Semantic Annotation of Data on Neurodegenerative Diseases in Patients Using Ontologies

Panče Panov Ana Kostovska Ilin Tolovski Sašo Džeroski
Jožef Stefan Institute & Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

Abstract

In this work we propose a mid-level ontology for representing various types of data on patients with neurodegenerative diseases. The proposed ontology describes the definition of the diseases as well as the whole process of diagnosing a patient through his/her visits, examinations and medical history. By defining the ontology in this way, we will further use it in semantic annotation of datasets that contain different diagnostic data (clinical, imaging, bio-marker data etc.) about neurodegenerative diseases and their progression in patients collected by different hospitals. Having an ontology describing data on patients with neurodegenerative diseases is important from two different perspectives:

- From a viewpoint of ontology-based data access, it would allow federation queries on data produced and stored at different hospitals.
- From a viewpoint of data analytics it would allow semi-automatic creation of data analysis workflows based on the datatypes that occur in the datasets, annotated with terms defined or used in the proposed ontology.

Definition of terms

- Ontology is a term that originates from philosophy where it represents a branch of metaphysics dealing with the nature of being.
- In computer science an ontology is defined as a set of concepts and categories in a certain subject area or domain that shows their properties and relations between them.
- Neurodegenerative diseases represent a heterogeneous group of disorders that are characterized by the progressive degeneration of the structure and function of the central or peripheral nervous system.
- Alzheimer's disease is a irreversible neurodegenerative disease that results in a loss of mental function due to the deterioration of brain tissue.
- Parkinson's disease is a disorder of the central nervous system that is a result from the loss of cells in various parts of the brain, categorized as a movement disorder.

Neurodegenerative Diseases Data Ontology (NDDO)

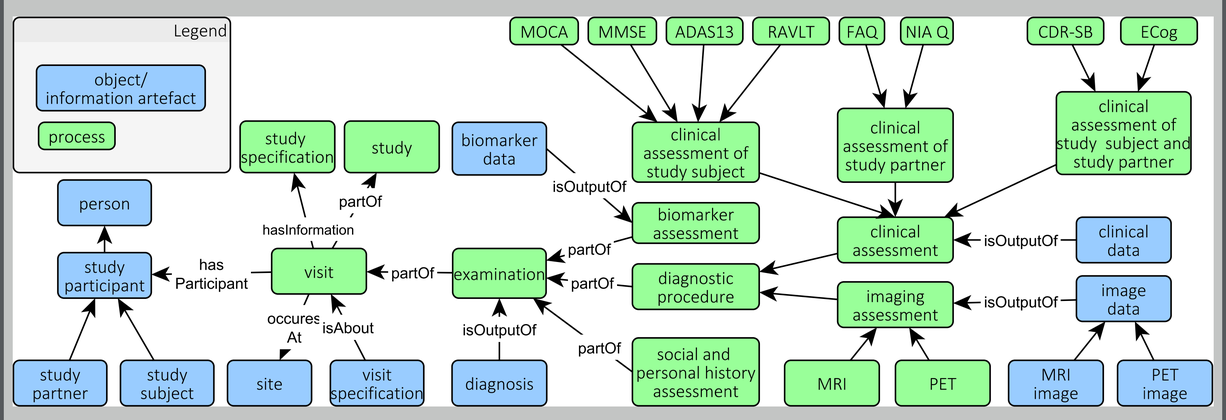

Figure 1: Scheme of the Neurodegenerative Diseases Data Ontology

Use case: Describing data on patients included in the PPMI study (Parkinson's disease)

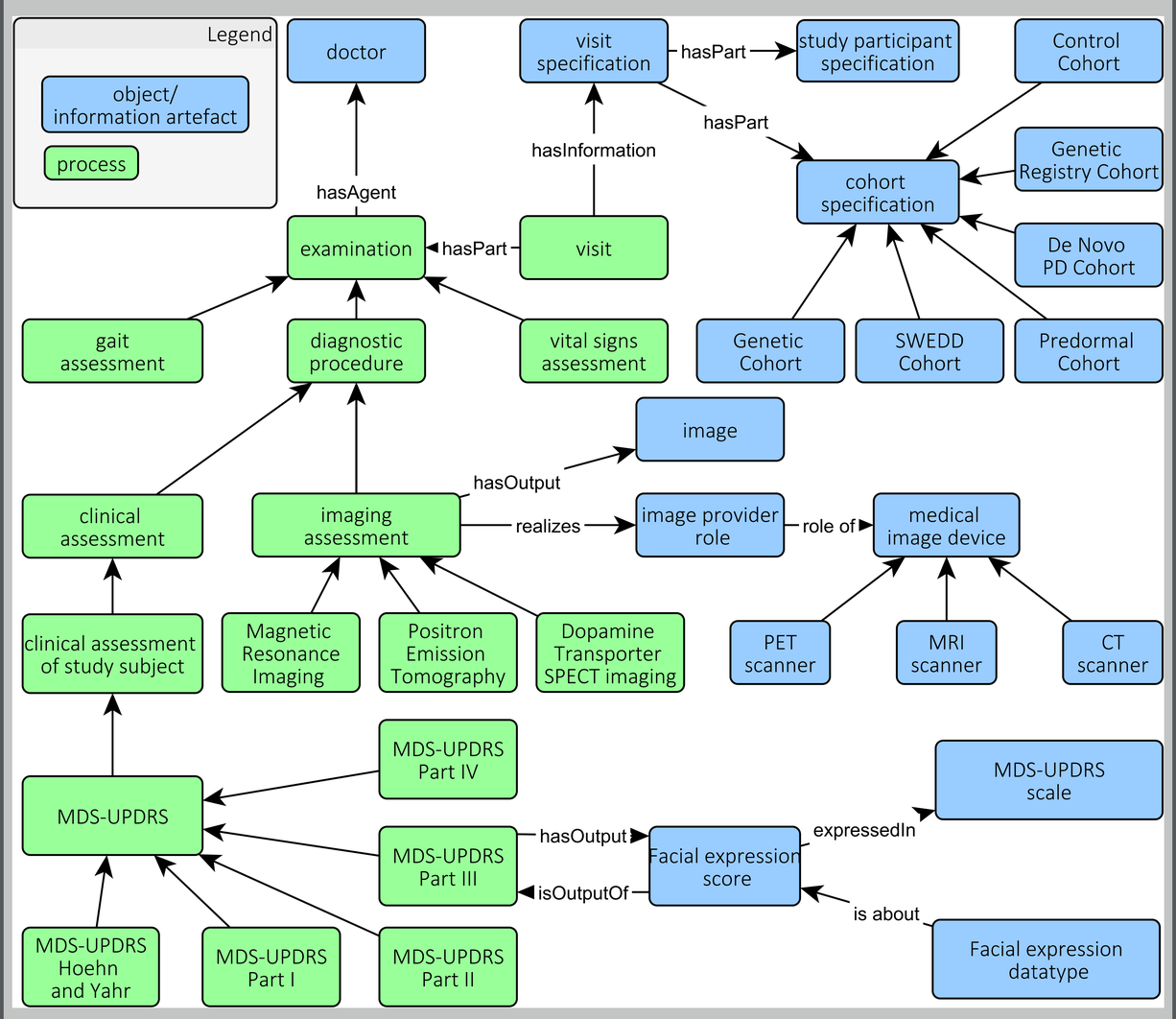

Figure 2: Scheme of the part of the ontology describing Parkinson's disease

Use case: Describing data on patients included in the ADNI study (Alzheimer's disease)

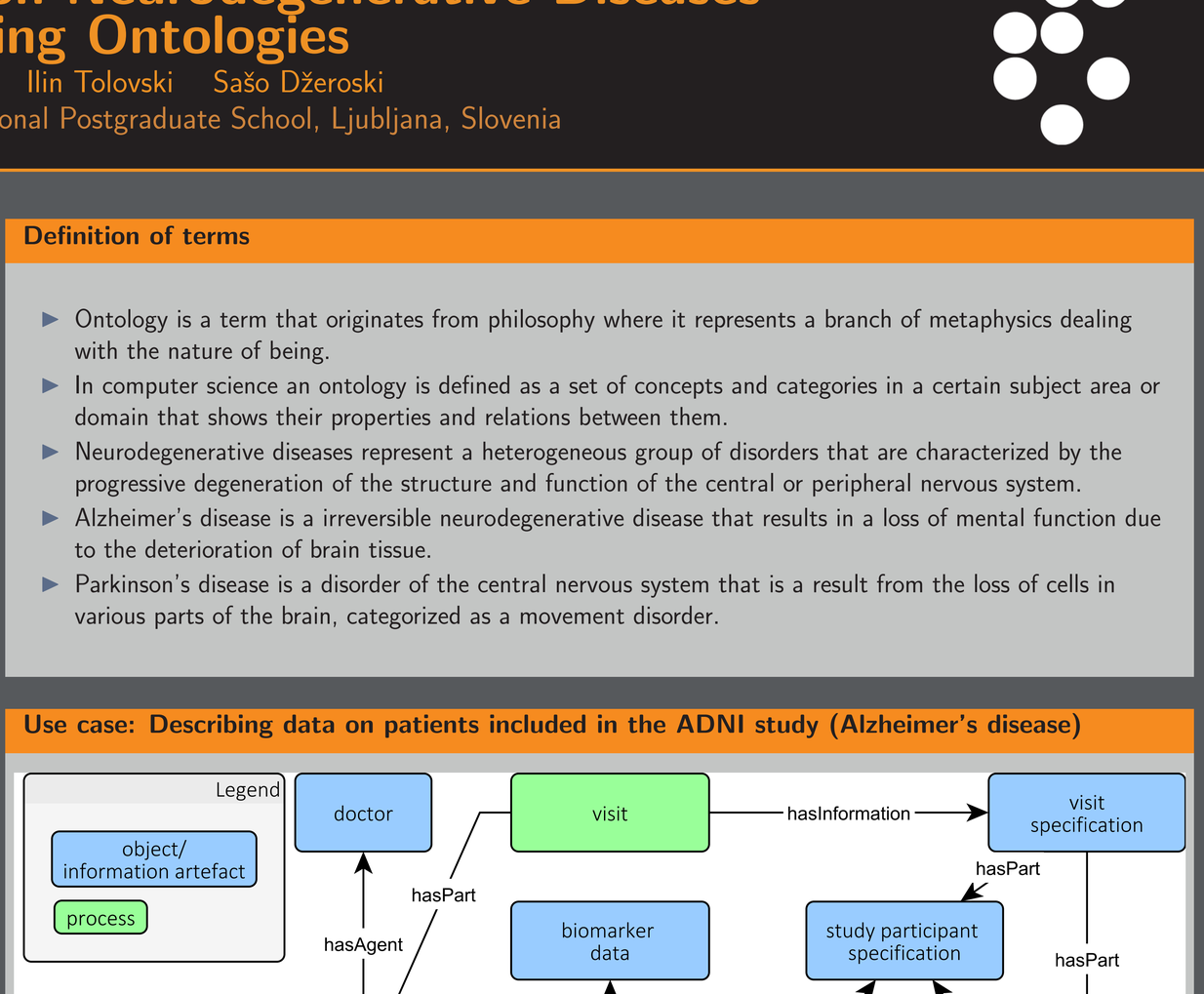

Figure 3: Scheme of the part of the ontology describing Alzheimer's disease

Example of annotation of dataset originating from the ADNI study

- We annotated an instance of ADNI dataset in a semi-automatic manner, using the developed ontology and Cellfie plug-in for Protg which enables importing axioms on data from Excel sheets. The rules are written in domain specific language (DSL) based on the Manchester OWL Syntax.
- On Figure 4 we can see an example of instantiating the Alzheimer's Disease Score (ADAS11) for each row (patient) of our dataset.
- Furthermore, the whole annotation pipeline is shown, starting from writing the rules in the Cellfie toolbox, towards the results i.e. generating an instance for each data example and its respective relations to all the parameters in the dataset, described by the classes in NDDO.

Semantic Annotation Pipeline

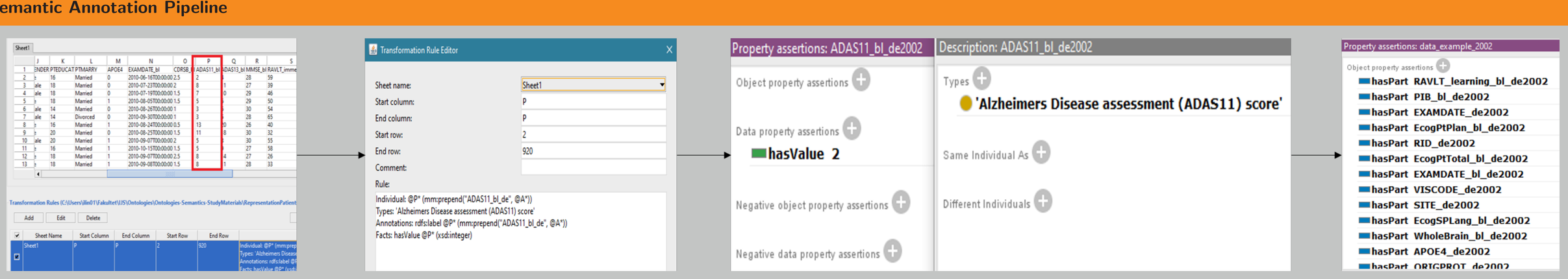

Figure 4: Pipeline of semantic annotation of the ADNI dataset using Cellfie

Conclusion & Future work

During this project, we created the NDDO mid-level ontology for describing data on patients with neurodegenerative diseases. Using this ontology, we semantically annotated two datasets from studies on Alzheimer's (ADNI) and Parkinson's disease (PPMI). In the future, we will publish the ontology on BioPortal, in order to make it publicly available and reusable. Furthermore, the annotated datasets will be stored in a triple store database, making them available for trying different querying scenarios. After having finished this, we plan to extend the ontology towards describing data from patients with other brain diseases.

Acknowledgments

The data used in this work was obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) and Parkinsons Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). We also acknowledge the European Commissions support through the H202 Human Brain Project SGA1 (Grant No. 720270) and the Slovene Human Resources Development and Scholarship Fund which conducted the Ad Futura Scholarship for Education.

References

[1] A. D. Spear B. Smith, R. Arp. *Building Ontologies with Basic Formal Ontology.* MIT Press, 7th edition, 2015.

[2] D. Calvanese G. De Giacomo M. Lenzerini A. Poggi, D. Lembo and R. Rosati. Linking data to ontologies. *Journal on Data Semantics*, pages 133–173, 2008.

[3] A. Mitelpunkt N. Shachar D. Gamberger, B. Ženko and N. Lavrač. Clusters of male and female alzheimers disease patients in the alzheimers disease neuroimaging initiative (ADNI) database. *Brain Informatics*, pages 169–179, 2016.

[4] L. N. Soldatova P. Panov and S. Džeroski. Generic ontology of datatypes. *Information Sciences*, pages 900–920, 2016.

[5] L. N. Soldatova P. Panov and S. Džeroski. Ontology of core data mining entities. *Data Mining and Knowledge Discovery 28*, pages 1222–1265, 2014.

[6] B. Draganski V. Mileski, D. Kocev and S. Džeroski. Multi-dimensional analysis of PPMI data. *Proceedings of the 8th Joef Stefan International Postgraduate School Students Conference*, pages 175–178, 2016.

[7] Alzheimers Disease Neuroimaging Initiative (ADNI) database. URL: http://adni.loni.usc.edu/). 20.10.2017.

[8] Parkinson's Progression Markers Initiative (PPMI) database. URL: http://www.ppmi info.org/. 20.10.2017.

[9] ADNI Procedures Manual. URL: https://goo.gl/UbR2Ad. 20.10.2017.

[10] ADNI GO Procedures Manual. URL: https://goo.gl/od8yUC. 20.10.2017.

[11] ADNI 2 Procedures Manual. URL: https://goo.gl/VQ4pL7. 20.10.2017.

[12] PPMI Study Protocol. URL: https://goo.gl/QrD7p8. 20.10.2017.

[13] PPMI Case Report Forms. URL: https://goo.gl/JCpZRw. 20.10.2017.