

# The Epilepsy Ontology: a community-based ontology tailored for semantic interoperability and text-mining

**Astghik Sargsyan**

Fraunhofer SCAI <https://orcid.org/0000-0002-5860-6369>

**Philipp Wegner**

Fraunhofer SCAI <https://orcid.org/0000-0003-1159-9269>

**Stephan Gebel**

Fraunhofer SCAI <https://orcid.org/0000-0001-5086-8324>

**Shounak Baksi**

Causality Biomodels <https://orcid.org/0000-0001-7225-5377>

**Geena Mariya Jose**

Causality Biomodels

**Vanessa Lange-Rupprecht**

Fraunhofer SCAI <https://orcid.org/0000-0003-1374-8184>

**Johannes Darms**

ZB Med

**Bruce Schultz**

Fraunhofer SCAI <https://orcid.org/0000-0003-4492-9463>

**Jürgen Klein**

Fraunhofer SCAI

**Marc Jacobs**

Fraunhofer SCAI <https://orcid.org/0000-0003-4153-3930>

**Philippe Ryvlin**

University Hospital CHUV <https://orcid.org/0000-0001-7775-6576>

**Sumit Madan**

Fraunhofer SCAI

**Martin Hofmann-Apitius**

Fraunhofer SCAI <https://orcid.org/0000-0001-9012-6720>

**Alpha Tom Kodamullil** (✉ [alpha.tom.kodamullil@scai.fraunhofer.de](mailto:alpha.tom.kodamullil@scai.fraunhofer.de))

Fraunhofer SCAI <https://orcid.org/0000-0001-9896-3531>

**Keywords:** Epilepsy, Ontology, Text Mining, Disease Ontology, BFO, ILAE

**Posted Date:** January 14th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1259791/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# The Epilepsy Ontology: a community-based ontology tailored for semantic interoperability and text-mining

Astghik Sargsyan<sup>1,2†</sup>, Philipp Wegner<sup>1†</sup>, Stephan Gebel<sup>1</sup>, Shounak Baksi<sup>3</sup>, Geena Mariya Jose<sup>3</sup>, Vanessa Lage-Rupprecht<sup>1</sup>, Johannes Darms<sup>4</sup>, Bruce Schultz<sup>1</sup>, Jürgen Klein<sup>1</sup>, Marc Jacobs<sup>1</sup>, Philippe Ryvlin<sup>5,6</sup>, Sumit Madan<sup>1</sup>, Martin Hofmann-Apitius<sup>1,2</sup>, Alpha Tom Kodamullil<sup>1,3\*</sup>

<sup>1</sup> Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), 53754 Sankt Augustin, Germany, <sup>2</sup> Bonn-Aachen International Center for Information Technology (B-IT), University of Bonn, 53113 Bonn, Germany, <sup>3</sup>Causality Biomodels, Kinfra Hi-Tech Park, Kalamassery, Cochin, Kerala 683503, India, <sup>4</sup>German National Library of Medicine (ZBMED) - Information Centre for Life Sciences, 53115 Bonn, Germany, <sup>5</sup> Department of Clinical Neurosciences, University Hospital CHUV, Lausanne, Switzerland, <sup>6</sup>Department of Neurology and Acute Neurorehabilitation Unit, University Hospital CHUV, Lausanne, Switzerland <sup>7</sup>Institute of Computer Science, University of Bonn, 53115 Bonn, Germany

\*corresponding author

† These authors contributed equally to this work.

## Abstract

**Motivation:** Epilepsy is a multi-faceted complex disorder that requires a precise understanding of the classification, diagnosis, treatment, and disease mechanism governing it. Although scattered resources are available on epilepsy, comprehensive and structured knowledge is missing. In contemplation to promote multidisciplinary knowledge exchange and facilitate advancement in clinical management, especially in pre-clinical research, a disease-specific ontology is necessary. The presented ontology is designed to enable better interconnection between scientific community members in the epilepsy domain.

**Results:** The Epilepsy Ontology (EPIO) is an assembly of structured knowledge on various aspects of epilepsy, developed according to Basic Formal Ontology (BFO) and Open Biological and Biomedical Ontology (OBO) Foundry principles. Concepts and definitions are collected from the latest International League against Epilepsy (ILAE) classification, domain-specific ontologies, and scientific literature. This ontology consists of 1,879 classes and 28,151 axioms (2,171 declaration axioms, 2,219 logical axioms) from several aspects of epilepsy. This ontology is intended to be used for data management and text mining purposes.

**Keywords:** Epilepsy, Ontology, Text Mining, Disease Ontology, BFO, ILAE

**Contact:** alpha.tom.kodamullil@scai.fraunhofer.de

**Supplementary information:** Supplementary data can be found below.

## 1 Introduction

Epilepsy is a common chronic neurological disorder, affecting more than 60 million individuals worldwide, and characterized by an enduring predisposition to generate epileptic seizures that might be associated with severe cognitive, psychological and social issues. Epileptic seizures reflect the transient occurrence of various neurological signs (motor, behavioral) and/or symptoms (cognitive, emotional, sensory), often including loss of awareness. Such seizures differ widely in severity, appearance, cause, lasting consequences, and management depending on the type<sup>2</sup>. Despite all available antiepileptic treatments, up to one third of patients still suffer from uncontrolled seizures. Hence, the management of epilepsy has become a critical challenge for physicians, due in part to the difficulty in properly classifying the form of epilepsy being treated<sup>3</sup>. To aid healthcare workers in this, The International League Against Epilepsy (ILAE) developed a classification system for determining the type of epilepsy affecting an individual, and serves as a critical tool for practicing physicians to better understand these disease variants and their underlying mechanisms<sup>4</sup>. However, in order to ensure that this classification system remains as accurate as possible, it must be updated regularly with the latest findings and reports on epilepsy. Therefore, there exists a need to create an aggregated resource of all available knowledge on epilepsy that is also compatible with the demands of

clinicians, scientists, patients, and industry as well as to serve as a means for accurate data collection and semantic integration. Ontologies are commonly used as semantic frameworks in biomedical and clinical knowledge disciplines for addressing these types of data management challenges. Disease-specific ontologies, such as the Cardiovascular Disease Ontology (CVDO)<sup>5</sup>, incorporate disease-specific concepts and terminologies thus providing a platform for more in-depth research. The lack of a standard, specific ontology describing clinical and experimental epilepsy is a hindering detriment for both physicians and researchers. Development of a comprehensive epilepsy disease ontology that can accommodate diverse aspects, terminologies, and well-defined vocabularies can serve as a core resource for epilepsy data and knowledge-driven research. In this paper, we describe an integrated epilepsy ontology (EPIO), which was developed in a community-based manner by various stakeholders from both Europe and the United States. The ontology was developed based on the ILAE classification of epilepsy as well as by integrating already existing epilepsy-related ontologies such as those developed according to BFO 2.0 classes and OBO Foundry principles. Many of the existing ontologies focus more on specific aspects pertaining to epilepsy, while in this work we sought to build a more generalized ontology that describes the multiple facets of epilepsy. Because the EPIO includes clinical, experimental, and theoretical aspects of epilepsy, it benefits a wider range of individuals in comparison to the more focused available ontologies.

## 2 Materials and methods

Epilepsy-specific concepts were added under the upper level classes of BFO. Protégé OWL editor (version 5.5.0) was used to build the epilepsy ontology using the ontology web language (OWL) format (<https://protege.stanford.edu/>). The initial version of the ontology was composed of 285 concepts with definitions from the ILAE (<https://www.epilepsydiagnosis.org/>). It was further enhanced using terms from other Epilepsy related ontologies such as EPSO (<https://bioportal.bioontology.org/ontologies/EPSO>)<sup>6</sup>, ESSO (<https://bioportal.bioontology.org/ontologies/ESSO>), EPILONT (<https://bioportal.bioontology.org/ontologies/EPILONT>). Furthermore, the preclinical aspects of epilepsy include preclinical disease models, cellular and molecular disease mechanisms. Knowledge of these aspects was curated from available epilepsy-related ontologies and research articles available on PubMed. Epilepsy-associated concepts such as genes, miRNA, SNPs were not included in the ontology because of the ambiguity of the evidence associated with these. The EPIO was constructed based on guidelines and principles defined by OBO Foundry (<http://www.obofoundry.org/>). The Ontofox tool (<http://ontofox.hegroup.org/>) was used to integrate concepts from other OBO ontologies.

## 3 Results

The ontology's structure was designed to be accessible for all users. Concepts, as well as the major classes added from the ILAE classification, were classified using the BFO 2.0 higher-order classes for proper interoperability. Multiple clinical and non-clinical sources were parsed in order to gather knowledge from epilepsy-related domains, resulting in an ontology which consists of 1,879 classes and 28,151 axioms (2,171 declaration axioms, 2,219 logical axioms). From the clinical domain, diagnoses, clinical screening tests, and treatment-related concepts were collected and added to the EPIO. Etiological, biological mechanisms, anatomical, cellular and molecular causes and effects, and preclinical disease models comprise the non-clinical epilepsy information in the ontology.

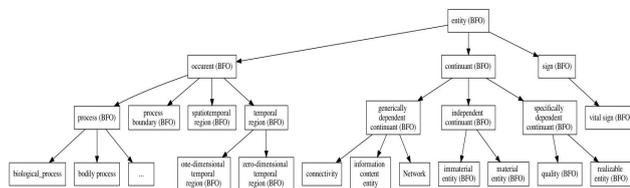


Figure 1: High-level structure of EPIO

**Text mining Applications of Epilepsy Ontology:** The EPIO presented here acts as the semantic layer for the text mining-based knowledge discovery software *Neuro SCAIView* (<https://neuro.scaiview.com/>). Here, we would like to illustrate some of the specific applications of *Neuro SCAIView* using the EPIO (more details on use cases are included in the Supplementary document).

- Filtering relevant documents in the epilepsy context:** Using EPIO, we can filter and further process domain-specific documents. Specific filters provide information on anatomical entities, cellular processes, diagnoses, disease classifications, imitators, syndromes, etiology and risk factors in the context of epilepsy.
- Data mining/information extraction of metadata from epilepsy-related articles:** Finding the most mentioned brain region in a certain search context, then gene-specific analysis to generate a list of genes mentioned in the specific search context were applied. In addition, the EPIO, in combination with DrugBank, was used as a semantic layer to search and retrieve the epilepsy related drugs mentioned in the literature.
- Grouping epilepsy articles into specific topics:** The combinatorial search using epilepsy related concepts and filters allowed for a targeted search for information and the acquisition of new knowledge based on text analytic methods that may contribute to future research.

## 4 Discussion

We have integrated in-depth knowledge from several sources and described many crucial aspects of the disease important to the clinical

and scientific communities. From the clinician's perspective, modern intelligent systems backed up by an extensive disease specific ontology like EPIO, may help in diagnosing and identifying the patient's epilepsy type. The presented ontology can incorporate screening tests and the clinical history of the patient which can be utilized to gather the health record of the patient, thus serving as a useful tool by aiding clinicians in planning the treatment regimen for patients. From the researcher's point of view, this ontology could provide all available information on associated molecular mechanisms, cell signaling pathways, preclinical models, as well as the clinical details of the disease. In conclusion, the EPIO represents a substantial effort to semantically assemble the knowledge of the complex and multidimensional disease known as epilepsy. Due to the rapid growth of information and research in the field of this disease, the knowledge contained within this ontology will never be complete, and thus continuous enrichment is required. The EPIO was made publicly available so that members of the scientific community can submit open-sourced updates to it, thus helping to ensure the ontology stays as current as possible.

## Abbreviations

**OWL:** Web Ontology Language  
**ILAE:** International League Against Epilepsy  
**EPIO:** Epilepsy Ontology  
**BFO:** Basic Formula Ontology  
**OBO:** Open Biological Ontology Foundry

## Acknowledgments

The authors thank Thrishna KG, Abish K, and Priya S from Causality Biomodels for the ontology curation work.

## Funding

This work was supported and funded from the EU's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2 by the HBP Medical Informatics Platform).

## Competing interests

None declared.

## Availability of data and materials

The current release of the ontology is publicly available under a Creative Commons 4.0 License and shared via <http://purl.obolibrary.org/obo/epso.owl> and is a community-based effort to assemble various facets of the complex disease epilepsy. The ontology is also deposited in BioPortal at <https://bioportal.bioontology.org/ontologies/EPIO>.

## References

- Falco-Walter J. Epilepsy-Definition, Classification, Pathophysiology, and Epidemiology. *Semin Neurol*. 40(6), 617-623 (2020);
- Oto, M.M., The misdiagnosis of epilepsy: Appraising risks and managing uncertainty. *Seizure* 44, 143-146 (2017).
- Fisher, R.S. et al., ILAE official report: a practical clinical definition of epilepsy. *Epilepsia* 55 (4), 475-482 (2014).
- Arguello Casteleiro, M. et al., Deep learning meets ontologies: experiments to anchor the cardiovascular disease ontology in the biomedical literature. *J Biomed Semantics* 9 (1), 13 (2018).
- Sahoo, S.S. et al., Epilepsy and seizure ontology: towards an epilepsy informatics infrastructure for clinical research and patient care. *J Am Med Inform Assoc* 21 (1), 82-89 (2018).



## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [EpilepsyOntSupplementaryfile.pdf](#)