frontiers in ICT

EDITORIAL
published: 17 July 2017
doi: 10.3389/fict.2017.00018

Editorial: MAPPING: MAnagement and Processing of Images for Population ImagiNG

Michel Dojat[1], David N. Kennedy[2] and Wiro Niessen[3]*

[1] U1216-GIN, INSERM, Site Santé, La Tronche, France, [2] Medical School, University of Massachusetts, Worcester, MA, United States, [3] Erasmus University, Rotterdam, Netherlands

Keywords: data sharing, neuroimaging, brain, magnetic resonance imaging, image processing, computer-assisted

Editorial on the Research Topic

MAPPING: MAnagement and Processing of Images for Population ImagiNG

Several recent papers underline methodological points that limit the validity of published results in imaging studies in the life sciences and especially the neurosciences (Ioannidis, 2005; Carp, 2012; Button et al., 2013; Ingre, 2013). At least three main points are identified that lead to biased conclusions in research findings: endemic low statistical power, selective outcome, and selective analysis reporting. Because of this, and in view of the lack of replication studies, false discoveries or solutions persist. To overcome the poor reliability of research findings, several actions should be promoted including conducting large cohort studies, data sharing, and data reanalysis. The construction of large-scale online databases should be facilitated, as they may contribute to the definition of a "collective mind" (Fox et al., 2014) facilitating open collaborative work or "crowd science" (Franzoni and Sauermann, 2014). Although technology alone cannot change scientists' practices (Wicherts et al., 2011; Wallis et al., 2013; Poldrack and Gorgolewski, 2014; Roche et al., 2014), technical solutions should be identified, which support a more "open science" approach. Also, the analysis of the data plays an important role. For the analysis of large datasets, image processing pipelines should be constructed based on the best algorithms available and their performance should be objectively compared to diffuse the more relevant solutions. Also, provenance of processed data should be ensured (MacKenzie-Graham et al., 2008). In population imaging, this would mean providing effective tools for data sharing and analysis without increasing the burden on researchers. This subject is the main objective of this research topic (RT), cross-listed between the specialty section "Computer Image Analysis" of Frontiers in ICT and Frontiers in Neuroinformatics. First, it gathers works on innovative solutions for the management of large imaging datasets possibly distributed in various centers. The paper of Danso et al. describes their experience with the integration of neuroimaging data coming from several stroke imaging research projects. They detail how the initial NeuroGrid core metadata schema was gradually extended for capturing all information required for future meta-analysis while ensuring semantic interoperability for future integration with other biomedical ontologies. With a similar preoccupation of interoperability, Shanoir relies on the OntoNeuroLog ontology (Temal et al., 2008; Gibaud et al., 2011; Batrancourt et al., 2015), a semantic model that formally described entities and relations in medical imaging, neuropsychological, and behavioral assessment domains. The mechanism of "Study Card" allows to seamlessly populate metadata aligned with the ontology, avoiding fastidious manual entrance and the automatic control of the conformity of imported data with a predefined study protocol. The ambitious objective with the BIOMIST platform is to provide an environment managing the entire cycle of neuroimaging data from acquisition to analysis ensuring full provenance information of any derived data. Interestingly, it is conceived based on the product lifecycle management approach used in industry for managing products (here

OPEN ACCESS

Edited and Reviewed by:
Kaleem Siddiqi, McGill University, Canada

****Correspondence:***
Michel Dojat
michel.dojat@univ-grenoble-alpes.fr

Specialty section:
This article was submitted to Computer Image Analysis, a section of the journal Frontiers in ICT

Received: *28 April 2017*
Accepted: *29 June 2017*
Published: *17 July 2017*

Citation:
Dojat M, Kennedy DN and Niessen W (2017) Editorial: MAPPING: MAnagement and Processing of Images for Population ImagiNG. Front. ICT 4:18. doi: 10.3389/fict.2017.00018

neuroimaging data) from inception to manufacturing. Shanoir and BIOMIST share in part the same OntoNeuroLog ontology facilitating their interoperability. ArchiMed is a data management system locally integrated for 5 years in a clinical environment. Not restricted to Neuroimaging, ArchiMed deals with multimodal and multi-organs imaging data with specific considerations for data long-term conservation and confidentiality in accordance with the French legislation. Shanoir and ArchiMed are integrated into FLI-IAM,[1] the national French IT infrastructure for *in vivo* imaging.

Second, dedicated software and hardware infrastructures are proposed for the sharing and execution of image-processing workflows making easier the replication and comparison of data analysis procedures. The contribution of Das et al. presents the functionalities added to the LORIS-CBRAIN software ecosystem to fulfill the technical challenges raised by supporting an Open Science approach. Specific mechanisms have been introduced for ensuring privacy and security of the stored data, quality control checking, and heterogeneous tools integration. Fastr is a workflow engine dedicated to the automation of complex medical imaging processing pipelines. It allows the composition of different software elements to design pipelines, checks datatype compatibility of linked outputs and inputs, ensures data provenance, and finally creates a list of jobs for execution. In the same vein, OpenMOLE is designed to optimize execution of workflows on distributed computing architectures. Although no specific application domain is targeted by OpenMOLE, case studies are reported to illustrate its suitability to neuroimaging data processing. How to document data provenance to facilitate processed data sharing and reuse is the question explored by Pauli et al. from datasets processed using the most common software package used in Neuroimaging. They provide a set of results as a benchmark for testing automated provenance software.

Finally, two papers are more concerned with the usage of such platforms. Serag et al. propose SEGMA, a supervised solution for brain tissue and structure segmentation combining sparse training data selection, linear registration, and random forest classifier for processing large MR datasets with a reduced computational time. Brain atlases are often used by automated workflows for imaging population studies. The paper by Dickie et al. reviews the brain MRI atlases currently available, which appear of modest size, based on limited image sequences and where some populations are underrepresented. The next challenge is then to develop non-parametric brain atlases including a wide number of parameters extracted from different imaging sequences from a large set of individuals, representative of more different classes of population.

To conclude, this RT demonstrates that, since the pioneer experiments of neuroimaging data sharing with the fMRIDC project (Van Horn and Gazzaniga, 2013) or the BIRN initiative (Keator et al., 2008), many technical efforts have been performed or are currently underway to facilitate data and tools sharing. Solutions now exist that are mature enough to help us make substantial changes to how we conduct health research (Chan et al., 2014), improving reproducibility, and quality of published research findings.

AUTHOR CONTRIBUTIONS

The authors contributed equally to this editorial.

[1]https://project.inria.fr/fli/en/.

REFERENCES

Batrancourt, B., Dojat, M., Gibaud, B., and Kassel, G. (2015). A multilayer ontology of instruments for neurological, behavioral and cognitive assessments. *Neuroinformatics* 13, 93–110. doi:10.1007/s12021-014-9244-3

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., et al. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376. doi:10.1038/nrn3475

Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *Neuroimage* 63, 289–300. doi:10.1016/j.neuroimage.2012.07.004

Chan, A.-W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gøtzsche, P. C., et al. (2014). Increasing value and reducing waste: addressing inaccessible research. *Lancet* 383, 257–266. doi:10.1016/S0140-6736(13)62296-5

Fox, P. T., Lancaster, J. L., Laird, A. R., and Eickhoff, S. B. (2014). Meta-analysis in human neuroimaging: computational modeling of large-scale databases. *Annu. Rev. Neurosci.* 37, 409–434. doi:10.1146/annurev-neuro-062012-170320

Franzoni, C., and Sauermann, H. (2014). Crowd science: the organization of scientific research in open collaborative projects. *Res. Policy* 43, 1–20. doi:10.1016/j.respol.2013.07.005

Gibaud, B., Kassel, G., Dojat, M., Batrancourt, B., Michel, F., Gaignard, A., et al. (2011). "NeuroLOG: sharing neuroimaging data using an ontology-based federated approach," in *AMIA 2011 Annual Symposium*, ed. R. Scott Evans.

Ingre, M. (2013). Why small low-powered studies are worse than large high-powered studies and how to protect against "trivial" findings in research: comment on Friston (2012). *Neuroimage* 81, 496–498. doi:10.1016/j.neuroimage.2013.03.030

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med.* 2:e124. doi:10.1371/journal.pmed.0020124

Keator, D. B., Grethe, J. S., Marcus, D., Ozyurt, B., Gadde, S., Murphy, S., et al. (2008). A national human neuroimaging collaboratory enabled by the Biomedical Informatics Research Network (BIRN). *IEEE Trans. Inf. Technol. Biomed.* 12, 162–172. doi:10.1109/TITB.2008.917893

MacKenzie-Graham, A. J., Van Horn, J. D., Woods, R. P., Crawford, K. L., and Toga, A. W. (2008). Provenance in neuroimaging. *Neuroimage* 42, 178–195. doi:10.1016/j.neuroimage.2008.04.186

Poldrack, R. A., and Gorgolewski, K. J. (2014). Making big data open: data sharing in neuroimaging. *Nat. Neurosci.* 17, 1510–1517. doi:10.1038/nn.3818

Roche, D. G., Lanfear, R., Binning, S. A., Haff, T. M., Schwanz, L. E., Cain, K. E., et al. (2014). Troubleshooting public data archiving: suggestions to increase participation. *PLoS Biol.* 12:e1001779. doi:10.1371/journal.pbio.1001779

Temal, L., Dojat, M., Kassel, G., and Gibaud, B. (2008). Towards an ontology for sharing medical images and regions of interest in neuroimaging. *J. Biomed. Inform.* 41, 766–778. doi:10.1016/j.jbi.2008.03.002

Van Horn, J. D., and Gazzaniga, M. S. (2013). Why share data? Lessons learned from the fMRIDC. *Neuroimage* 82, 677–682. doi:10.1016/j.neuroimage.2012.11.010

Wallis, J. C., Rolando, E., and Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS ONE* 8:e67332. doi:10.1371/journal.pone.0067332

Wicherts, J. M., Bakker, M., and Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS ONE* 6:e26828. doi:10.1371/journal.pone.0026828

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.