information about the experimental conditions to which the groups of subjects are submitted, such as tasks to perform and stimulus to apply (which); the information about the laboratory where data was collected (where) and even publications or other results that have arisen from the study of the collected data. Scientists can also record additional details for each experiment's volunteer, such as the information about his/her clinical history and sociodemographic data.

In NES, each primary data acquired from an experiment is always associated to a study participant and to a specific step of the experimental protocol. For this reason, before storing the primary data collected in an experiment, the researcher needs to record in detail each step involved in the experimental protocol (e.g., the specific preparation for the realization of the experiment). The experimental protocol is described as a workflow, that can contain both sequential and parallel steps. With this, NES provides a structured and comprehensive platform with a robust tracking of data provenance that is fundamental to enable the reproduction of the experiment.

It is worth to mention that NES is not a new proposal to standardize the experimental data representation. There are several models and formats (e.g. NeoHDF5 [3], NWB [4], NIX [5]) currently in development to address this issue. These models are appropriate for organizing and exchanging data of a particular type and from a particular experiment. However, they do not replace the function of a database. A database system keeps large data volumes and provides functionalities for access control, data consistency, fault tolerance and efficient data recovery. Furthermore, on a database it is possible to store the relationships among different types of data from different experiments, allowing for more sophisticate data analysis which are specially valuable to support advances in the healthcare domain.

The data model used in NES is aligned with several formats used in neuroscience, enabling interoperability with the most promising initiatives for standardization of data representation for electrophysiology and neuroimaging, as much as with guidelines to report neuroscience experiments (e.g., MINI [6]; MINEMO [7]; fMRI [8]). Currently, NES is able to manage several types of electrophysiological data and metadata used by the neuroscience community, and the neuroimaging module is under development.

To handle the heterogeneity of experimental data, NES adopts a polyglot persistence strategy, combining different kinds of storage technologies, such as relational databases, NoSQL databases and distributed file systems.

NES is an initiative of the Research, Innovation and Dissemination Center for Neuromathematics (NeuroMat), hosted at University of São Paulo, Brazil. It is licensed under the Mozilla Public License version 2.0 and its source code and documentation are available at https://github.com/neuromat/nes. One of the goals of NeuroMat is to establish in Brazil a leading research group in Neuroinformatics, specialized in curating experimental data. With the development of free software tools such as NES, we want to encourage and support the collection of high-quality data and metadata and the

creation of more open databases in the Neurosciences area.

Acknowledgements

This work was produced as part of the activities of FAPESP Research, Innovation and Dissemination Center for Neuromathematics (grant 2013/07699-0, S.Paulo Research Foundation)

References

1. Buckow, K., Quade, M., Rienhoff, O., & Nussbeck, S. Y. (2014). Changing requirements and resulting needs for IT-infrastructure for longitudinal research in the neurosciences. Neuroscience research. doi:10.1016/j.neures.2014.08.005
2. Goble, C. (2002, October). Position statement: Musings on provenance, workflow and (semantic web) annotations for bioinformatics. In Workshop on Data Derivation and Provenance, Chicago (Vol. 3).
3. Garcia, S., Guarino, D., Jaillet, F., Jennings, T., Pröpper, R., Rautenberg, P. L., ... & Davison, A. P. (2014). Neo: an object model for handling electrophysiology data in multiple formats. Front. Neuroinform, 8(10). doi: 10.3389/fninf.2014.00010
4. Teeters, J. L., Godfrey, K., Young, R., Dang, C., Friedsam, C., Wark, B., ... & Denisov, G. (2015). Neurodata Without Borders: Creating a Common Data Format for Neurophysiology. Neuron, 88(4), 629-634. doi:10.1016/j.neuron.2015.10.025
5. Neuroscience information exchange format (NIX). https://github.com/G-Node/nix.
6. Gibson, F., Overton, P. G., Smulders, T. V., Schultz, S. R., Eglen, S. J., Ingram, C. D., ... & Cunningham, M. (2008). Minimum information about a neuroscience investigation (MINI): electrophysiology. Nat. Precedings. doi:10.1016/j.pbiomolbio.2011.07.001
7. Frishkoff, G., Sydes, J., Mueller, K., Frank, R., Curran, T., Connolly, J., ... & Malony, A. (2011). Minimal Information for Neural Electromagnetic Ontologies (MINEMO): A standards-compliant method for analysis and integration of event-related potentials (ERP) data. Standards in genomic sciences, 5(2), 211-223. doi:10.4056/sigs.2025347
8. Poldrack, R. A., Fletcher, P. C., Henson, R. N., Worsley, K. J., Brett, M., & Nichols, T. E. (2008). Guidelines for reporting an fMRI study. Neuroimage, 40(2), 409-414. doi:10.1016/j.neuroimage.2007.11.048

Keywords: Neurosciences, Experimental data, Electrophysiology, Provenance information, databases
Conference: Neuroinformatics 2016, Reading, United Kingdom, 3 Sep - 4 Sep, 2016. Presentation Typeº Demo Topicº Data systems
Citation: Ruiz Olazar M, Rocha ES, Rabaça SD, Ribas CE, Vargas CD, Nascimento AS and Braghetto KR (2016). NES: a free software to manage data from neuroscience experiments. Front. Neuroinform. Conference Abstract: Neuroinformatics 2016. doi: 10.3389/conf.fninf.2016.20.00043
Received: 29 Apr 2016; Published Online: 18 Jul 2016.
* Correspondence: PhD. Kelly R Braghetto, University of São Paulo, Department of Computer Science, Institute of Mathematics and Statistics, São Paulo, São Paulo, 05508-090, Brazil, kellyrb@ime.usp.br