Domain Ontology of a PhD Program

Dana Malcova[1]

[1]Prague University of Economics and Business, Nam. W. Churchilla 1938/4, 130 67 Prague, Czech Republic

Abstract

This poster paper aims to aid in bridging the gap between funding organizations, incl. industry partners, and PhD Program providers, allowing for a more transparent and value-creating collaboration between both parties and the PhD Candidates. It introduces the initial version of domain ontology of a PhD program as an entry point for development of a more robust future version of a broader domain ontology of PhD studies, which will serve as an input for unpacking of both PhD Program's and PhD Candidate's funding needs. The domain ontology is modelled using OntoUML modelling language, which anchors this work in the UFO foundational ontology. This enables to capture the human understanding of this domain, which will help to articulate domain specific nuances to all parties involved, for purposes of analysis and clear communication of each party's needs, expectations and agreed outcomes.

Keywords

PhD program, PhD studies, PhD program ontology, domain ontology, OntoUML, UFO

1. Introduction and background

PhD studies and PhD research funding is a constant topic of discussions on institutional, national and international level. Unfortunately not everyone who is involved in such discussions, including industry partners providing PhD funding, has a holistic understanding of the problem area and its many nuances. Usually only researchers, who have undergone PhD training themselves, have a more complete understanding of the domain. While even their understanding is rather partial – thanks to the diversity in PhD programs' designs and funding models as well as unique attributes and circumstances of every individual PhD candidate.

This leads to confusion, myths and results in design of suboptimal policies and PhD program designs, which add complexity, uncertainty and require additional resources and effort from many of those involved in this very specific domain. Most of these effects then disproportionately affect the PhD candidates themselves and the departments, faculties and institutions, where they conduct their early research career training. These are the levels of the system with the most understanding and appreciation of the topic – however also the levels with the least power to influence the overarching policies anchored mostly in national legislation and designs of funding programs of third parties.

The ambition of this work includes providing a shared view of a nuanced domain understanding to all PhD program stakeholders, incl. PhD program providers, PhD fuding providers and the PhD candidates, who need to navigate the research domain as they only step by step learn about its terrain and realities.

Understanding the domain's structure is fundamental for clear understanding of the types of PhD funding needs and when these occur during a PhD candidate's lifecycle. This will aid management of all stakeholders' expectations and enable better planning and prioritization of allocation of existing financial resources, as well as strategic planning and execution of efforts required to secure and allocate PhD funding from external and temporary sources – which are the key types of scientific research funding.

ER2025: Companion Proceedings of the 44th International Conference on Conceptual Modeling: Industrial Track, ER Forum, 8th SCME, Doctoral Consortium, Tutorials, Project Exhibitions, Posters and Demos, October 20-23, 2025, Poitiers, France

dana.malcova@vse.cz (D. Malcova)

0009-0006-8707-3668 (D. Malcova)

2. Ontology description

The presented ontology covers the very core elements of a PhD program setup. This ontological model is being created for purposes of initial analysis of the domain at hand – it is basically a model to aid the "requirements analysis" phase of the broader project with the overall goal of designing a refined ontological model of the PhD Funding domain using the OntoUML modelling language.

While various ontologies exist that cover various aspects of the higher education or human education domains, no available ontology covers the core of the PhD domain in its entirety.

The key use case behind the development of this ontology was creation of a 1-page sufficiently nuanced shared view of the PhD program domain concepts for all domain stakehoders. This initial view will serve as a starting point for discussions with representatives of each stakeholder group which is involved in PhD funding or is affected by it. So far only the core elements have been defined – in the next steps this version will be reviewed by stakeholder groups involved in various PhD programs – at first at Prague University of Economics and business with further refinement based on an international group of respondents.

The outcome of this next phase of the research project, will then serve for identification and definition of funding needs on the PhD program as well as individual PhD candidate's level. This information will then provide a more generic and more holistic picture of PhD funding needs, which will be used for further research purposes. But will also be available for the community to reuse for various research purposes in the domain of PhD studies. There will be also an immediate value add to the international PhD community and to everybody considering commencement of their personal PhD journey. Having this information readily available in an easy to understand format will help PhD candidates to optimize their PhD funding and spending – and will help to manage expectations and funds to both groups. Also university leadership teams and their staff involved in securing and realization of PhD programs will have a shared view of PhD funding needs which will aid communication with their internal and external stakeholders as well as optimized allocation of available PhD funds.

Following this initial stage, the ontology development will continue using the FORZA methodology [1] following its four main phases:

1. Preparation and analysis
2. Ontology reuse (Foundational and domain ontologies)
3. Modeling in ontoUML using Visual Paradigm
4. Validation and refinement

This ontological model will be also expanded with the concepts of Funding need and Funding resources – and related Research funding concepts.

Ontologies to be reused include the following:

1. **Funding, Research Administration and Projects Ontology (FRAPO)**
   FRAPO is a CERIF-compliant ontology for describing administrative information related to grant applications, projects, and research funding. Originally published in 2010.
   Link: SPAR Ontologies - FRAPO, http://www.sparontologies.net/ontologies/frapo
2. **Data INtegration for Grants Ontology (DINGO)**
   DINGO provides a machine-readable, extensible framework to model data related to projects, funding, actors, and funding policies within the research landscape.
   DINGO 2020 FULLTEXT PDF: https://arxiv.org/pdf/2006.13438
3. **The Open Funder Registry (OFR, formerly FundRef) (FundRef – Funder Registry)**
   An open and unique registry of persistent identifiers for grant-giving organizations around the world. Link: https://fairsharing.org/FAIRsharing.odf1nG

A thorough review of other relevant resources including [2] is currently in progress.

The ontological model and its future versions will be reviewed using the MENTOR app and approach. And the content will be reviewed by domain experts as well as experienced ontology and ontoUML experts. Current version already builds on consultation of previous draft version with prof. Nicola

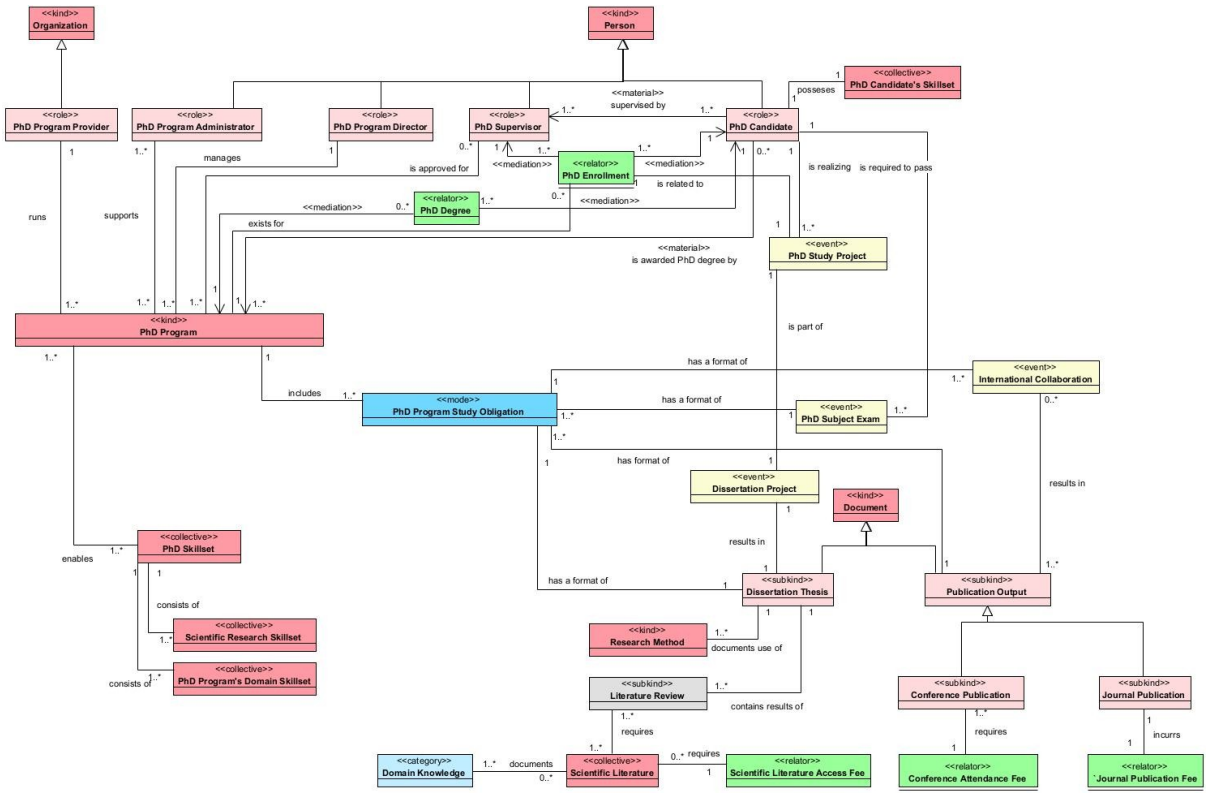

Figure 1: OntoUML model of domain ontology of a PhD program – work in progress

Guarino, who pointed out overuse of Relators as one of the key issues with the initial draft and provided also other useful feedback. Another improvement has been introduction of Event type concepts based on input from recent C-FORS Summer School in Foundational Ontologies, where prof. Giancarlo Guizzardi dove into more details of this concept.

The ontological model will be placed in the OntoUML/UFO Catalog to enable its reuse by the community and FAIRification of the resources.

3. Definitions of the concepts in scope

We consider a PhD Program Provider as a type **Organization**.

PhD Program Provider is any organization which is offering a PhD Program. Further evaluation will be done to determine the definition of this object class more specifically.

PhD Program represents the key entity in this ontological model. A PhD Program represents a program of study which grants the graduates the Ph. D. degree upon successful completion. There are a number of key PhD Program stakeholders. So far the following stakeholders of the **Person** kind are captured in the model:

PhD Program Director – person responsible for running the entire PhD Program;

PhD Program Administrator – person responsible for administrative support of the PhD Program;

PhD Supervisor – supervisor of the PhD candidate, appointed by the PhD Program;

PhD Candidate – a person enrolled in the PhD program, who receives the PhD degree upon completion of all study requirements and graduation.

PhD Enrollment represents the participation of an accepted PhD Candidate in a PhD Program.

PhD Study Project represents a "container" style concept enabling a holistic view of a PhD candidate's activities from a project management perspective. This entity is important as a container of all tasks that are required to be completed by a PhD candidate during the course of their PhD studies, and

as such will serve as a basis for identification of all relevant types of resulting PhD candidate's funding needs.

PhD Candidate's Skillset is a container type object representing a PhD candidate's set of skills

Study Obligations are the conditions a PhD candidate has to fulfill during their participation in the PhD program in order to successfully graduate from the program.

International Collaboration is one of the types of study obligations of the PhD program.

Dissertation Project represents a project management view on a research project, results of which are presented in the format of a dissertation thesis.

Dissertation Thesis is a document summarizing a PhD candidate's work on their dissertation research project and its outcomes in a standardized way. It is also the key deliverable of one's PhD studies required to fulfill a key study obligation.

Research Method is a method established as eligible for use in scientific research.

Literature Review represents a part of Dissertation Thesis, which is a result of the activity of reviewing existing scientific literature for purposes of a specific research project and/or output of such activity.

Scientific Literature represents a body of knowledge documented in a standardized format and published as outputs of scientific research activities.

Scientific Literature Access Fee represents cost related to ensuring access to scientific literature.

Domain Knowledge is knowledge of realities pertaining to a specific domain of human interest.

PhD Subject Exam represents one format of study obligation, which is satisfied by passing an exam associated with a PhD subject of study.

Publication Output represents published outputs of a research project, here specifically as a PhD project's output required to satisfy specific study obligations.

Conference Publication is an output of dissertation research project published in format of an article published in conference proceedings.

Conference Attendance Fee is the fee incurred by a PhD candidate's attendance of a scientific conference.

Journal Publication is an output of dissertation research project published in format of a journal article.

Journal Publication Fee represents cost associated with publication of a paper in a scientific journal.

PhD Degree is an academic degree for a doctor of philosophy.

PhD Skillset represents the skillset which a PhD candidate shall develop during their participation in the PhD program and which shall be possessed by the PhD program's graduate.

Scientific Research Skillset consists of skills relevant for the scientific research domain.

PhD Program's Domain Skillset consists of skills relevant for a specific domain of a PhD program.

In the presented version of the domain model, we are abstracting from time-related dimensions. The model will be expanded by introduction of relevant time dimensions in its future versions. The author is aware of the importance of these dimensions, as one of the practical challenges in this domain is synchronization of the cycle of application for funding with the PhD candidate's PhD lifecycle as well as the PhD Program's lifecycle. This is another existing challenge, which shall be mitigated by application of the artefacts resulting from author's dissertation project.

Declaration on Generative AI

The author has not employed any Generative AI tools.

This work has been supported by the EU's Horizon Europe grant no. 101058682 (Onto-DESIDE).

References

[1] C. Keet, M. T. Khan, C. Ghidini, Ontology authoring with forza, in: 22nd ACM International Conference on Information and Knowledge Management, CIKM'13, ACM proceedings, 2013.

[2] V. B. Nguyen, V. Svatek, G. Rabby, O. Corcho, Ontologies supporting research-related information foraging using knowledge graphs: Literature survey and holistic model mapping, in: D. M. Keet, C.M. (Ed.), Knowledge Engineering and Knowledge Management, volume 12387 of *EKAW 2020. Lecture Notes in Computer Science*, Springer, Cham, 2020.