Semantic Web 0 (2012) 1–0
IOS Press

The Bowlogna Ontology: Fostering Open Curricula and Agile Knowledge Bases for Europe's Higher Education Landscape

Editor(s): Carsten Keßler, University of Münster, Germany; Mathieu d'Aquin, The Open University, UK; Stefan Dietze, L3S Research Center, Germany
Solicited review(s): Mathieu d'Aquin, The Open University, UK; Carsten Keßler, University of Münster, Germany; anonymous reviewer

Gianluca Demartini [a,*], Iliya Enchev [a], Joël Gapany [b], and Philippe Cudré-Mauroux [a]
[a] *eXascale Infolab, University of Fribourg, Switzerland*
[b] *Faculty of Humanities, University of Fribourg, Switzerland*
E-mail: {firstname.lastname}@unifr.ch

Abstract. The Bologna Process initiated a radical change within higher education institutions. This change triggered the creation of new administrative procedures in the every day life of European universities. It also gave rise to the emergence of new concepts for the description of curricula. It is critical for the successful continuation of this process to support the publication and exchange of information among universities. With this aim in mind, we created the *Bowlogna* Ontology to model an academic setting as proposed by the Bologna reform. In this paper, we present our efforts to design this ontology and the entire process that lead to its creation starting from the definition of a linguistic lexicon derived from the Bologna reform and its conversion to a formal ontology. We also describe practical applications of our ontology for end-users at universities (such as a faceted search and browsing system for course information).

Keywords: Bologna Process, Ontology Engineering, Faceted Search

1. Introduction

In 1999, a process of renovation and standardization of Europe's higher education landscape started with the signing of the Bologna declaration. This process lead to the definition of standard ways to create university curricula, assign degrees, and evaluate students. This is of course a long process which, in order to be successful, requires time and several revisions of the proposed framework and its implementation by each participating country.

After more than ten years, the reform process is well under way and a number of higher education policies and processes have already been consolidated. For instance, the European Credit Transfer and Accumulation System (ECTS) has now been adopted by most European countries and allows students to attend courses in foreign universities and to automatically transfer the credits they earned abroad to their home university.

While the Bologna process surely has helped to consolidate the heterogeneous higher education systems across Europe, there is still a long way to go before creating a coherent, interconnected and integrated higher education system in Europe. For example, today an exchange student who wants her ECTS credits recognized at her home university has to provide a written certificate from the foreign university about her performance. Once earned, the credits are typically entered once at the foreign university, transferred by post to the home institution, and entered a second time manually

*Corresponding author. E-mail: gianluca.demartini@unifr.ch

at the home university. This scenario clearly points out the lack of cohesion and automation across university systems inside Europe.

To support the ongoing Bologna process, we describe in this paper our efforts towards automatizing both higher education processes and workflows by creating a standard and reusable ontology which formally defines the concepts and relations involved in an academic setting as described by Bologna. The motivation to create such an ontology is to push the standardization process that will enable the connection among academic systems and improve educational workflows, communication, and collaboration between universities. We can imagine different usage scenarios where a standard ontology describing the Bologna university environment would have a beneficial impact. One example is the sharing of information among different universities. The adoption of a standard schema would facilitate the mapping between heterogeneous university schemas and enable information flows among universities (e.g., Erasmus students who have taken exams in foreign universities). Such an adoption would also allow the creation of general tools for internally searching and browsing data. As the same schema gets adopted by different institutions, a given query could be run over different knowledge bases (see an example of such a process in Section 6). Processes in European universities currently follow the Bologna reform directives. For this reason, a common schema based on such directives would be the easiest to integrate in current workflows. In summary, there are several benefits in the definition and progressive adoption of a standard ontology targeting the Bologna university reform. In this paper, we first describe our effort in modeling the Bologna reform through the creation of a lexicon. We then describe how we converted the lexicon to a standard ontology. We also present different applications of the proposed ontology for end-users in university administrations (e.g., a faceted search system).

The rest of this article is structured as follows: Section 2 and Section 3 give a few details about the Bologna reform and our efforts to identify and define the key concepts related to the Bologna reform using a linguistic approach. This effort lead to the creation of a lexicon of about 60 key terms used for the implementation of the Bologna Process in our institution [7]. Then, we describe (Section 4) the process which, in collaboration with the linguists who defined the lexicon, lead to the definition of the Bowlogna ontology. In this part we also mention the challenges that we encountered in, for example, modeling *part-of* relations in this context. We present the Bowlogna ontology with its classes and properties in Section 5. More specifically, we discriminate between public and private parts of the ontology and also focus on the multilingual aspect of our work. We describe (Section 6) the envisioned usage of the proposed ontology and various applications fostering the integration of information from different knowledge repositories. Finally, we conclude the paper in Section 7 by summarizing the lessons learned during the process of creating a lexicon and an ontology about the Bologna Reform, and propose future directions with the use of Bowlogna to improve knowledge flows between higher education institutions in Europe.

2. The Bologna Process: Convergence within Diversity

In June 1999, the ministries of education of 29 European countries signed the so-called "Bologna Declaration". Ten years later, in March 2010, the European Higher Education Area (EHEA) was officially launched at the Budapest and Vienna Ministerial Anniversary Conference. This newly created area—which was on the same occasion enlarged to Kazakhstan—encompasses at present 47 countries, whose broad variety of cultures and languages unquestionably influence the way the reforms are locally understood and implemented. This diversity may be regarded as a cultural asset and there is of course a legitimacy to take into account the specificities related to the individual national contexts in which the Bologna Process takes place. On the other hand, the need for a common and mutual understanding of the ongoing reforms has progressively been more apparent with the enlargement of the EHEA.

The train of reforms in which higher education institutions have been involved during the last decade has been investigated by disciplines as different as pedagogy, management or sociology, to mention only a few. With many respects, this interest led to significant evolutions in teaching and curricula, in the organization and in the governance of universities, and in the perception that members of the academic community have of their role in the civic society. But, to some extent, the lack of communication between specialists from various disciplines has prevented higher education institutions from really developing a coherent view of the transformations they are currently experiencing.

Improving the mobility of students within the EHEA is often mentioned as one of the main goals of the Bologna process. But, as such, it evokes a more abstract type of mobility that has to do with the capacity to share information. Although this dimension is rarely mentioned as a specific goal, we think that, on the contrary, it is at the heart of the Bologna process. To some extent, it is even the necessary counterpart of the diversity of approaches and implementations we have just mentioned. Technically, it will require the development of integrated, next-generation information systems, such as the one we describe in the rest of this article. We are confident that this movement will not only be useful for new information systems, but will also contribute to a form of convergence that is at the same time constitutive of the Bologna process and necessary to the achievement of its main objectives.

3. The Bologna Linguistic Lexicon

The Bologna process takes place in a multilingual environment, which is certainly not a barrier to good communication, but implies the capacity to explicitly link local discourses to a common understanding—or at least to parallel and comparable understandings—of the various higher education systems. The issue can practically be addressed by constructing local ontologies that should relate or map to a set of common entities.

We recently published [7] in the official Bologna Handbook a basic set of entities that correspond to the new system of studies and that seem to be rather stable across time and institutions. In that article, we described how the compilation of a lexicon of about 60 definitions contributed to the implementation of the Bologna Process in our institution. We adopted a purely text-based approach to identify the various concepts defined in the lexicon: We hired a terminologist, whose work was to carefully review the texts underpinning the Bologna reform, and to extract all essential concepts (rather than single words) from the text while transcribing their implied meaning through succinct definitions and explicit relations to other concepts. The resulting lexicon has been spread among academic and administrative staff, in order to create a common understanding of Bologna.

4. Converting a Lexicon into an Ontology

In this section we describe the process that lead us to the creation of the Bowlogna ontology, a formal model of the academic setting as described in the Bologna reform. Once our domain of discourse was well understood and a standard lexicon was produced by domain experts together with terminologists, we were able to start the modeling process that later produced a digital formalization of the concepts described in the lexicon, along with their relationships.

As we did not designed the ontology by directly analyzing the existing domain bur rather transforming a model (i.e., the lexicon) into an ontological format, we did not use standard ontology engineering methods [8]. However, we relied on the high-quality modeling that was performed while creating the Bologna lexicon. For instance, properties and suggestions (e.g, the Unity criteria from [8]) had already been considered at the lexicon definition level and have been preserved in the Bowlogna ontology[1].

4.1. Previous Approaches to Thesaurus conversion

Several approaches have been proposed for (semi-)automatically translating lexicons or thesauri to Semantic Web formats with the goal of enabling interoperability between the converted models. Efforts in converting standard thesauri such as, for instance, the medical thesaurus MeSH used to index scientific articles, have been proposed [13]. In that case, an individual thesaurus was converted to RDF without the goal of having a general framework for such conversions.

Other researchers have proposed methods for converting a general thesaurus to an ontological format. With such approaches, the challenge is usually to model any type of relationship that can be found in a thesaurus into a model that focuses on taxonomic (e.g., *is-a*) relationships. An example of such an approach is [12] where, starting from a thesaurus from the Food and Agriculture Organization, an ontology meta-model is first defined, and a set of transformation rules is then applied to produce a formal conceptualization, which is edited manually in the end.

A last category of approaches consists of those using SKOS as a general framework for converting a thesaurus into an ontology. SKOS[2] provides a basic

[1] We also took particular attention in modeling part-of relations. See the discussion in Section 4.3.

[2] `http://www.w3.org/TR/skos-primer/`

framework to model the core structure of controlled vocabularies (e.g., thesauri and taxonomies). The use of such a common framework allows the produced models to be interconnected, that is, it allows to link concepts which are present in different thesauri. As examples of such approaches, [15] proposes generic methods to convert thesauri, producing a mapping to SKOS via an intermediate RDF meta-model. In [14], van Assem *et al.* propose a general method to convert thesauri to SKOS with the goal of interoperability between the converted models. Despite that goal, they do not provide evidence of any interaction between different models. The main limitation of this class of approaches is that information is typically lost in the conversion process, since not all information present in thesauri can be represented using SKOS (as, for example, for the MeSH lexicon [14]).

Compared to previous work, our setting does not require the produced model to be compatible with other ontologies but, rather, to well represent the modeled domain, and, thus, to encompass all important information present in the original lexicon. For this reason, we did not adopt SKOS classes and properties in the Bowlogna ontology. Another important aspect to be mentioned is the fact that we did not adopt an automated approach to create our ontology in order to have a high-quality outcome. Instead, we manually designed the ontology together with the help of the domain experts and terminologists who defined the lexicon containing the original concepts and relations.

4.2. Our Conversion Approach and Steps

Our approach to lexicon conversion has a lexical perspective. After domain experts have defined the main concepts and some of their relations in natural language, we performed a conversion of the lexicon to a formal ontology. The main goal of our approach was to reach *high coverage* in terms of the concepts mentioned and defined in the lexicon. As seen in previous work, coverage is not always the main goal of conversion approaches since, for the sake of compatibility between the generated ontology and legacy ontologies, previous approaches typically focus on a subset of all concepts defined in the thesaurus (see, for example, [14] where not all MeSH terms and relations are included in the final ontology).

In more detail, the steps we took in our approach are as follows. In the initial step, a lexicon is created by domain experts through a linguistic analysis of the sources defining the domain of discourse. The second step consists of creating, together with the help of domain experts, a mapping of lexicon terms to classes, instances, and relationships. Then, relationships are further refined to distinguish between different types (e.g., subclass, part-of). In a subsequent step, additional concepts are added to the ones imported from the lexicon in order to guarantee the full coverage of the domain. This step is also essential to integrate elements which are present in legacy systems and are not always directly defined in the lexicon or in the Bologna setting (such as local or legal information). Finally, a further revision is done to exclude concepts which are not useful in the deployment of the ontology for its final use (e.g., the class about the Bologna reform itself).

Technically speaking, in our setting, the conversion process started from a PDF document describing the linguistic lexicon. We then transformed the PDF document into an XML file, which was parsed by a conversion program that, exploiting the Jena framework[3], produced a minimal RDF file containing all the concepts defined in the lexicon. Those first RDF definitions were then revised manually with the help of domain experts and the use of Jena and the Protégé editor[4]. Our approach is similar to previous work in the sense that it adopts a semi-automatic approach to translate the lexicon into an ontology. Of course, a substantial effort was made to understand the lexicon before designing the first RDF schema and the translation rules which were then applied in our conversion program.

Another key property of our conversion method is the fact that we maintained all the definitions from the lexicon in four different languages (i.e., French, German, Italian, and English) in the final ontology. This is a critical multi-lingual first step, as we aim to establish an ontology for many countries and their respective universities. As a possible application scenario, we can imagine two different universities from Germany and France sharing data about the same concepts (e.g., courses and ECTS credits). Thanks to the multilinguality of our ontology, different actors can understand the meaning of concepts used to label the shared data.

Although we were taking a comprehensive approach, not all lexicon terms became concepts in the ontology at the time of our conversion process. This was the case for terms that are relevant to the Bologna

[3] http://jena.sourceforge.net/
[4] http://protege.stanford.edu/

setting which are, however, neither classes to be instantiated nor entities that can be part of a relation. An example of such a lexicon term is “Bologna Reform”. Further lexicon elements that were not included in the final ontology are terms that are either local or that are not generic enough to be of interest for each European university (e.g., “general study qualification for psychology students”). On the other hand, we added to the ontology classes which were not explicit terms defined in the lexicon but rather concepts mentioned in the definition of other terms and used by current information systems in our university. Examples of such classes include “Department” and various classes related to students and professors. The examples we have reported so far motivate our choice of a manual step in the conversion process even further, in addition to initial automated conversion step to syntactically convert the entire original lexicon to an ontology.

4.3. Mereology Relations

An interesting challenge we have faced during the conversion process is the handling of mereology relations (i.e., “part of”). Mereology defines the part-of relation using three axioms. It must be:

- **Transitive**: If A is part of B and B is part of C, then A is part of C
- **Reflexive**: A is part of A
- **Antisymmetric**: If A is part of B and A != B then B is not part of A.

In our setting, an example of a transitive relation is “Learning Activity” being part of “Teaching Unit” which is part of an “Educational Module”. During the conversion process we have encountered different types of part-of relations that we modeled using different OWL relationships. For example, we can see that a professor is part of a department. In order to distinguish between different part-of relations, we have used a set of OWL relationships to model them in our ontology (e.g., “Professor” - *bb:isFromDepartment* - “Department”). An example of a different type of a part-of relation is the Evaluation class where a student, a professor, a teaching unit, and a grade are components of the evaluation concept in the sense that, together, they create an instance of an evaluation. For this case, we have created appropriate relationships to model such scenario: “Evaluation” - *bb:performedByStudent* - “Student”, “Evaluation” - *bb:evaluatedByProfessor* - “Professor”, and “Evaluation” - *bb:evaluatesTeachingUnit* - “Teaching_Unit”. Moreover, given the lack of support of OWL for different types of part-of relationships, we took particular care while modeling them to avoid problems such as creating cycles in the part-of graph.

5. The Bowlogna Ontology

The outcome of the process described in the previous two sections is the Bowlogna ontology. This ontology describes the university setting as defined by the Bologna reform started in 1999. Figure 1 shows some of the classes in the Bowlogna ontology together with their relations. Different arrow colors indicate different relation types. For example, purple arrows (e.g., between Person and Professor) indicate a *rdfs:subClassOf* relation, green arrows (e.g., between Student and Major Area) indicate a *bb:follows* relation. The ontology currently contains 25 top level classes (66 in total) describing concepts like students, professors, student evaluations (i.e., exams where students get a grade from a lecturer in the context of a teaching unit), teaching units (i.e., courses given by lecturers and attended by students), ECTS credits, as well as formal documents such as, for instance, study programs, certificates, or grade transcripts. Additionally, succinct definitions for all the concepts included in the original lexicon are available in German, French, Italian and English in the ontology.

One of the key classes in this ontology is “Evaluation” in which a Student, a Professor and a Teaching Unit are involved. This class models the event when a student, after attending a course given by a professor (i.e., a teaching unit) is evaluated and given a grade. Properties of teaching units include the semester and the language in which the units are given, as well as the number of corresponding ECTS credits earned by students who successfully follow the course. The class Student also has a number of properties including the name of the student, his enrollment or graduation dates for Bachelor and Maser degrees. The same student instance can be, over time, registered for different academic degrees (e.g., first a Bachelor and then a Master). The class “ECTS credit” is linked to all classes directly or indirectly related to the measurement of students advancement: for example, teaching units and study programs. The class “Semester”, with start and end dates, enables queries on the temporal dimensions (e.g., the number of students enrolled at the university during a certain semester).

The full ontology can be accessed at the following URL: `http://diuf.unifr.ch/xi/bowlogna/`

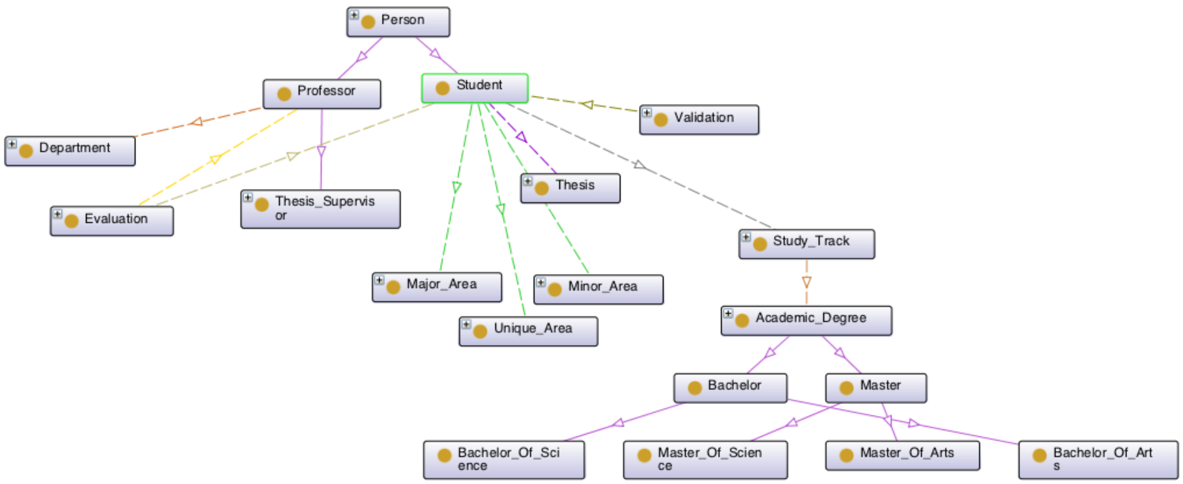

Fig. 1. The Bowlogna ontology: key classes and relations.

5.1. *Public and Private Ontology Parts*

The Bowlogna ontology can be divided in two important parts according to the type of information stored. Some information are public and can be shared with other universities as well as with the general public. Examples of such public information include the departments, the teaching units together with information about their ECTS credits and teaching language. The second part of the ontology consists of information which should not be publicly available, such as grades given to students. In real instantiations of the ontology, the private part will contain many more instances as compared to the public part. It is also clear that aggregations over private information might often be of general interest (e.g., the number of students currently enrolled at the university), and can be safely shared.

While such a separation is made clear at the ontology level by defining which classes, sub-classes, properties, and sub-properties are public or private, Bowlogna does not incorporate any security mechanisms at this point. Security, privacy, and access policies are obviously of utmost importance when deploying information systems based on Bowlogna, but they are infrastructure and application-dependent and as such cannot be fully specified in the ontology from our perspective.

5.2. *Multi-Linguality*

One of the key properties of Bowlogna is the definition of all its concepts in four different languages: English, French, Italian, and German. This has been done as the University of Fribourg, which is a trilingual institute (where all Bachelor-level courses are given in both German and French, while Master courses are given in English) in a trilingual country (French, German and Italian are all official languages in Switzerland[5]).

We kept all the original definitions available in the lexicon (which defined all terms in both German and French) for our ontology. Classes are linked to their multi-lingual concept definitions. Additionally, we have included in the ontology English and Italian definitions to ease its use at the European level. In order to cover the most important languages involved in the European higher-education landscape, we aim at extending the set of languages used for the definitions in the near future.

5.3. *Related Ontologies*

Other ontologies describing similar domains exist. For example, the “Metadata for Learning Opportunities” (MLO) is a model (also available as RDFS) which can be used to create and share course descrip-

[5]Romansh is the forth official Swiss language but only account for 0.5% of the population.

tions. Compared to it, Bowlogna adopts a course (i.e., teaching unit) modeling approach which is based on the Bologna process definition of teaching units. Mappings between the two ontologies are naturally possible (as for example *http://purl.org/net/mlo/assessment*, which is similar to *bb:Evaluation*). Another related project is "eXchanging Course Related Information"[6] (XCRI) where a language to share course descriptions is defined. Possible mappings with Bowlogna exist as well: for example, *http://xcri.org/profiles/catalog/1.2/course* is related to *bb:Teaching_Unit*. The "Academic Institution Internal Structure Ontology"[7] (AIISO) has a different focus, which targets the structure of an academic organization. An example mapping to Bowlogna is *http://purl.org/vocab/aiiso/schema#Module*, which relates to *bb:Teaching_Unit*. Finally, LinkedUniversties[8] is an initiative aiming at listing all available vocabularies and resources on linked data for higher-education.

6. A Standard Ontology for European Universities: Usage Scenarios

Once created, the Bowlogna ontology can be deployed to automate information workflows and facilitate data exchange in both open and closed higher education settings. We detail below how we took advantage of the ontology at the university of Fribourg to power next generation services for various members of the academic community, and how open higher education interoperability can be fostered by adopting or mapping to the ontology.

6.1. A Faceted Search and Browsing System over Bowlogna

We now describe a first system that we built on top of the Bowlogna ontology. The application scenario is the following. In an university administration, there is a constant need to access the large amounts of data stored about students. In the case where the institution is using the Bowlogna ontology for storing such data, it is possible to adopt external search and browsing tools supporting RDF or OWL. As a proof-of-concept, we have developed a search and browsing interface over the Bowlogna schema using the SIMILE libraries[9]. With such an interface, it is easy to find, for example, all exams performed by a specific student, all the professors affiliated with one department, or the courses taught in one specific language[10]. A screenshot of the system is shown in Figure 2.

[6] `http://www.xcri.co.uk/`
[7] `http://vocab.org/aiiso/schema`
[8] `http://linkeduniversities.org/`
[9] `http://simile.mit.edu/`

6.2. Agile Knowledge Bases for University Administration

Beyond the simple search system described above, the management of data across research groups and departments inside a university is increasingly becoming difficult, both because of the rapid augmentation of information to manage, and the increasing complexity of the processes and queries to support. Our current information infrastructure at the university of Fribourg is based on standard commercial relational database technologies; over the years, two main data management problems have surfaced and are today severely hampering information workflows inside the university:

Super-linear increase in data: the relational system currently in place is used to store data for both current course offerings, registered students, etc. and also to keep historical data about the university. Year after year, the constant augmentation of students and academic programs leads to a steady increase in the sheer amount of data to manage in the database. Semi-automatic processes used so far to extract and analyze data, such as CSV export of historical course offerings and manual inspection of the data through a spreadsheet, are no longer possible today (the database contains for example several thousands of entries in the various "course" tables). Tackling this problem requires a complete overhaul of the current data infrastructure inside the university, which is technically difficult since relational data were never meant to be shared across systems (i.e., transferring relational data from one system to the other is a challenging process typically implying some loss of data either because of unsupported data types, constraints, or event handlers).

Multiplication of Relations: As another consequence of the constant augmentation of course offerings and the complexification of teaching organiza-

[10] A demo of our faceted search system over Bowlogna is available online at: `http://diuf.unifr.ch/xi/bowlogna/`

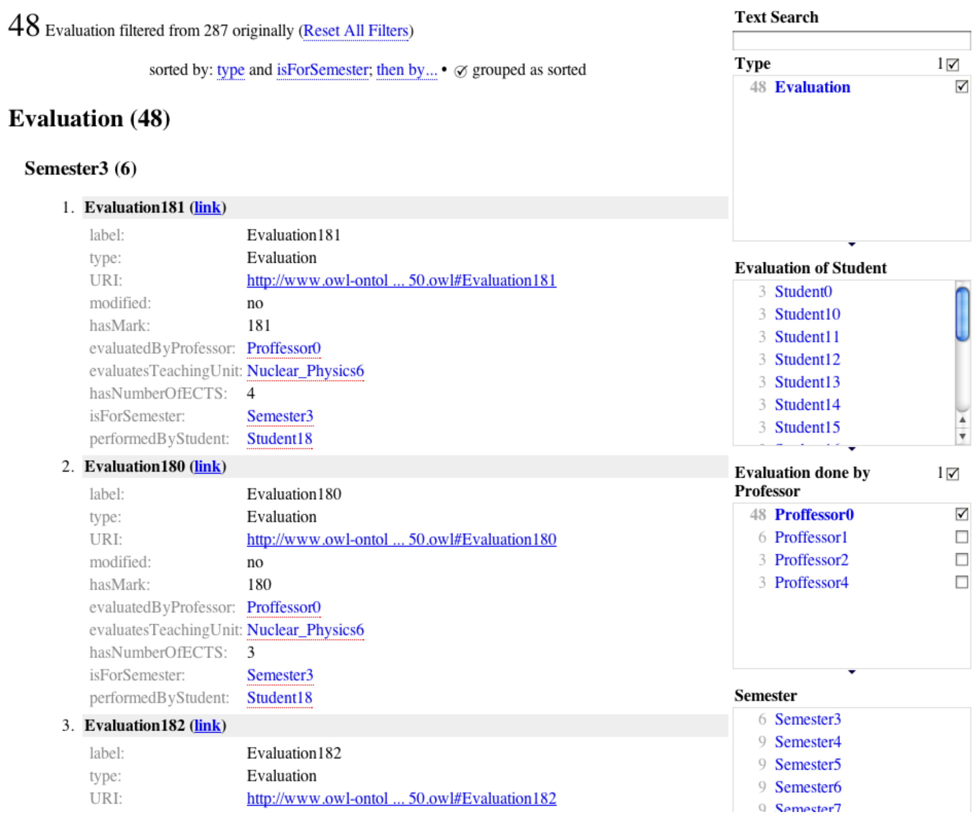

Fig. 2. A faceted search and browsing system over the Bowlogna ontology. In the figure we can see the selection of all the evaluations done by a specific professor.

tions, the number of tables used by the relational engine to store data is also soaring. There are for example more than fifty relations used to represent students, accounting for various study programs, part-time students, external students, students registered to special programs, external auditors, etc. The relationship between those tables are never formally described, though the tables are often implicitly linked through automated triggers (e.g., adding a student to the "Part-Time Students" tables automatically triggers the addition of the student in other student tables as well). This implicit increase in schema complexity is today becoming unbearable and creates awkward situations where relatively straightforward queries such as "what is the total number of students registered at the university?" cannot be answered precisely today since the query touches a web of relations, whose semantic relationships are only vaguely and implicitly defined.

The two problems described above are both tackled by the proof-of-concept system that we are building around Bowlogna. As our infrastructure is centered around semi-structured OWL data, sharing or exporting both schema and instance data is straightforward, as all data can be exchanged using standardized serialization mechanisms (for example through an XML serialization). Thus, offloading data to specialized tools, for instance for large-scale reasoning [1] or business intelligence tasks [17] is straightforward. As for the multiplication of relations, Bowlogna provides two key advantages: i) it allows for more dynamic schemas, where new concepts or relationships can be easily added without having to modify or update former conceptualization (e.g., by adding new object properties to an existing ontology, or adding new properties linking pairs of classes) and ii) it defines the relationships between classes through formal constructs (e.g., using class hierarchies or formal properties such as *owl:TransitiveProperty*). Taking advantage of inference mechanisms, queries such as the one mentioned above ("what is the total number of students registered at the university?") are then straightforward to answer, and generally boil down to simple selections/projections on one class in the class hierarchy (i.e., a selection on the "student" superclass).

Another salient problem in current information management systems inside universities is the difficulty to trace information lineage. For example, it is not easy to trace all exams (successful and not) a student has been taking since she started her academic studies, including also the ones given by foreign universities during Erasmus exchanges. The gradual adoption of a standard ontology within European universities would enable an easier traceability of the data. In our Bowlogna ontology it is, for instance, possible to run temporal queries (e.g., "what is the percentage of students who continue their university studies beyond the Bachelor?") as instances contain time-related properties which allow to trace, for example, student activities from their first registration to the end of their studies.

6.3. Large-Scale Curricula Integration for European Students

Transparent integration of course offerings is another breakthrough which can be made possible by the adoption of Semantic technologies by universities. Today, while basically each university in Europe provides some HTML pages to describe its courses and programs, no common infrastructure exists to automatically integrate all curricula across Europe, in order for example to provide a single query interface to prospective students. Integrating heterogeneous data displayed on HTML pages is an extremely intricate problem [10]. The adoption of Semantic technologies to manage higher-education data offers a number of advantages in this context. The systematic use of URIs to identify both conceptualizations and data instances, as well as the availability of standard publication mechanisms such as SPARQL end-points enable the publication of vast quantities of structured data on the Web. Once the course offerings are available as raw data, the *a posteriori* integration of information can be made possible by schema [3,5] and instance matching [4] techniques. The publication of a reference ontology for course offerings (like the Bowlogna ontology) is a key-element to minimize the amount of manual work required to integrate data, as it can be used as a central hub or *mediator* [16] to facilitate the mappings between various university systems.

Once data mappings are in place, automated query reformulation or gossiping mechanisms [2] can be used to create transparent data integration and a unique query interface for the end-user, for example to create faceted search functionalities such as those described above to let students browse all potential courses offered in European universities. Beyond the extension of the simple search system described above to include several institutions, one can easily envision the potential benefits of using or mapping to a unique ontology across all higher education institutes in Europe. This would for instance open the door to the adoption of standard software built on top of the common conceptualization as well as enable the publication and sharing of information among institutions. Semantic data integration also opens the door to fully automated data exchange [6], where universities following the Bologna precepts could for instance exchange grade or ECTS information automatically and securely [9] across heterogeneous and distributed information systems.

7. Conclusions

The Bologna process, which started in 1999, established a new era for higher education across Europe. Many institutions adapted their course offerings to comply to the standards suggested by this reform. This harmonization process has taken several years to unfold. As a consequence, we can today observe striking similarities among very disparate universities in Europe.

To support this renovation process from the point of view of information systems, there is an urgent need for standard data representation in order to enable automated information flows and data sharing among institutions.

In this paper, we described Bowlogna, an ontology modeling an academic setting as defined by the Bologna reform of higher education in Europe. We presented the efforts made by terminologists to create a lexicon regrouping the terms used to describe the academic setting as prescribed by the Bologna reform. We then described the semi-automated process of converting such a lexicon to an ontology, which could then be used as a standard conceptualization for modeling data about students, courses, evaluations, etc. We also showed how such an ontology could underpin various applications tackling crucial problems of today's IT systems in European universities. Finally, we presented our current faceted-search prototype, and discussed how Bowlogna could be used to implement more agile and integrated information platforms for managing, integrating and serving higher education data across Europe.

7.1. Lessons Learned

We described above the creation process of a lexicon of concepts capturing the Bologna reform, its conversion to a standard ontology, and the application of such an ontology in various contexts (e.g., a faceted search application). During those various steps, we made important observations which can be shared as practical guidelines for transitioning from a legacy relational system to a next-generation Semantic Web infrastructure in a university. In detail, we can share the following recommendations:

1. Asking linguistic experts to model the domain of discourse through a direct and careful analysis of the sources describing the domain is essential in order to extract a comprehensive set of key concepts:

   – In the domain we analyzed, one critical aspect was the variety of culture and languages involved. This lead us to the decision that all terms present in the lexicon should be defined in different languages to ease communication among the different actors involved. Currently, the Bowlogna ontology include definitions in English, German, Italian, and French.
   – Another important aspect is the interdisciplinarity that the Bologna process implies. Different disciplines such as pedagogy, management, and sociology are involved. It is therefore necessary to consider this aspect while modeling the domain and to invest efforts in maintaining natural language description of all important conceptualization in order to foster communication among different disciplines.

2. Transforming a lexicon of concepts into an ontology is a non-trivial task and must again involve domain-experts for optimal results:

   – Majors efforts had to be invested in understanding which concepts and relations from the lexicon should be kept in the ontology, and which should be discarded.
   – Not all the terms defined in the lexicon had to be translated to ontology classes, even though our goal was to obtain a high coverage of the domain of discourse.
   – Particular care had to be taken to model different types of part-of relations used rather informally in the lexicon.

3. If properly defined, the common conceptualization foster information exchange and data integration through various higher-level applications:

   – In the case of end-user applications, most of the user requirements were already defined during the lexicon definition and the ontology conversion. This allowed quick prototyping of further applications thanks to the careful modeling of the domain of discourse.
   – Once available, a formal conceptualization can considerably lower the efforts required to integrate heterogeneous information sources, by acting as a common representation or as a central mediator upon which local schemas or ontologies can be mapped.

7.2. Future Directions

As mentioned in the previous sections, we envision several important future directions involving the Bowlogna ontology. Most importantly, we plan to push for the gradual adoption of the Bowlogna ontology as an integral part of the overall description of the Bologna reform, in order to lower the efforts required to integrate information systems across European universities. As a first step forward, we are currently assessing the various benefits of adopting such an ontology in our local institution. We have developed a demonstrator based on the Bowlogna ontology that allows university administration employees as well as students to navigate and query real data from our institution. In this way, (prospective) students can be better informed about the current educational offers (e.g., “Which degrees can I obtain in Fribourg about History?”) at our University. We hope to gather experience on the potential pitfalls and risks related to the wide deployment of Semantic technologies in a university setting, and on the safeguards to put in place in order to allow for a smooth transition from legacy systems. We are also currently approaching various higher education institutions in Europe in order to push for the wider adoption of Semantic technologies in university settings, and hope to deploy a first distributed prototype supporting the sharing and exchange of information between a small set of institutions in the near future.

8. Acknowledgment

This work is supported (in part) by the Swiss National Science Foundation under grant number PP00P2_128459.

References

[1] Barry Bishop, Atanas Kiryakov, Damyan Ognyanoff, Ivan Peikov, Zdravko Tashev, and Ruslan Velkov. OWLIM: A family of scalable semantic repositories. *Semantic Web - Interoperability, Usability, Applicability*, 2(1):33–42, 2011.

[2] Philippe Cudré-Mauroux. *Emergent Semantics*. EPFL & CRC Press, 2008.

[3] Philippe Cudré-Mauroux, Karl Aberer, and Andras Feher. Probabilistic Message Passing in Peer Data Management Systems. In Ling Liu, Andreas Reuter, Kyu-Young Whang, and Jianjun Zhang, editors, *Proceedings of the 22nd International Conference on Data Engineering, ICDE 2006, 3-8 April 2006, Atlanta, GA, USA*, page 41. IEEE Computer Society, 2006.

[4] Philippe Cudré-Mauroux, Parisa Haghani, Michael Jost, Karl Aberer, and Hermann de Meer. idMesh: Graph-based disambiguation of linked data. In Juan Quemada, Gonzalo León, Yoëlle S. Maarek, and Wolfgang Nejdl, editors, *Proceedings of the 18th International Conference on World Wide Web, WWW 2009, Madrid, Spain, April 20-24, 2009*, pages 591–600. ACM, 2009.

[5] Jérôme Euzenat. An API for Ontology Alignment. In McIlraith et al. [11], pages 698–712.

[6] Ronald Fagin, Phokion G. Kolaitis, Renée J. Miller, and Lucian Popa. Data exchange: semantics and query answering. *Theor. Comput. Sci.*, 336(1):89–124, 2005.

[7] Joël Gapany and Guido Vergauwen. Curricular design and computerisation: are information systems useful in curricular reorganisation? *European University Association: the Bologna Handbook*, 10 (C 3.9-3), 2010.

[8] Nicola Guarino and Christopher A. Welty. An Overview of OntoClean. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, pages 151–159. Springer Verlag, 2004.

[9] Amit Jain and Csilla Farkas. Secure resource description framework: an access control model. In *Proceedings of the eleventh ACM symposium on Access control models and technologies*, SACMAT '06, pages 121–129, New York, NY, USA, 2006. ACM.

[10] Jayant Madhavan, David Ko, Lucja Kot, Vignesh Ganapathy, Alex Rasmussen, and Alon Y. Halevy. Google's Deep Web crawl. *PVLDB*, 1(2):1241–1252, 2008.

[11] Sheila A. McIlraith, Dimitris Plexousakis, and Frank van Harmelen, editors. *The Semantic Web - ISWC 2004: Third International Semantic Web Conference,Hiroshima, Japan, November 7-11, 2004. Proceedings*, volume 3298 of *Lecture Notes in Computer Science*. Springer, 2004.

[12] Dagobert Soergel, Boris Lauser, Anita C. Liang, Frehiwot Fisseha, Johannes Keizer, and Stephen Katz. Reengineering Thesauri for New Applications: The AGROVOC Example. *Journal of Digital Information*, 4(4):1–23, 2004.

[13] Lina Fatima Soualmia, Christine Golbreich, and Stéfan Jacques Darmoni. Representing the MeSH in OWL: Towards a semi-automatic migration. In Udo Hahn, editor, *KR-MED 2004, First International Workshop on Formal Biomedical Knowledge Representation, Whistler, BC, Canada, 1 June 2004*, volume 102 of *CEUR Workshop Proceedings*, pages 81–87. CEUR-WS.org, 2004.

[14] Mark van Assem, Véronique Malaisé, Alistair Miles, and Guus Schreiber. A Method to Convert Thesauri to SKOS. In York Sure and John Domingue, editors, *The Semantic Web: Research and Applications, 3rd European Semantic Web Conference, ESWC 2006, Budva, Montenegro, June 11-14, 2006, Proceedings*, volume 4011 of *Lecture Notes in Computer Science*, pages 95–109. Springer, 2006.

[15] Mark van Assem, Maarten R. Menken, Guus Schreiber, Jan Wielemaker, and Bob J. Wielinga. A Method for Converting Thesauri to RDF/OWL. In McIlraith et al. [11], pages 17–31.

[16] Gio Wiederhold. Mediators in the Architecture of Future Information Systems. *IEEE Computer*, 25(3):38–49, 1992.

[17] Marcin Wylot, Jigé Pont, Mariusz Wisniewski, and Philippe Cudré-Mauroux. dipLODocus[RDF] - Short and Long-Tail RDF Analytics for Massive Webs of Data. In Lora Aroyo, Chris Welty, Harith Alani, Jamie Taylor, Abraham Bernstein, Lalana Kagal, Natasha Fridman Noy, and Eva Blomqvist, editors, *The Semantic Web - ISWC 2011 - 10th International Semantic Web Conference, Bonn, Germany, October 23-27, 2011, Proceedings, Part I*, volume 7031 of *Lecture Notes in Computer Science*, pages 778–793. Springer, 2011.