Procedia Computer Science
Volume 80, 2016, Pages 1780–1790
ICCS 2016. The International Conference on Computational Science

Using Ontology Engineering Methods to Improve Computer Science and Data Science Skills

Svetlana Chuprina[1], Vassil Alexandrov[2], and Nia Alexandrov[3]
[1] *Perm State University, Bukireva Str. 15, 614990, Perm, Russia*
chuprinas@inbox.ru
[2] *ICREA-Barcelona Supercomputing Centre, Barcelona, Spain*
vassil.alexandrov@bsc.es
[3] *Barcelona Supercomputing Centre, Barcelona, Spain*
nia.alexandrov@bsc.es

Abstract
This paper focuses on issues of ontology construction process, Computing Classification System and Data Science domain ontology all used to help not only IT-students but any IT-specialists from industry and academia also to tackle the problems addressing the Big Data and Data Science skills gap. We discuss some methodological aspects of ontology design process and enriching of existing free accessible ontologies and show how suggested methods and software tools help IT-specialists including master students to implement their research work and participate in real world projects. The role of visual data exploration tools for certain issues under discussion and some use cases are discussed.

Keywords: Ontology Engineering Methods, Data Science Skills, Computer Science Skills, Data Science Ontology, Computing Classification System Ontology, Bridging Talent Gap, Visual Data Exploration Tools

1 Introduction

Nowadays, it is evident that both industry and academia have met a growing gap in Computational Science research skills and Data Science (DS) skills in particular. The situation is escalated by a rapidly growing technological diversity, and because the advance of more and more new methods that require the massive data processing in particular. The success of Big Data initiatives will depend upon the ability of companies and organizations to find the right talent, as well as the talent with the right abilities and to effectively integrate them into the business infrastructure (Chuprina, 2015).

There is no unique definition what is a Data Scientist although everybody knows that the Data Scientist skill set is a rare and highly multidisciplinary. Moreover different industries use Data Scientists for different purposes. So the role of a Data Scientist is very complicated and difficult to understand. We agree with a common and well-known point of view that main aim of a Data Scientist

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2016

doi:10.1016/j.procs.2016.05.447

is “to find meaning in the chaos of Big Data”. But this is a very common point of view and it is a very complex problem to choose a comprehensive subset from large set of technologies to teach the students to find Big Data meaning mentioned above. It is necessary also to take into account that any Faculty or Department that deals with the training of highly qualified IT-professionals and are going to open a new specialization in DS area, when choosing a profile of training and/or a focus of study, must take into account not only the needs of the employers, but also an existing infrastructure of the University, a presence of relevant professionals and teachers who have sufficient experience in DS training, as well as many other factors.

Up to now, there is no generally accepted methodology to help universities to tackle these problems especially for such diverse domain of educational training as a Data Science. In addition, from our point of view, it is not enough to launch new programs in Big Data and Data Science in order to meet the growing demand in DS specialization. Because DS projects usually are interdisciplinary it is necessary also to make changes in existing and more traditional IT educational programs (for example, within the remit of existing courses of Master Programs, in our case “Applied Mathematics and Computer Science”) to introduce some introductory DS courses and new more advanced courses aimed at developing the special collaborative skills to work together in a team with professionals from DS. The authors have long experience in training master and postgraduate students in this field (for more detail, see (Alexandrov & Alexandrov, 2015; Chuprina, 2015) and to engage them in real-word industry projects.

The authors have taken part in Severo Ochoa project (https://www.bsc.es/about-bsc/press/bsc-in-the-media/bsc-severo-ochoa-scholarships-high-performance-computing/), which is a very interdisciplinary project. Some problems under discussion were around “What data scientist skills does it take for a Big Data professional to work in collaboration within data science projects?”

Taking into account a big difference in backgrounds of master/postgraduate students, we stand in solidarity that it is very important to use some special methodology to automate an explicit visual representation building of these backgrounds in terms of concepts and relationships related to DS knowledge area context. We use for these purposes ontology engineering methods and suggest exploring the ontologies of personal student skills by its mapping to ACM Computing Classification System (http://www.acm.org/about/class/2012/) and Data Science Ontology (http://DataScienceOntology.com) to obtain clearer understanding of the needs of personal educational paths or pathways in each particular case. In some cases we have not only extended the ontologies mentioned above but also have made changes, for example, in the Data Science Ontology performing pre-conversion of the ontology into OWL standard format representation.

2 Data Scientist vs Data Engineer and Other Computer Science Skills

We have no aim to give a definitive answer to the question “What are the major differences between Data Scientist and Data engineer”? Moreover, it seems to us that there is no definitive answer to this question. But some demystifying of the Data Scientist’s skills and responsibilities is necessary to explain our point of view how it can be facilitated the training of students to be Data Scientist or to work effectively in collaboration with professionals from DS.

At the beginning, we demonstrate in Table 1 a brief matching of some definitions, responsibilities and skills of Data Scientist vs Data Engineer according to “Job Comparison – Data Scientist vs Data Engineer vs Statistician” (Saswat, 2015).

Table 1: Data Scientist vs Data Engineer (adapted from (Saraswat, 2015))

	Data Scientist	Data Engineer
Who is Who	These people use their analytical and technical capabilities to extract meaning insights from data	These people ensure uninterrupted flow of data between servers and applications. They are responsible for data architecture
Responsibilities	• Develop and plan required analytic project in response to business needs • Contribute to data mining (DM) architectures, modeling standards, reporting, and data analysis methodologies • Collaborate with stakeholders to integrate data mining results with existing systems • Monitor DM system performance and implement efficiency improvements	• Design, construct, install, test and maintain highly scalable data management systems • Improve data foundational procedures, guidelines and standards • Integrate new data management technologies and software engineering tools into existing structures • Create custom software components and analytics applications
Skills	• Programming • Mathematics • Business Understanding • Statistics • Data Visualization • Machine Learning • Attention to detail	• Database design • Production coding • Data Collection • Data Warehousing • Data Transformation • Work diligently with data

This table does not reflect some latest trends in the massive data analysis, which will become increasingly valuable in the long term. Not only applying Data Mining and Machine Learning to Big Data analysis are very useful, but also other practically relevant methods from the field of Artificial Intelligence (AI) are important tackling Big Data analytical problems.

According to (Council, 2013) a number of challenges in both data management and data analysis require new approaches to support the Big Data era. These challenges span generation of the data, preparation for analysis, and policy-related challenges in its sharing and use, contextual semantic search, which use for their tackling among others the modern AI tools and techniques, and as we point in ontologies as the models for information representation and analysis in particular. A challenge is to merge the benefits of different technologies that are being integrated and used to drive toward more systematic and valuable approaches in massive data analytics.

For example, most technologies applied to social networking, which is a growing source of massive data, do a little more than collect data with simply counting the frequency of messages, key words, hashtags, etc. and provide simple visualizations. Truly making use of this data requires scalable clustering techniques, real-time ontology abstraction, and on-the fly thesauri creation for extracting the complete network associated with a topic of interest (Council, 2013).

So the demand for data scientists with the wider range of multidisciplinary skills to understand and make decisions based on the analysis of Big Data is increasing. According to the well-known McKinsey Global Institute research study, by end of 2018 the US alone may face a 50% to 60% gap between supply and requisite demand of deep analytic talent. And this forecast is representative of the general state-of-the-arts. We know (see, for example, the review "Data Science Revealed: A Data-Driven Glimpse into the Burgeoning New Field" of EMC, http://www.emc.com/collateral/about/news/emc-data-science-study-wp.pdf) that "trends tend to follow a cycle, where the initial opportunity leads to ever increasing demand for a certain set of skills, while later demand wanes as many of those initial skills are automated by even newer tools".

In order to remain competitive in the world of Data Science, the universities need not only launch new DS education specialties or increase the number of DS graduates but also to make changes in the curricula of some other widespread educational specializations like Computer Science, Business

Intelligence, Business Analytics and others to train specialists, which must be ready to develop new software and upgrade the existing tools that are conducive to the data-driven decision making. The graduates from mentioned above specializations must also have some high-performance skills to work within cross-functional teams that include a variety of roles, such as programmers, statisticians, data scientists, data engineers, graphic designers, and others. Besides that, the data science toolkit is more varied and more technically sophisticated than the BI toolkit. DS professionals use not only SQL, but also advanced statistical packages, NoSQL databases, big-data tools like Hadoop, advanced visualization tools like Tableau, and others.

Taking into account a different students' backgrounds in a computing area one of the possible ways to help universities to close the gap in DS multidisciplinary skills is to use some special educational methodic not only to improve a learning process especially focused on the big data analytics students' skills development but also to give them a more strict identification state of the art of their current skills and competences in an explicit form and means/instruments to compare them with the necessary subset of DS skills. Because ontologies have a great potential to accumulate the knowledge in a systematic form and to depict it in clear graphic form ontologies in our opinion are very suitable for these purposes.

We have had good experiences in the use of ontology engineering methods for different purposes within the educational process and students research work according to varied learning objectives (Alexandrov & Alexandrov, 2015; Alexandrov, N. Cherenkova & Chuprina, 2010; Chuprina & Statsenko, 2010; Chuprina, 2015; Ryabinin & Chuprina, 2015b). According to the topic of this paper we also suggest using a domain or an application ontology not only as a subject of study but also as a learning tool helping students to construct their own ontological profile within DS area and related educational trajectory.

3 Methodological aspects of Ontology Design Training

In this paper we describe how we teach students to build their research domain ontology, how it can help them to improve their Computer Science skills and enable them to improve their research work and take part in real-world projects. According (Gómez-Pérez, Fernández-López, & Corcho, 2007), ontology engineering is "the set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies". It must be well-organized process supervised by the teacher or the specialist from industry with a good experience in ontology engineering. At the beginning of this section, we give some methodological recommendations about how to help students having different backgrounds and different specialties to learn ontology engineering methods and construct their own first ontology. Then we discuss some use cases and describe our suggestions how to transfer the existing experience in the improvement of students' Computer Science skills by means of research domain ontology construction of the DS field. We believe that our suggestions can help students to meet the requirements for the new DS specializations.

One of our aims is to teach the students from the different IT-specializations (and not only IT) to construct their own research area domain ontology. We suggest to involve a learning of ontology engineering methods at the Bachelor level not only for Computer Science professionals training but also for every IT-specialist training and even, ideally for any university Bachelor program (at 4-th year of education) to help students to improve her/his research skills and to get the methodological basis for further collaborative skills development to work in an interdisciplinary projects.

We recommend beginning with the brief Introduction to Ontologies and Semantic Web (see, for example, http://www.obitko.com/tutorials/ontologies-semantic-web/introduction.html) and discuss the Thomas Gruber ontology definition as "explicit specification of conceptualization". Then instead of more formal declaration we give a simple clear understandable example showing the necessity of semantic contextual search and the role of ontology to enrich the query for more relevant ranking of retrieval results similar to the following query (without inverted commas or any other quotes): <Red List of Perm>.

As the left part of Figure 1 shows, the first (!) and seven references from the top positions of the first page of ranking (and a number of next ones not shown in this picture) of this query retrieval results are not pertinent (data of access via Google search engine is 30.01.2016). To do our demo ontology more applicable and reusable we integrate it with the data base of regions (<DB of regions>) and use it's interpretation for automated transformation of source query text about Red List of any region to the follows "Red List" + <Name of Region> during the pre-processing. In our case the source query is automatic transformed to the query "Red List" + Perm. The right part of Figure 1 demonstrates example of demo ontology "Red List of Regions" and the related retrieval results which are more pertinent to the source query due to the automated pre-processing.

After that, because ontology enables "to share common understanding of the structure of information among people or software agents" (Noy & McGuinness, 2001) we discuss the issues related to the features of ontology representation of knowledge such as following:

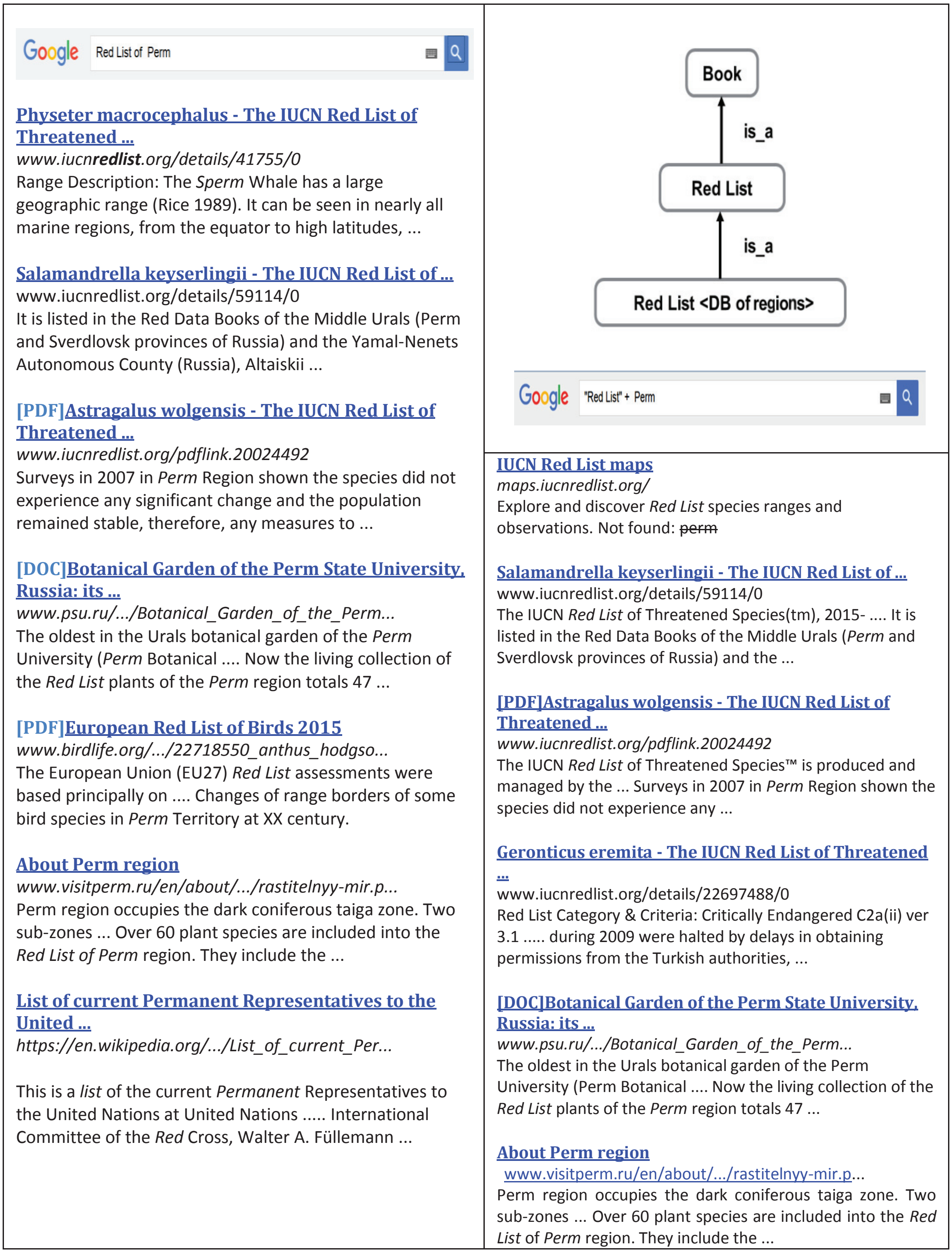
- enabling reuse of domain knowledge;
- making domain assumptions explicit;
- separating domain knowledge from the operational knowledge;
- analyzing domain knowledge and others.

Any student from any specialization including Data Scientist and Data Engineer must after the first lesson to prepare the list of definitions, main concepts related to its research domain and to send this list to the educator. The teacher can select one or two domain area descriptions that are more understandable for most students and at the next lesson can demonstrate all the life cycle steps on ontology development beginning with requirement and domain analysis. It is possible to use for this purpose the results of ontology development of the previous year students.

First of all, it is necessary to determine the domain of interest and scope. In our case we suggest bachelors to construct such domain ontology that can be used as a basis of their graduation work overview chapter. Additionally, we recommend using the result of identification step as a glossary of graduation research work (bachelor/master thesis).

The main steps of the life cycle of ontology development are as follows:

1. Determine the domain and scope of the ontology.
2. Consider reusing existing ontologies.
3. Enumerate important terms in the ontology (identification step).
4. Define the classes and the class hierarchy.
5. Define attributes and relationships.
6. Define the restrictions of the properties.
7. Create instances.



Physeter macrocephalus - The IUCN Red List of Threatened ...
*www.iucn**redlist**.org/details/41755/0*
Range Description: The *Sperm* Whale has a large geographic range (Rice 1989). It can be seen in nearly all marine regions, from the equator to high latitudes, ...

Salamandrella keyserlingii - The IUCN Red List of ...
www.iucnredlist.org/details/59114/0
It is listed in the Red Data Books of the Middle Urals (Perm and Sverdlovsk provinces of Russia) and the Yamal-Nenets Autonomous County (Russia), Altaiskii ...

[PDF]Astragalus wolgensis - The IUCN Red List of Threatened ...
www.iucnredlist.org/pdflink.20024492
Surveys in 2007 in *Perm* Region shown the species did not experience any significant change and the population remained stable, therefore, any measures to ...

[DOC]Botanical Garden of the Perm State University, Russia: its ...
www.psu.ru/.../Botanical_Garden_of_the_Perm...
The oldest in the Urals botanical garden of the *Perm* University (*Perm* Botanical .... Now the living collection of the *Red List* plants of the *Perm* region totals 47 ...

[PDF]European Red List of Birds 2015
www.birdlife.org/.../22718550_anthus_hodgso...
The European Union (EU27) *Red List* assessments were based principally on .... Changes of range borders of some bird species in *Perm* Territory at XX century.

About Perm region
www.visitperm.ru/en/about/.../rastitelnyy-mir.p...
Perm region occupies the dark coniferous taiga zone. Two sub-zones ... Over 60 plant species are included into the *Red List of Perm* region. They include the ...

List of current Permanent Representatives to the United ...
https://en.wikipedia.org/.../List_of_current_Per...

This is a *list* of the current *Permanent* Representatives to the United Nations at United Nations ..... International Committee of the *Red* Cross, Walter A. Füllemann ...

IUCN Red List maps
maps.iucnredlist.org/
Explore and discover *Red List* species ranges and observations. Not found: ~~perm~~

Salamandrella keyserlingii - The IUCN Red List of ...
www.iucnredlist.org/details/59114/0
The IUCN *Red List* of Threatened Species(tm), 2015- .... It is listed in the Red Data Books of the Middle Urals (*Perm* and Sverdlovsk provinces of Russia) and the ...

[PDF]Astragalus wolgensis - The IUCN Red List of Threatened ...
www.iucnredlist.org/pdflink.20024492
The IUCN *Red List* of Threatened Species™ is produced and managed by the ... Surveys in 2007 in *Perm* Region shown the species did not experience any ...

Geronticus eremita - The IUCN Red List of Threatened ...
www.iucnredlist.org/details/22697488/0
Red List Category & Criteria: Critically Endangered C2a(ii) ver 3.1 ..... during 2009 were halted by delays in obtaining permissions from the Turkish authorities, ...

[DOC]Botanical Garden of the Perm State University, Russia: its ...
www.psu.ru/.../Botanical_Garden_of_the_Perm...
The oldest in the Urals botanical garden of the Perm University (Perm Botanical .... Now the living collection of the *Red List* plants of the *Perm* region totals 47 ...

About Perm region
www.visitperm.ru/en/about/.../rastitelnyy-mir.p...
Perm region occupies the dark coniferous taiga zone. Two sub-zones ... Over 60 plant species are included into the *Red List* of *Perm* region. They include the ...

Figure 1: The example of automated pre-processing of source query to achieve more pertinent results due to the demo ontology “Red List of Regions”

The step 2 “Consider reusing existing ontologies” is very important because if suitable existing ontology is found then its reusing can significantly reduce the time of ontology development. The existing ontologies may be a requirement if student’s system development within graduation research work needs to interact with other applications that have already committed to particular ontologies or controlled vocabularies; a lot of ontologies are already available in electronic form and can be imported into an ontology-development environment that the students are using. So it is necessary to discover potential reuse candidates, evaluate their usability, and customize ontology to be reused, integrate and merge student’s ontology under development to the target ontology.

During the ontology reusing consideration a special attention should be given to the issues of ontology reusing limits such as mismatches (do not exist), finding “Right Ontology” (difficult to find), inconsistency and incompleteness (doesn’t fit), modularity (needs only a part of ontology), integration (alignment, mapping, merging).

At the conceptualization phase (step 4 and step 5) the knowledge of the domain of interest is structured via the designing of conceptual graph with nodes representing the concepts (terms are described on the step 3) and links representing the relationships between related concepts. First of all, it is recommended to construct the taxonomy relations (“is-a”, “class-subclass”). After building taxonomy of concepts other types of relations are established. Based on our experience we recommend restricting the set of relationships’ types and suggesting students try to use in addition to the ‘is-a” relations only the follows types of relations: “a-part-of”, “has”, “synonym”, “use”, and to monitor the amount of every relations’ type used. In masters educational program not only domain ontologies but also application ontologies are constructed to improve the adaptability of developed systems and shells within masters research work. In addition of the above mentioned such types of relations as “element-of-set”, “input”, “output” (for specifications of input and output parameters) are used.

We recommend not to use such type of relations as “see also” during the first steps of creation and interactive improvement of student’s ontology because our practice demonstrates that students are very likely to use overwhelmingly this relation type and without bothering to find out whether another type of relationship is more appropriate.

Formalization phase adds taxonomic and ad-hoc to the common terminology yielding lightweight (for bachelors) but formal ontologies. The implementation of ontology we recommend to do within suitable visual ontology editor. As the most accessible ontology editors are intended for ontology engineers we use ONTOLIS, which is an adaptable visual editor intended for domain experts and casual users (Chuprina, 2015; Ryabinin & Chuprina, 2015a, 2015b) and can import ontology specification in *OWL* standard. Inter alia, the recent version of ONTOLIS allows saving a needed part of ontology as a new separate ontology.

Ones more activities involved in life cycle of ontology are developed: knowledge acquisition and evaluation, the educator monitors the first one and makes the second one by itself. The refining of first ontology is repeated until consensus between the teacher and the student is reached. The process of interactive improvement of ontology is an adaptation of the Delphi method, a technique for collecting views of several stakeholders.

4 Using ACM Computing Classification System and Data Science Ontology to Improve Computer Science Skills

At the “Consider reusing existing ontologies” phase (step 2) we propose to use the mentioned above ACM Computing Classification System (CCS) or Data Science Ontology as target ontology, depending on the specifics of the students’ research subject area. We show bellow how to refine and expand this ontology resources, and what to do if student do not find the right or similar concept within them, or it is available but is used repeatedly in different contexts.

The 2012 ACM Computing Classification System has been developed as a poly-hierarchical ontology that can be utilized in semantic web applications. It is integrated into the search capabilities and visual topic displays of the ACM Digital Library and various other libraries, companies, and publishers such as Springer, IEEE, and Emerald have made use of the ACM CCS. The full CCS classification tree has flat and interactive views and is freely available for educational and research purposes in these downloadable formats: *SKOS (xml)*, *Word*, and *HTML*. As we recommend for students using the *OWL* standard to represent ontology it is necessary firstly to find a suitable place in the context of parents' concepts hierarchy and nearest siblings and to depict this hierarchy within ONTOLIS environment.

If somebody doesn't find the needed place it is necessary to find similar concepts in terms of their meaning or in terms of similarity of their contexts. Let's see related use case. For example, the research area of one Master's Student is Competitive Intelligence (CI). There is no "Competitive Intelligence concept in CCS, but there are some concepts, which have the term "Intelligence" as a part of its designation. One of them is "Business intelligence", but this term is represented by several nodes with the same name in different contexts as it is shown on the Figure 2.

• Applied computing •Enterprise computing • Business process management • Business process modeling • Business process management systems • Business process monitoring • Cross-organizational business processes • Business intelligence	• Information systems • Information retrieval • Retrieval tasks and goals • Question answering • Document filtering • Recommender systems • Information extraction • Sentiment analysis • Expert search • Near-duplicate and plagiarism detection • Clustering and classification • Summarization • Business intelligence

Figure 2: A part of ACM CCS with the different context of the same concept – "Business Intelligence".

An exploration of parents' concepts hierarchy and nearest siblings including close study of their definitions helps the student to acquire a deeper understanding of the context of his/her research, an additional understanding of "Business Intelligence vs Competitive Intelligence" issues, and maybe there is a need for further consultation with his supervisor to choice the adequate ontology context.

Due to the limitations on the scope of the paper's volume we show only one example of a student ontology, which also demonstrates the role of visualization tools. Figure 3 shows application ontology created by MSc student, which is integrated with ACM CSS terms and is used in a real-world project, intended to visualize the dynamics of potash mine excavations. Figure 4 represents a screen short of the end-user interface of a Master's student mobile application that was automatically generated based on ontology represented on Figure 3.

How to use our experience in ontology engineering methods teaching to help the Computer Science students to become a Data Scientists and/or help the Computer Science students to tackle the DS problems by development more comprehensive assistant/supported tools? We suggest to use the ontology engineering methods for this purposes and by means of capabilities of visual ontology editor ONTOLIS to map the Master's students (or any other specialists) ontology profiles, which is a strong explicit specification of their current IT skills to the Data Science Ontology.

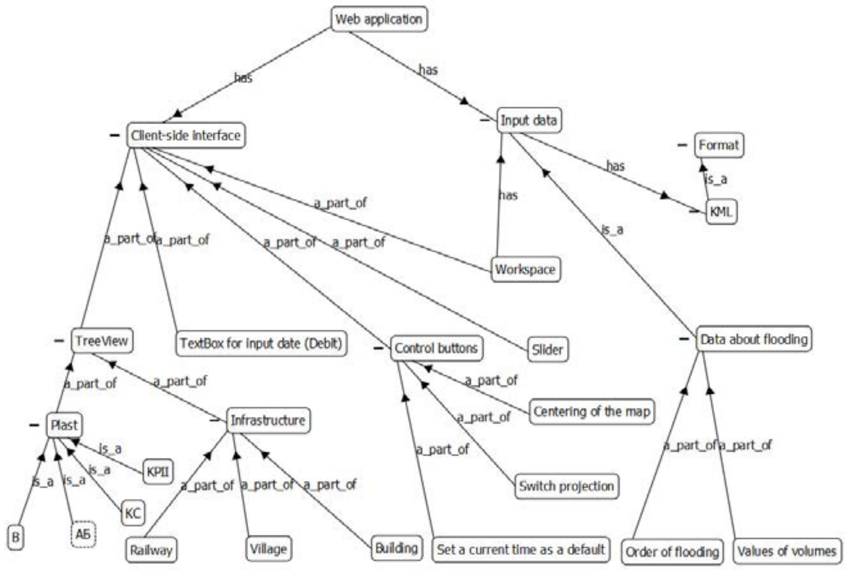


Figure 3: A part of student's application ontology integrated with ACM CSS terms within ONTOLIS

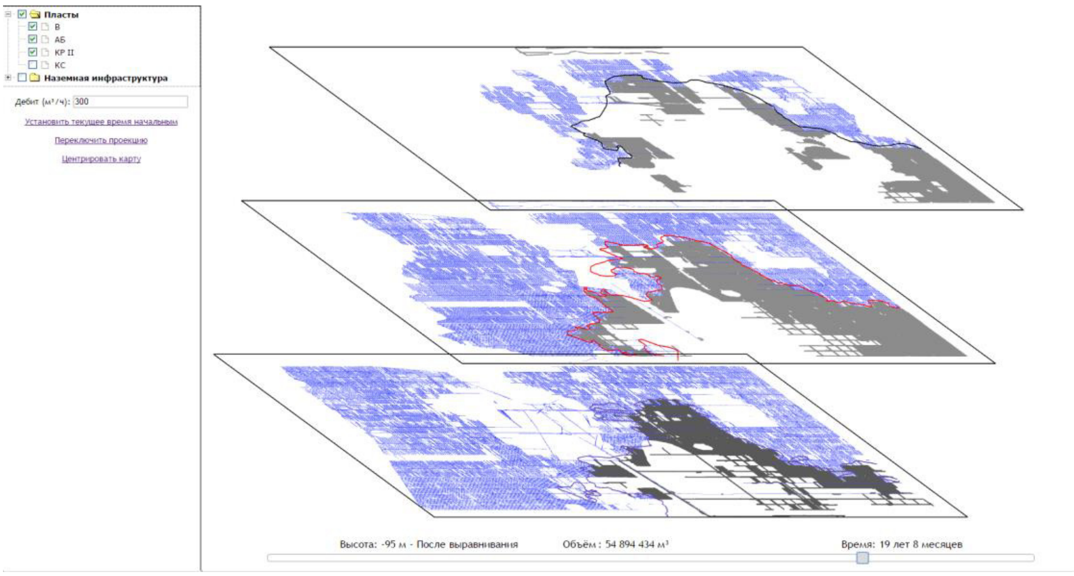


Figure 4: Mobile application interface automatically generated with the help of the ontology represented on Figure 3

Figure 5: A part of Data Science Ontology by Sean McClure (http://www.datascienceontology.com)

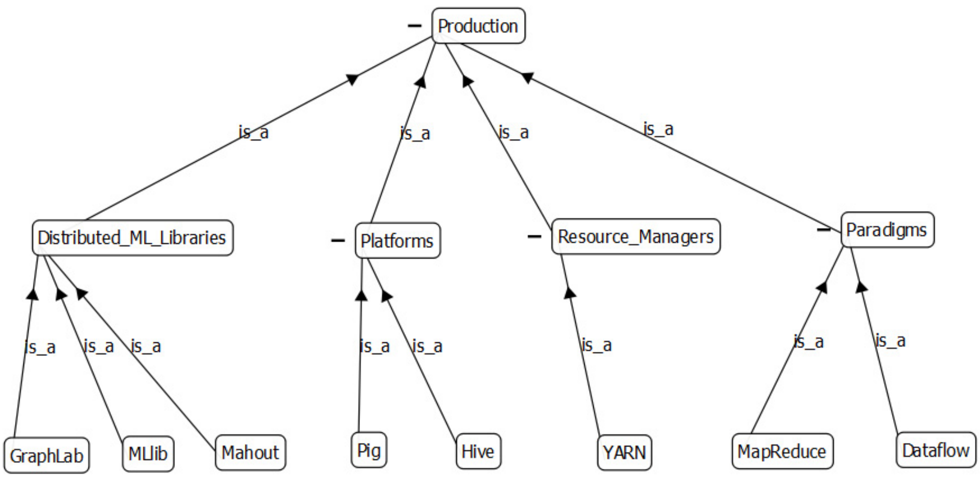


Figure 6: A part of Data Science Ontology showed in Figure 5 within ONTOLIS environment

Within Severo Ochoa project we have begun to create New Data Science Ontology (NDSO) to help teachers and students as well as researches both in industry and academia to find an explicit form of their personal trajectory to become Data Scientist and to understand their own gap. We use Data Science Ontology created by Sean McClure as target ontology (see Figure 5) and we automatically transform this ontology to *OWL* format (Figure 6 shows the part of this ontology within ONTOLIS). In future we have the plans to expand and to improve this ontology.

5 Conclusion

The aim of this paper has been to clarify the issues of students' ontology construction process and using Computer Classification System and Data Science domain ontology to help not only IT-students but also any IT-specialists from industry and academia also to tackle the problems addressing the Big Data and Data Science skills gap. Thanks to Sean McClure (Data Scientist, ThoughtWorks) who is the author of our target Data Science Ontology within Severo Ochoa project we have an opportunity to refine and to improve this ontology to give students and other specialists the means for explicit and strong identification of their Computer Science Skills in a form of personal ontology profile in comparison with the needed subset of DS skills.

Data science is a rapidly changing field, and so every student must be prepared to keep up with the state of the art. Due to a big amount of open ontology repositories for different domain areas (see, for example, http://taxonomies.labs.crossref.org/?page_id=11, http://mmisw.org/orr/, etc.) our approach is suitable to students who are not IT specialists to help them to take part in collaborative work within interdisciplinary projects.

References

Alexandrov, N., & Alexandrov, V. (2015). Computational Science Research Methods for Science Education at PG Level. *Procedia Computer Science*, *51*, 1685–1693. http://doi.org/10.1016/j.procs.2015.05.305

Alexandrov, N. Cherenkova, N., & Chuprina, S. (2010). Methods to Assess Students' Experiences in an Immersive 3D VR Environment. In *Proceedings of the 13th Interactive Computer Aided Learning International Conference ICL2010* (pp. 1248–1254).

Chuprina, S. (2015). Steps towards Bridging the HPC and Computational Science Talent Gap Based on Ontology Engineering Methods. *Procedia Computer Science*, *51*, 1705–1713. http://doi.org/10.1016/j.procs.2015.05.308

Chuprina, S., & Statsenko, N. (2010). Using ontology and metadata to integrate elearning resources and administrative information system of university. In *9th European Conference on eLearning 2010, ECEL 2010* (pp. 171–178).

Council, N. R. (2013). *Frontiers in Massive Data Analysis*. http://doi.org/18374

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2007). *Ontological Engineering: With Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web. (Advanced Information and Knowledge Processing)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. *Stanford Knowledge Systems Laboratory*.

Ryabinin, K., & Chuprina, S. (2015a). Development of ontology-based multiplatform adaptive scientific visualization system. *Journal of Computational Science*, *10*, 370–381. http://doi.org/10.1016/j.jocs.2015.03.003

Ryabinin, K., & Chuprina, S. (2015b). Using Scientific Visualization Tools to Bridge the Talent Gap. *Procedia Computer Science, 51*, 1734–1741. http://doi.org/10.1016/j.procs.2015.05.376

Saswat. (2015). Job Comparison - Data Scientist vs Data Engineer vs Statistician. Retrieved from http://www.analyticsvidhya.com/blog/2015/10/job-comparison-data-scientist-data-engineer-statistician/