

Data and Analysis Management for Functional Magnetic Resonance Imaging Studies

K. Jarrod Millman, Mark D’Esposito

Abstract—There is a critical shortage of efficient ways to record the information necessary to share fMRI data across labs or even within labs [1], [2], [3]. Recent developments in the field of neuroinformatics have provided the community with tools both for recording and annotating analyses even within complex customized processing streams. However, none of the current solutions integrate provenance capture within any of the current neuroimaging analysis tools. We propose an aspected-oriented provenance library integrated with existing tools to provide full support for documenting analyses and provenance tracking. Currently we are developing a prototype library, which we will integrate into the Neuroimaging Tools for Python (NiPy) project.

Index Terms—Functional MRI, data management, provenance tracking.

INTRODUCTION

There are numerous MRI methods that have been developed to study brain function such as high-resolution structural MRI, diffusion-tensor imaging and functional MRI (fMRI). Over the last decade, fMRI, in particular, has developed into a powerful method for studying human brain function [4]. fMRI is a non-invasive method that monitors blood flow and blood oxygenation changes correlated with neural activity. Although its predominant use by neuroscientists has been to examine the neural basis of cognitive and behavioral processes, it is more frequently being used to study patients with neurological and psychiatric disease. Its widespread availability, non-invasiveness, high spatio-temporal resolution and reasonable cost, especially when compared to positron emission tomography (PET) scanning, have all contributed to its increasing popularity. In addition, fMRI has earned a strong reputation over the last ten years by reproducing and extending many important neurophysiological findings, which have previously been the domain of experimental animal research. Importantly, the introduction of fMRI did not replace other neuroscientific methods for studying human brain function. Rather, it provided a new and unique way of examining brain-behavior relationships. Over the past few years, the success of fMRI has led to an exponential increase in its use to study brain function. A bibliographic search for the term “fMRI” in MEDLINE in the years 2003-2005 leads to over 7500 citations. The results of this search also reveal that fMRI studies are being published in a wide range of journals across many disciplines. Rapid advances in fMRI methods in both data acquisition and data

analysis are occurring at a rapid pace, which will likely fuel even more widespread use of this technology in the future.

For current fMRI studies to be available to future researchers, we need user-friendly, extensible, scalable, and interoperable tools to consistently store study data and analysis provenance. Several factors make this problem unusually challenging: (1) the massive volume of fMRI data in each study, (2) the numerous experimental design and MRI parameters that could significantly affect results, (3) the wide variety of analysis software and methodologies, (4) the disparate fields of knowledge required to fully understand the results of these studies, and (5) the inherent complexity of nervous system data [5], [6]. As a result of funding from the Human Brain Project (HBP), there are currently several efforts underway to address this need in the neuroimaging community; however, these efforts focus on either (1) creating a standalone suite separate from the major neuroimaging analysis packages or (2) providing a separate add-on toolbox to one or more of the current analysis packages. Ultimately to achieve the widespread deployment and adoption necessary to realize the goals of reproducible research, the facilities and tools created will need to be intimately wed to the actual analysis packages.

The typical fMRI experimenter is confronted with a bewildering complex of raw and processed data. Without knowing the precise acquisition parameters and conditions, raw fMRI data is meaningless. Without tracking the exact history of the analytic processing, the analysis is unreproducible. Thus, it is not surprising that developing methods to record and query this metadata is crucial to advancing the field. Such metadata (even if recorded) is scattered and too often lost. For example, subject information may be stored in a subject database; scanner parameters may be stored somewhere in the configurations of the data acquisition software; analysis, experimental, and stimulus details may each reside in different sets of files in different directories or even machines. Yet, effective collaboration between researchers or even the ability to reanalyze one’s own data depends critically on access to the totality of this disparate, unorganized information.

Further complicating the situation is the fact that researchers (most often post-docs) who conduct fMRI studies frequently leave the institution at which they carried out their study, thus geographically separating the different domain experts. Even if the departing researcher were to finish and publish their study before leaving, it is often not clear what files need to be saved in order to retain all the necessary information to make sense of their data. When new data or methods arise, the ability to apply them to existing data is limited by the extent to which this information is available and retrievable;

and once retrieved, uncertainty about the older data may make it unusable.

The problem compounds once collaboration extends to more than one neuroscience researcher. It would be easy to avoid such a situation were it not for the fact that some studies necessarily depend on the cooperation of multiple researchers at multiple locations. For example, patient studies may require the collaboration of researchers in different institutions since the patient pool in one location may not be complete enough. Despite thousands of published fMRI studies, only a few multi-site fMRI studies have been completed [7], which attests to the difficulty in sharing imaging data between labs. Another important consideration is that many imaging centers do not rely solely on fMRI data, but include multiple imaging modalities.

RELATED WORK

Currently, there are at least three different scales at which this problem exists:

- 1) small scale—the individual researcher or research laboratory level
- 2) medium scale—the research center or consortium level
- 3) large scale—the neuroimaging community level

At each scale the problem is manifest in a slightly different form given the different contexts and resources available.

Small Scale: Whether by using ad-hoc laboratory notebooks and manually recording the directory and file naming conventions or by using a more systematic data management system, every neuroimaging research facility deals with this data management problem in one way or another. Unfortunately, most rely on local, partial solutions out of necessity. However, several researchers have started replacing these homegrown systems with more sophisticated systems utilizing current industry and informatics best practices.

BrainVISA (<http://www.brainvisa.info>) was developed by methodologists [8] of the Institut Federatif de Recherche 49 (<http://www.ifr49.org>) to simplify the integration of various neuroimaging tools and to combine data from several acquisition modalities (aMRI, fMRI, dMRI, EEG, MEG, PET, etc). In order to remain compatible with existing software, BrainVISA's data management system is based on directories and file organization and naming.

UC Los Angeles' Laboratory of Neuroimaging (LONI) has developed a Pipeline Processing Environment [9], which provides an easy, graphical way for researchers to join together independent executable programs, or modules, into a processing stream. They have developed a simple graphical interface in Java and the description of the modules and pipelines is implemented in XML.

Medium Scale: At the center or consortium level, there is a substantial increase in the amount of data to track; traditional, straightforward data management practice is no longer a viable solution. Given the increased size of these research groups there has been a corresponding increase in the amount of informatics resources available. XCEDE and XNAT are two notable projects, which have been developed for these mid-sized research groups.

The Biomedical Informatics Research Network (BIRN) aids distributed collaborations in biomedical science by utilizing information technology innovations [10], [11]. To handle the increasingly large and diverse amounts of imaging and associated clinical data generated and maintained by the BIRN, they have developed an extensible database management system. This system is composed of 3 components: the Human Clinical Imaging Database (HCID), the XML-Based Clinical Experiment Data Exchange (XCEDE) schema [12], and a J2EE web interface [13]. HCID is a database schema with 1) an extensible framework for definition and storage of subject assessment data and 2) the incorporation of provenance tracking.

The eXtensible Neuroimaging Archive Toolkit (XNAT) was developed at the Howard Hughes Medical Institute at Washington University in St. Louis [14], [15]. XNAT is an easily customizable toolkit for managing fMRI data, which uses XML to represent and specify the structure of experimental data. Once a site-specific XML specification is created, XNAT tools will automatically generate a database schema, dynamic web pages, and data access code.

While both groups aim to make their systems available to smaller groups in easy to install bundles, the ability for small labs to effectively modify or customize these systems is greatly limited given the complexity of the underlying codebase and programming technologies, which rely on enterprise grade Java code.

Large Scale: The fMRI Data Center (fMRIDC) is a public repository of data associated with peer-reviewed fMRI studies. After overcoming immense sociological, scientific, and technical challenges in the early stages of its existence, the fMRIDC has grown to contain terabytes of data from nearly 100 studies [16]. The fMRIDC's database is organized according to an ontology [17], [18], [19], rather than a traditional database schema. An ontology is a formally defined collection of categorical concepts and the relationships between them. Ontological data annotation enables the application of automated inference, clustering, and other machine learning techniques to enhance higher order analysis and querying.

Berkeley is working with the fMRIDC to further develop their ontology for fMRI studies using a 2 year grant from the HBP (NIH P20 MH72580-01-Dartmouth). We are developing self-describing data formats incorporating terms from the ontology for data description. We are using XML as our preferred syntax for these data formats as it is inherently extensible, leaving room for future format evolution, which may not exist for simpler header-based formats [20], [21]. The data format will capture as much as is known about the data it contains, ensuring maximum interpretability. This includes, but is not limited to: scanner settings; details of the pulse sequences used; and any downstream analysis, such as image reconstruction algorithms and other analytic parameters the data may have been subject to (see also <http://nifti.nimh.nih.gov>). We are developing tools for converting these self-describing formats to and from other common fMRI image data formats.

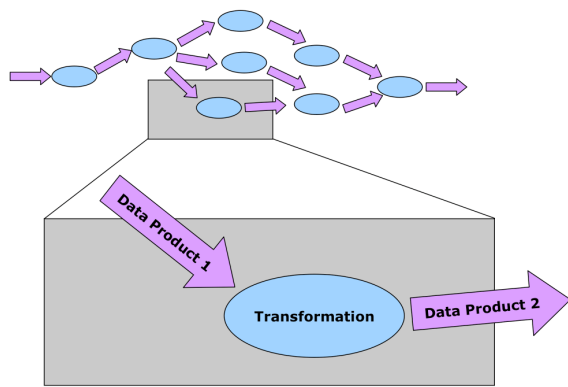


Fig. 1. Provenance as a directed acyclic graph (DAG).

METHODOLOGY

The proposed provenance library will leverage the fMRI ontology and XML data schema being jointly developed by the fMRIDC and Berkeley to enable sharing fMRI data and analyses with the wider imaging community. Provenance tracking is a known hard problem in computer science and an active area of research [22], [23], [24], [25].

In general terms, the provenance (also referred to as lineage) of a given data set includes two components: 1) ancestral data products and 2) the transformations those data products underwent. Often it is convenient to conceptualize the provenance as a graph (specifically, a directed acyclic graph or DAG) where the nodes are transformation processes and the edges are data products (see Fig. 1). The means by which the transformation processes execute, as well as consume and transform data, is referred to as the data processing architecture. The most common data processing architectures used with provenance tracking are service-oriented (e.g., [26], [27], [28]) associated with Grid computing and database oriented (e.g., [29], [22], [30]) associated with data warehouses. The advantages of these architectures lies in their ability to automate semantic annotation (in conjunction with a predefined ontology). The addition of this semantic annotation on the DAG allows for intelligent searching, categorization, and automated inference.

Script-based Processing Architecture

Service-oriented and database oriented processing architectures, however, do not represent the current computing environments nor common research practice in neuroimaging. In essence these enterprise grade architectures prevent individual researchers from directly disseminating their work to the community and instead require some data center to broker their publication.

Currently, there is little support for bridging the gap between the individual researcher and the enterprise level data systems. Typically, individual researchers responsible for creating data products use or write scripts to process or transform their raw data and use idiosyncratic naming conventions for their files and directories. Since they are not using such an enterprise level data processing architecture, there is no guarantee that

they will retain the necessary provenance information once a "publishable" result is achieved.

Our approach is very similar to a one developed at University of California, Santa Barbara for managing Earth science data [31], [32]. We will use objects to model our programs and files; thus, the metadata about our objects will be their attributes. Using this model, our goal is to provide a minimally disruptive analysis management infrastructure with the rigor of traditional enterprise architectures. To accomplish this goal, we must address two key questions:

- 1) How do we define the attributes of our objects?
- 2) How can we automatically populate these object attributes?

To define our object attributes, we will utilize an XML provenance schema to define our data products and the transformations on them. Since most research scientists use scripts or command-line programs to process their data, our approach will use an application programming interface (API) and wrapper scripts to automatically record the metadata as XML files.

Creating a stable, consistent, and easy to use application programming interface (API) exposing high-level underlying functionality allows for rapid development of utilities and functional plug-ins to the Neuroimaging tools for Python (NiPy) project (<http://neuroimaging.scipy.org>). NiPy is based largely on BrainStat, Jonathan Taylor's Python-based package for statistical analysis in neuroimaging [33]. BrainStat, a port of Keith Worsley's fmristat program from Matlab to Python, has a full system for general linear modeling of fMRI data and principal component analysis.

The plug-in architecture will mainly support statistical modeling. A well-designed plug-in architecture will greatly facilitate distributed development effort by providing a functionally encapsulated environment for collaboration [34].

Aspect-oriented Provenance Tracking

Our provenance library will utilize Python metaprogramming techniques. As the name implies, metaprogramming is programming that controls the behavior of other programs. In this way, rather than writing an application to track provenance, we will write a library that enables other programs to track their own provenance. In particular, we will use dynamic proxies to implement provenance tracking as an aspect-oriented pattern [35], [36]. Dynamic proxies are objects which transparently wrap a target object, potentially intercepting method calls to interpose new behavior, while leaving the original target unmodified. The new behavior provided by such a proxy can be referred to as an aspect.

Python provides extensive introspection support, which can be used to reveal useful information about a program's objects. In this case, the dynamic proxies in our provenance library not only intercept calls and create extensive records of which methods were called and how, but they will also use introspection to determine additional information about the object's state.

We will investigate the use of decorators and metaclasses for implementing dynamic proxies. Decorators are syntactical

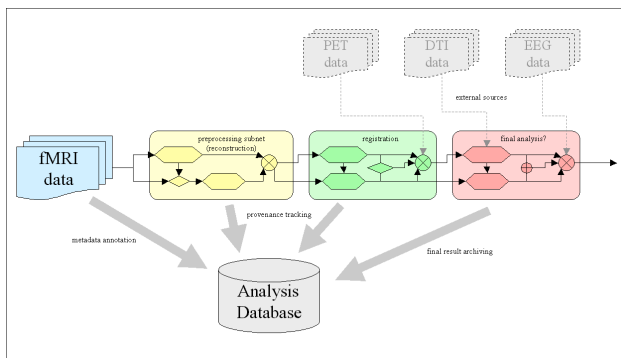


Fig. 2. FMRI Center Dataflow

sugar, which take some callable Python object (such as a function or class method) and return a new callable object, which may function differently from the original. A metaclass is a class factory. A custom metaclass can modify the way a Python class definition is used to define the class, enabling existing classes to incorporate new behavior, such as provenance tracking, without modification.

Code and data validation is another important component of our architecture. Our libraries will provide checksums for whatever software library, piece of code, or data set for a given analysis using Python introspection. Checksumming is a simple form of redundancy checking. It works by basically summing up the components of a given object (i.e., code or data) and storing the result. Later, you can determine whether the object has changed by recomputing the checksum and comparing it to the original checksum. This would allow a researcher to ensure that they are still using the data and code which were used previously.

Batch and Pipeline Processing

Pipeline analysis and batch processing are crucial aspects of modern scientific analysis software. Increasing volumes of raw data and a proliferation of experimental analysis techniques demand a more organized and formal data and dataflow management strategy. To this end, we are developing a flexible, modular analysis pipeline framework, which will enable exploratory analysis as well as capture best practices as they are developed.

A schematic representation of a typical dataflow for an imaging center is shown in Fig. 2. Based on scanner configuration data, fMRI raw scan data is initially annotated with metadata indicating parameters of the scan. This initial annotation will be completely automatic, and performed by an ontology-aware plug-in. As the data is processed and transformed by various tools, more annotation will occur using the aspect-oriented provenance library, capturing parameters and version information for each analysis algorithm involved in the pipeline. This will enable final results to be reliably reproduced from raw data sources.

SUMMARY AND CONCLUSION

The usefulness of provenance for a given domain is directly related to the level of granularity at which it is collected.

However, highly granular provenance data is more expensive to collect and maintain, due to its greater complexity. Thus, it is essential to determine what you need the provenance for as well as how it will be utilized.

Traditionally, scientists have published their ‘provenance history’ in the methods section of their scientific papers. Recently, neuroimagers have had extensive discussions regarding what are appropriate guidelines for presenting neuroimaging analyses [37]. The purpose of this discussion is stated as follows:

The methods which are used to collect and analyze fMRI, PET, SPECT, EEG and MEG data are quite varied. However, papers publishing results using such data often offer the most minimal descriptions (e.g. "Methods: SPM2 was used. Results: We found... "). The typical descriptions are deficient, in that they fail to meet a basic goal: Could another researcher, presented with an author’s data, reproduce the same results presented in the paper?

This effort is similar to work already done in the more mature field of event-related potentials (ERPs) [38].

For our purposes, a provenance system for neuroimaging data should imbue existing neuroimaging analysis software with the ability to record, store, query, and manipulate data provenance. We will first attempt to facilitate the collection of this information. As the project progresses we will refine the granularity to allow for new uses of the provenance. Attempting to form a system general enough to allow this site-specific knowledge to be incorporated at a later date risks either over- or under-generalizing. This is, perhaps, not to be avoided; instead of spending too much time planning it is better to sometimes proceed, developing a prototype [39]. With a prototype in place, it is easier to evaluate and assess the prototype’s ability to meet the needs it was designed to address. This leads to a natural cycle of development where old code is reviewed and refactored to better serve its purpose [40], [41].

A major constraint on this process is that continued development should always accommodate old data; that is, revisions of the provenance ontology will require either automated porting tools or remain backwards compatible. Another constraint on this work is that most researchers will be unable to manually enter all the requisite fields of the ontology due to the variety of knowledge required. So, where possible, it will be necessary to allow for the development of site-specific plug-ins to auto-populate much of this information.

ACKNOWLEDGEMENTS

We are grateful to Matthew Brett, Brian Hawthorne, Fernando Perez, and Jonathan Taylor for many useful discussions.

REFERENCES

- [1] S. H. Koslow, “Should the neuroscience community make a paradigm shift to sharing primary data?,” *Nat Neurosci*, vol. 3, no. 9, pp. 863–5, 2000. 1097-6256 Journal Article Review Review, Tutorial.

- [2] J. D. Van Horn, J. S. Grethe, P. Kostelec, J. B. Woodward, J. A. Aslam, D. Rus, D. Rockmore, and M. S. Gazzaniga, "The functional magnetic resonance imaging data center (fmridc): the challenges and rewards of large-scale databasing of neuroimaging studies," *Philos Trans R Soc Lond B Biol Sci*, vol. 356, no. 1412, pp. 1323–39, 2001. 0962-8436 Journal Article Meta-Analysis.
- [3] S. H. Koslow, "Opinion: Sharing primary data: a threat or asset to discovery?," *Nat Rev Neurosci*, vol. 3, no. 4, pp. 311–3, 2002. 1471-003x Journal Article.
- [4] M. D'Esposito, "Functional neuroimaging of cognition," *Semin Neurol*, vol. 20, no. 4, pp. 487–98, 2000. 0271-8235 Journal Article Review Review, Tutorial.
- [5] M. F. Huerta, S. H. Koslow, and A. I. Leshner, "The human brain project: an international resource," *Trends Neurosci*, vol. 16, no. 11, pp. 436–8, 1993. 0166-2236 Journal Article.
- [6] M. F. Huerta and S. H. Koslow, "Neuroinformatics: opportunities across disciplinary and national borders," *NeuroImage*, vol. 4, no. 3 Pt 2, pp. S4–6, 1996. 1053-8119 Journal Article.
- [7] B. J. Casey, J. D. Cohen, K. O'Craven, R. J. Davidson, W. Irwin, C. A. Nelson, D. C. Noll, X. Hu, M. J. Lowe, B. R. Rosen, C. L. Truwitt, and P. A. Turski, "Reproducibility of fmri results across four institutions using a spatial working memory task," *NeuroImage*, vol. 8, no. 3, pp. 249–61, 1998. 1053-8119 Journal Article Multicenter Study.
- [8] Y. Cointepas, J.-F. Mangin, L. Garnero, J.-B. Poline, and H. Benali, "BrainVISA: Software platform for visualization and analysis of multi-modality brain data," in *Proc. 7th HBM*, (Brighton, United Kingdom), p. S98, 2001.
- [9] D. E. Rex, J. Q. Ma, and A. W. Toga, "The LONI pipeline processing environment," *NeuroImage*, vol. 19, no. 3, pp. 1033–48, 2003. 1053-8119 Journal Article.
- [10] M. Ellisman and S. Peltier, *The Grid, Second Edition*, ch. Medical Data Federation: The Biomedical Informatics Research Network, pp. 109–120. Morgan Kaufmann, 18 Nov. 2003.
- [11] J. Jovicich, M. F. Beg, S. Pieper, C. Priebe, M. M. Miller, R. Buckner, and B. Rosen, "Biomedical Informatics Research Network: Integrating Multi-Site Neuroimaging Data Acquisition, Data Sharing and Brain Morphometric Processing," in *18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05)*, 2005.
- [12] S. Gadde and D. B. Keator, "Design and use of the Human BIRN XML schema." 2004.
- [13] I. Ozyurt, "J2EE based Web-Interface Development Framework for Clinical Imaging Databases." 2004.
- [14] D. Marcus, T. Olsen, M. Ramaratnam, A. Snyder, and R. Buckner, "XNAT - The extensible neuroimaging archive toolkit: informatics tools for managing and exploring neuroimaging data," in *uman Brain Project Annual Conference*, 2004.
- [15] D. Marcus, T. Olsen, M. Ramaratnam, and R. Buckner, "XNAT: A software framework for managing neuroimaging laboratory data," in *Society for Neuroscience*, 2004.
- [16] J. D. Van Horn, S. T. Grafton, D. Rockmore, and M. S. Gazzaniga, "Sharing neuroimaging studies of human cognition," *Nature Neuroscience*, vol. 7, pp. 473 – 481, Apr. 2004.
- [17] T. Gruber, "A translation approach to portable ontologies," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.
- [18] N. Guarino and P. Giaretta, "Ontologies and knowledge bases: towards a terminological clarification," in *Towards very large knowledge bases* (N. Mars, ed.), pp. 25–32, Amsterdam: IOS Press, 1995.
- [19] M. Guninger and J. Lee, "Ontology applications and design," *Communications of the ACM*, vol. 45, no. 2, pp. 39–41, 2002.
- [20] E. Shaya, B. Thomas, and C. Cheung, "Specifics on a xml data format for scientific data," in *Astronomical Data Analysis Software and Systems X, ASP* (F. R. Harnden, Jr., F. A. Primini, and H. E. Payne, eds.), (San Francisco), 2001.
- [21] D. M. Sawyer, L. I. Reich, and S. Nihkinson, "Comparing emerging xml based formats from a multi-discipline perspective," in *American Geophysical Union, Fall Meeting 2002*, 2002.
- [22] P. Buneman, S. Khanna, and W. C. Tan, "Why and Where: A Characterization of Data Provenance," *ICDT*, pp. 316–330, 2001.
- [23] M. Greenwood, C. Goble, R. Stevens, J. Zhao, M. Addis, D. Marvin, L. Moreau, and T. Oinn, "Provenance of e-Science Experiments - experience from Bioinformatics," in *Proceedings of the UK OST e-Science second All Hands Meeting*, 2003.
- [24] R. Bose and J. Frew, "Lineage retrieval for scientific data processing: a survey," *ACM Comput. Surv.*, vol. 37, pp. 1–28, 2005.
- [25] Y. L. Simmhan, B. Plale, and D. Gannon, "A Survey of Data Provenance Techniques," tech. rep., Indiana University, Department of Computer Science, Aug. 2005.
- [26] I. T. Foster, J.-S. Vöckler, M. Wilde, and Y. Zhao, "Chimera: A Virtual Data System for Representing, Querying, and Automating Data Derivation," *SSDBM*, 2002.
- [27] J. D. Myers, T. C. Allison, S. Bittner, B. T. Didier, M. Frenklach, W. H. Green, Jr., Y.-L. Ho, J. C. Hewson, W. S. Koegler, C. Lansing, D. Leahy, M. Lee, R. McCoy, M. Minkoff, S. Nijsure, G. V. Laszewski, D. Montoya, C. M. Pancerella, R. Pinzon, W. Pitz, L. A. Rahn, B. Ruscic, K. Schuchardt, E. Stephan, A. Wagner, T. L. Windus, and C. L. Yang, "A Collaborative Informatics Infrastructure for Multi-scale Science," *CLADE*, 2004.
- [28] J. Zhao, C. Wroe, C. A. Goble, R. Stevens, D. Quan, and R. M. Greenwood, "Using Semantic Web Technologies for Representing E-science Provenance," *International Semantic Web Conference*, 2004.
- [29] A. Woodruff and M. Stonebraker, "Supporting Fine-grained Data Lineage in a Database Visualization Environment," *ICDE*, 1997.
- [30] J. Widom, "Trio: A System for Integrated Management of Data, Accuracy, and Lineage," *CIDR*, 2005.
- [31] J. Frew and R. Bose, "Earth System Science Workbench: A Data Management Infrastructure for Earth Science Products," in *Proceedings of the 13th International Conference on Scientific and Statistical Database Management*, 2001.
- [32] R. Bose and J. Frew, "Composing Lineage Metadata with XML for Custom Satellite-Derived Data Products," in *Proceedings of the 16th International Conference on Scientific and Statistical Database Management*, 2004.
- [33] J. Taylor and K. Worsley, "Inference for magnitudes and delays of responses in the FIAC data using BRAINSTAT/FMRISTAT," in *Human Brain Mapping*, 2006.
- [34] M. Fowler, *Patterns of Enterprise Application Architecture*. The Addison-Wesley Signature Series, Addison-Wesley, 2003.
- [35] G. Kiczales, J. Irwin, J. Lamping, J.-M. Loingtier, C. V. Lopes, C. Maeda, and A. Mendhekar, "Aspect-Oriented Programming," *ACM Computing Surveys*, vol. 28, p. 154, Dec. 1996.
- [36] K. De Volder and T. D'Hondt, *Meta-Level Architectures and Reflection: Second International Conference, Reflection'99, Saint-Malo, France*, vol. 1616, ch. Aspect-Oriented Logic Meta Programming, pp. 250–272. Springer Berlin / Heidelberg, Jan. 1999.
- [37] T. Nichols, "Guidelines for Presenting Neuroimaging Analyses." <http://www.sph.umich.edu/nichols/NIPub/>, 2005.
- [38] T. Picton, S. Bentin, P. Berg, E. Donchin, S. Hillyard, R. Johnson JR., G. Miller, W. Ritter, D. Ruchkin, M. Rugg, and M. Taylor, "Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria," *Psychophysiology*, vol. 37, no. 2, pp. 127–152, 2000.
- [39] A. Hunt and D. Thomas, *The Pragmatic Programmer: From Journeyman to Master*. Addison-Wesley, 2000.
- [40] B. Foote and W. F. Opdyke, "Lifecycle and refactoring patterns that support evolution and reuse," in *Pattern Languages of Program Design* (J. Coplien and D. Schmidt, eds.), Reading, Mass: Addison-Wesley, 1995.
- [41] M. Fowler, *Refactoring: Improving the Design of Existing Code*. The Addison-Wesley Object Technology Series, Addison-Wesley, 2000.