



Published in final edited form as:

*Epilepsia*. 2012 July ; 53(Suppl 2): 28–32. doi:10.1111/j.1528-1167.2012.03556.x.

## From Classification to Epilepsy Ontology and Informatics

Guo-Qiang Zhang, PhD\*, Satya S Sahoo, PhD\*, and Samden D Lhatoo, MD, FRCP (Lon)<sup>†</sup>

\*Division of Medical Informatics, Case Western Reserve University and University Hospitals Case Medical Center Cleveland, OH

<sup>†</sup>Department of Neurology, Case Western Reserve University and University Hospitals Case Medical Center Cleveland, OH

### Summary

The 2010 International League Against Epilepsy (ILAE) classification and terminology commission report proposed a much needed departure from previous classifications to incorporate advances in molecular biology, neuroimaging, and genetics. It proposed an interim classification and defined two key requirements that need to be satisfied. The first is the ability to classify epilepsy in dimensions according to a variety of purposes including clinical research, patient care, and drug discovery. The second is the ability of the classification system to evolve with new discoveries. Multi-dimensionality and flexibility are crucial to the success of any future classification. In addition, a successful classification system must play a central role in the rapidly growing field of epilepsy informatics. An epilepsy ontology, based on classification, will allow information systems to facilitate data-intensive studies and provide a proven route to meeting the two foregoing key requirements. Epilepsy ontology will be a structured terminology system that accommodates proposed and evolving ILAE classifications, the NIH/NINDS Common Data Elements, the ICD systems and explicitly specifies all known relationships between epilepsy concepts in a proper framework. This will aid evidence based epilepsy diagnosis, investigation, treatment and research for a diverse community of clinicians and researchers. Benefits range from systematization of electronic patient records to multi-modal data repositories for research and training manuals for those involved in epilepsy care. Given the complexity, heterogeneity and pace of research advances in the epilepsy domain, such an ontology must be collaboratively developed by key stakeholders in the epilepsy community and experts in knowledge engineering and computer science.

### Keywords

Classification; informatics; ontology

### Introduction

Epilepsy is heterogeneous in etiology, pathophysiology, semiology and a variety of other parameters. Classification of the epilepsies has proven to be a complex (Berg, 2011, Berg and Scheffer, 2011, Shorvon, 2011, Panayiotopoulos, 2011, Engel, 2011, Luders et al. 2012), often controversial and sometimes thankless undertaking. Henri Gastaut proposed the

---

Correspondence to: Samden Lhatoo, MD, FRCP (Lon), Case Western Reserve University and University Hospitals Case Medical Center, 11100 Euclid Avenue, Cleveland, OH 44106, samden.lhatoo@uhhospitals.org, Phone 216-8445855, Fax 216-8443160.

#### Disclosures:

None of the authors have any conflicts of interest to disclose.

We confirm that we have read the Journal's position on issues involved in ethical publications and affirm that this report is consistent with those guidelines. No undisclosed groups or persons have had a primary role in the preparation of this manuscript.

first formal epilepsy classification system in 1969, which was updated by the Classification and Terminology Commission of the International League Against Epilepsy (ILAE) in 1981 (Commission, 1981) and in 1989 (Commission, 1989). However, there is broad agreement that existing classification systems are inadequate since major recent advances in molecular genetics and electrophysiology are not incorporated and several epilepsy syndromes are excluded. In 2010, the ILAE Commission put forth two key requirements for the creation of a flexible, multi-dimensional, and extensible classification system that (Berg, 2010):

- Adapts to the evolving understanding of epilepsy in the context of advances in “epidemiology, electrophysiology, imaging, developmental neurobiology, genetics, systems neurobiology, and neurochemistry,”
- Allows dynamic classification of epilepsy along the appropriate dimensions or features as required by different applications (e.g. drug discovery, clinical research, patient care, training and education).

With the burgeoning quantities of multi-modal clinical and electrophysiological data produced by epilepsy centers worldwide, our current approaches often result in missed opportunities for data-driven research that can potentially address many unanswered questions in epilepsy. A new epilepsy informatics paradigm should impact a range of clinical and research areas, including epilepsy phenomenology, etio-pathogenesis, drug development and gene discovery. An essential first step to accomplish this objective is to develop an epilepsy ‘ontology’ for modeling epilepsy concepts which will serve as a basis for such a new paradigm. It will be a framework with a multi-faceted, multi-dimensional, dynamic environment, supported by a set of existing tools for authoring, sharing, and quality assurance. Incorporating knowledge and development of other ontologies such as gene, anatomical and neuro-electromagnetic ontologies, epilepsy ontology has great potential to flexibly accommodate changing new knowledge in epilepsy and to provide a core resource for the evolving world of epilepsy informatics.

## What is Ontology?

An ontology is an organized and coherent structure of domain knowledge, represented in a formal, logic-based language, which reduces terminological heterogeneity, facilitates data interoperability and enables knowledge discovery (Cimino and Zhu, 2006) (Bodenreider, 2010, Bodenreider, 2008). Ontologies can be used seamlessly as components of information systems and data management tools. They represent not only the concepts/classes used in scientific work, but just as importantly, the relationships between the concepts/classes. Such relationships help determine the semantics of a concept in the context of other concepts. For example, the term “generalized 3–4 hertz spike wave” is an electroencephalographic description of an ictal or inter-ictal epileptic phenomenon. The term however, is also used in literature to describe or typify a group of genetic generalized epilepsies. These epilepsies may be linked not just through electrophysiological similarities but through phenotype (e.g. typical absence seizures) and conceivably genotype. In large database or electronic medical record query management, definition of such terms and linking their relationships can allow tailored data mining to “lump” or “split” according to the clinician or researcher’s questions. Ontologies have thus become a central component in biomedical information management. Familiar in their role in supporting application menus similar to those generated by MeSH Headings, ontologies are also becoming valuable for designing intuitive and novel interfaces to query, access, and visualize large sets of distributed biomedical datasets (Zhang et al., 2010a).

Researchers increasingly rely on biomedical ontologies as critical resources throughout their experimental work flows. In the last decade, 5437 publications indexed by PubMed involve the use of ontologies, of which nearly 1000 contain the key word “ontology” in the title. The

BioPortal resource at the National Center for Biomedical Ontologies (NCBO) lists nearly 300 ontologies consisting of 5.3M terms used in a range of biomedical informatics applications from bench experiments (Ashburner, 2000) (Sahoo, 2009) to patient care at the bedside (Zhang, 2010a).

The complexity and heterogeneity of epilepsy therefore begs the creation of an epilepsy ontology, which if successful, will serve as a core resource for epilepsy informatics. Such an ontology will enhance rather than straitjacket approaches to classification by accommodating diverse terminologies and nomenclature. Any impasse can be gracefully managed, as the ontology framework is open and inclusive rather than exclusive. Because it is designed to accommodate, the outcome will benefit the entire field, with involvement of stakeholders varying according to need.

There are three main challenges addressed by an epilepsy ontology with corresponding advantages if each are met. These are:

### Challenge 1 – Enabling Terminology and Classification

The International Classification of Diseases (ICD-9, 10 and the impending version 11 (Bergen et al., 2012), the ILAE classifications, the ILAE's recommendations on standardization of epidemiological studies and surveillance in epilepsy (Thurman et al., 2011) and the NINDS Common Data Elements (CDE) project (Loring et al., 2011) all have the common goal of providing usable, reliable, reproducible and standardized epilepsy diagnoses and terminologies. This is done, however, in the near absence of agreed-upon, standardized terminology and concepts. Creation of a common language using an informatics approach allows a move away from restrictive, paper-based systems to a computer-based paradigm that allows cross-talk between diverse approaches to classification. Extending Hughlings-Jackson's depiction of the empirical and scientific approaches to the classification of animals and plants (the hunter versus the zoologist, the gardener versus the botanist (Hughlings-Jackson, 1888), a simple modern analogy is that of a book purchase on amazon.com. The Count of Monte Cristo is in French or English, in Classic Literature, Fiction, Teen Books, Historical Fiction, Historical Fantasy or under "Dumas", the author. All comprise methods of classification, depending on the search approach but are impossible to reproduce in a physical book shop where the book can usually only be placed in one section on a single shelf. In the digital world, access to the book can be achieved through multiple angles and this computerized digital content access, unlike physical or printed media, has unlimited potential for us to take advantage of. Of course, the complexity of the epilepsy domain deserves great respect, but this is not an insurmountable problem.

Very importantly, in large parts of the world, professionals involved in the care of patients with epilepsy range from tertiary care epileptologists in urban settings to primary care physicians and health workers in rural areas of low-income countries. The latter may comprise individuals of diverse socio-medical backgrounds and levels of training who use different epilepsy terms/diagnoses across many different settings. But, they still require access to and communication with the wider epilepsy community. Hence, what is a focal dyscognitive seizure to one domain expert may remain a complex partial seizure for some time yet for another epilepsy professional, just as the terms dialeptic seizure or automotor seizure may for yet another practitioner. Large database research, particularly when it comes to annotated nomenclature, stands to suffer in this manner when data is contributed by a wide variety of centers and countries but cannot be harmonized according to true intended semantics.

In addition, epilepsy research encompasses a wide range of sub-domains that involve animal as well as human subjects. Improving accessibility of modern knowledge, research methodologies and expertise to all end-users is a goal that can be made realistic through modern informatics technology. This technology can be used to leverage epilepsy informatics and data resources such as clinical or research databases, diagnostic manuals, and teaching resources tailored to the sophistication of the end-user.

## Challenge 2 – Incorporation of Existing Terminological Resources

Clinical and experimental epilepsy encompasses a wide range of sub-domains. Any epilepsy informatics resource has to incorporate epilepsy terminology extensively as well as include related terminologies that do not directly fall under the remit of epilepsy but are integral to practice and research. These include neuroanatomical terms, EEG, MEG and MRI nomenclature, and genetic and proteomics terminology. Many of these allied domains have already created standardized terminologies, including gene-related terminology standardized in the Gene Ontology (GO) (Ashburner, 2000), anatomical features in the Foundational Model of Anatomy (FMA) (Rosse, 2003), and electrophysiological concepts in the Neural ElectroMagnetic Ontologies (NEMO) (Dou, 2007). Others, such as neuroimaging, have not yet been developed. Due to the current lack of an epilepsy ontology, investigators have been unable to leverage these existing standards for use in their respective fields in clinical and experimental epilepsy. Doing so is likely to significantly enrich epilepsy ontology as well as integrate advances in genetics and imaging for example into the epilepsy domain. Gene Ontology is particularly important as the impact of genetics on the field is likely to greatly increase in the future.

## Challenge 3 – Technological Challenges of Multi-modal Epilepsy Data

Modern epilepsy practice uses sophisticated imaging, modern neurophysiological techniques (including digital EEG, EEG source imaging and magnetoencephalography) and polygraphic data acquisition in the epilepsy monitoring unit (oximetry, capnography, sleep etc). Usage and sharing of such data, particularly in large scale research, demands major informatics infrastructure and expertise. In the current era of cloud computing, easy and efficient collaboration across international boundaries, even using large neurophysiological and imaging datasets is now a realistic possibility. The International Epilepsy Electrophysiology Portal of the International Collaborative Seizure-Prediction Group is a model for such collaboration (<http://braintrust.seas.upenn.edu>), with a focus mainly on intracranial EEG and seizure prediction. Multicenter collaboration has also been successfully achieved in the sleep domain as result of a major effort by experts in sleep medicine, knowledge engineering, informatics and computer science. In the latter, successful development of ontology has been central to this success. Ontology is increasingly recognized as the key to drive data capture, data search and query, and data integration in research involving alphanumeric as well as electrophysiological signal data.

## A Roadmap for Epilepsy Ontology and Informatics

Development of an epilepsy ontology involves several sequential steps requiring close collaboration between informatics experts and epilepsy domain experts. There is firstly an exhaustive and comprehensive listing of controlled vocabulary using terms from all existing and proposed classification systems, ICD systems, NINDS CDE and epilepsy literature. This is then extended and enriched with terms and concepts from all related terminological systems that are integral to the epilepsy domain. These include already advanced systems such as Gene Ontology and the Foundational Model of Anatomy. The latter for example, provides detailed terminology and relationships for brain structures involved in epileptogenesis, symptoms and treatment (epilepsy surgery, deep brain stimulation). A

standard set of ontological relationships is then used to link together all these terms in a formal logic language (e.g. Web Ontology Language (OWL) that the informatician uses to create a rich epilepsy domain ontological structure that can be used for databases, electronic medical records, diagnostic manuals and other digital applications with diverse utility.

The Gene Ontology Consortium (The Gene Ontology Consortium, 2012) is a model for the development of an epilepsy ontology which can drive the subsequent development of epilepsy informatics infrastructures. Funded by the National Human Genome Research Institute, the project has three structured ontological components that describe gene products in terms of their 1) associated biological processes, 2) cellular components and 3) molecular functions in a species-independent manner. The project effort itself has three aspects: first, the development and maintenance of the ontologies themselves; second, the annotation of gene products, which entails making associations between the ontologies and the genes and gene products in the collaborating databases; and third, development of tools that facilitate the creation, maintenance and use of ontologies. An attractive and innovative feature of the Consortium's model which is likely to be highly relevant to the epilepsy field is the open and interactive nature of the project where comments and suggestions are solicited for incorporation.

As epilepsy ontology takes its shape, benefits can be derived by plugging them into software and informatics tools such as PhysioMIMI (<http://physiomimi.case.edu>). Physio-MIMI enhances the efficiency of whole data-integration, data access and data exploration life-cycle, leveraging ontology to directly drive the federated query interface VISAGE (VISual Aggregator and Explorer [Zhang, 2010b]), and PhysioMap, the database-to-ontology mapper. This architecture makes use of ontology beyond its traditional role for terminology standardization, resulting in a flexible framework with a domain ontology, such as the epilepsy ontology, as a "plug-and-play" component capable of harmonizing disparate data sources without requiring adherence to a uniform data model.

Epilepsy ontology however, has different challenges and a broad collaborative framework is required with the endorsement and support of adult and pediatric epileptologists, neurophysiologists, general neurologists, epilepsy researchers, epidemiologists and other professionals involved in clinical and research epilepsy. Experts in knowledge engineering, informatics and computer sciences are crucial to success as is investment by the appropriate funding bodies.

## Acknowledgments

The authors' epilepsy ontology development work is supported by NIH/NINDS (Grant P20-NS076965 – 01–The PRISM Project)

## References

- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet.* 2000; 25:25–9. [PubMed: 10802651]
- Berg AT. Classification and epilepsy: the future awaits. *Epilepsy Curr.* 2011; 11:138–40. [PubMed: 22016666]
- Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, Van Emde Boas W, Engel J, French J, Glauser TA, Mathern GW, Moshé SL, Nordli D, Plouin P, Scheffer IE. Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005–2009. *Epilepsia.* 2010; 51:676–685. [PubMed: 20196795]

- Berg AT, Scheffer IE. New concepts in classification of the epilepsies: entering the 21st century. *Epilepsia*. 2011; 52:1058–62. [PubMed: 21635233]
- Bergen DC, Beghi E, Medina M. Revising the ICD-10 Codes for Epilepsy and Seizures. *Epilepsia*. 2012 (THIS ISSUE).
- Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support. *Yearb Med Inform*. 2008:67–79. [PubMed: 18660879]
- Bodenreider, O. Technical report. Bethesda: Lister Hill National Center for Biomedical Communications, National Library of Medicine; 2010. Quality assurance in biomedical terminologies and ontologies.
- Cimino JJ, Zhu X. The practical impact of ontologies on biomedical informatics. *Yearb Med Inform*. 2006:124–35. [PubMed: 17051306]
- Commission of Classification and Terminology of the International League Against Epilepsy. Proposal for revised clinical and electrographic classification of epileptic seizures. *Epilepsia*. 1981; 22:489–501. [PubMed: 6790275]
- Commission of Classification and Terminology of the International League Against Epilepsy. Proposal for revised classification of epilepsies and epileptic syndromes. *Epilepsia*. 1989; 30:389–399. [PubMed: 2502382]
- Dou, D.; Frishkoff, G.; Rong, J.; Frank, R.; Malony, A.; Tucker, D. Development of NeuroElectroMagnetic Ontologies (NEMO): A framework for mining brain wave ontologies. Thirteenth International Conference on Knowledge Discovery and Data Mining (KDD2007); San Jose, CA. 2007. p. 270-279.
- Engel J Jr. The etiologic classification of epilepsy. *Epilepsia*. 2011; 52:1195–7. discussion 1205–9. [PubMed: 21671927]
- Hughlings-Jackson J. Remarks on the diagnosis and treatment of diseases of the brain. *BMJ*. 1888; ii: 58–62.
- Loring DW, Lowenstein DH, Barbaro NM, Fureman BE, Odenkirchen J, Jacobs MP, Austin JK, Dlugos DJ, French JA, Gaillard WD, Hermann BP, Hesdorffer DC, Roper SN, Van Cott AC, Grinnon S, Stout A. Common data elements in epilepsy research: development and implementation of the NINDS epilepsy CDE project. *Epilepsia*. 2011; 52:1186–91. [PubMed: 21426327]
- Luders HO, Amina S, Baumgartner C, Benbadis S, Bermeo-Ovalle A, Devreux M, Diehl B, Edwards J, Baca-Vaca GF, Hamer H, Ikeda A, Kaiboriboon K, Kellinghaus C, Koubeissi M, Lardizabal D, Lhatoo S, Luders J, Mani J, Mayor LC, Miller J, Noachtar S, Pestana E, Rosenow F, Sakamoto A, Shahid A, Steinhoff BJ, Syed T, Tanner A, Tsuji S. Modern technology calls for a modern approach to classification of epileptic seizures and the epilepsies. *Epilepsia*. 2012
- Panayiotopoulos CP. The new ILAE report on terminology and concepts for organization of epileptic seizures: a clinician's critical view and contribution. *Epilepsia*. 2011; 52:2155–60. [PubMed: 22004554]
- Rosse C, Mejina JLV Jr. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *Journal of Biomedical Informatics*. 2003; 36:478–500. [PubMed: 14759820]
- Sahoo, SS.; Weatherly, DB.; Muttharaju, R.; Anantharam, P.; Sheth, A.; Tarleton, RL. Ontology-driven Provenance Management in eScience: An Application in Parasite Research. In: Meersman, TDeaR, editor. The 8th International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE 09). Vilamoura, Algarve-Portugal: Springer Verlag; 2009. p. 992-1009.
- Shorvon SD. The etiologic classification of epilepsy. *Epilepsia*. 2011; 52:1052–7. [PubMed: 21449936]
- Thurman DJ, Beghi E, Begley CE, Berg AT, Buchhalter JR, Ding D, Hesdorffer DC, Hauser WA, Kazis L, Kobau R, Kroner B, Labiner D, Liow K, Logroscino G, Medina MT, Newton CR, Parko K, Paschal A, Preux PM, Sander JW, Selassie A, Theodore W, Tomson T, Wiebe S. Standards for epidemiologic studies and surveillance of epilepsy. *Epilepsia*. 2011; 52(Suppl 7):2–26. [PubMed: 21899536]
- The Gene Ontology Consortium: Enhancements for. *Nucleic Acids Res*. 2011; 40:D559–64. [PubMed: 22102568]



- Zhang GQ, Siegler T, Saxman P, Sandberg N, Mueller R, Johnson N, Hunscher D, Arabandi S. VISAGE: A Query Interface for Clinical Research. AMIA Summits Transl Sci Proc. 2010a; 2010:76–80. [PubMed: 21347154]
- Zhang, GQ.; Siegler, T.; Saxman, P.; Sandberg, N.; Mueller, R.; Johnson, N.; Hunscher, D.; Arabandi, S. AMIA Clinical Research Informatics Summit. San Francisco: 2010b. VISAGE: A Query Interface for Clinical Research; p. 76-80.