Towards a Linked and Reusable Conceptual Layer around Higher Education Programs

Fouad Zablith

Olayan School of Business
American University of Beirut
PO Box 11-0236, Riad El Solh, 1107 2020
Beirut, Lebanon
fouad.zablith@aub.edu.lb

Abstract. From a knowledge representation perspective, higher education programs can be exhibited as a set of concepts exchanged in learning environments to achieve specific learning objectives. Such concepts are often grouped in different blocks such as courses, modules or topics that count towards a degree. Currently, most of program representations are performed at a high level in the form of course descriptions that are part of course catalogs, or at more specific levels using for example course syllabi. While this is great for informative purposes, systematically capturing, processing and analyzing the concepts covered in a program is not possible with this text-based representation. We present in this chapter a data model based on the linked data principles to create a conceptual layer around higher education programs. We follow a collaborative approach using a semantic Mediawiki to build the knowledge graph around a business school curriculum. The impact of this linked open layer is highlighted at the level of (1) enriching online learning environments by extending the graph through learning material and selectively pushing it to existing course pages; and (2) enabling a more in-depth analysis of the program during review activities.

1 Introduction

Higher eduction is undergoing major changes in the way students and teachers interact. Such changes are highly dictated by the latest advancement of mobile, Internet and Web technologies. Those advancements are changing the perception of information consumers. The expectations from what can be done today in a classroom, offline as well as online, are increasing. The ease of access to information, coupled with the ever-improving sophistication of computing devices are opening new opportunities for educational environments. For example we have seen lately the increase in the number of open online courses, where teachers can reach out to thousands of students in a course deliverable. Another example is the availability of blended learning options that universities are also experimenting with.

While technology is infiltrating higher educational settings, one major question that is yet to be answered is how can existing forms of curricula representation cope with information that is increasingly being exchanged in online environments? Most higher education curricula are currently represented in textual formats, where concepts covered are buried deep into course catalogs and syllabi. This form of representation works well for describing courses and degrees, however it fails when more sophisticated and systematic processing is required. Furthermore, inferring cross connection among courses, or between courses and learning material is a labor-intensive process. This will result in creating boundaries around courses that are hard to break not only in traditional classroom settings, but also in online environments. In other words, a textual representation of curricula cannot keep up and move at the pace of how learning environments are evolving.

We believe that the goodness provided by the latest efforts of the linked data community can be exploited to improve curricula representation. We present in this chapter[1] our effort towards creating a linked layer around a higher education program. We follow the linked data principles [4] to create unique reference-able entities of the concepts exchanged in courses. We propose a data model that we used as a starting point to build the linked data graph. We deploy a semantic Mediawiki [8] to collaboratively generate and inter-connect the linked data graph around a Business School curriculum. We consume the generated data through two pilot applications. The first is in the context of enriching online learning environments, in which we use moodle to showcase how relevant material and information can be easily integrated in course pages without any modifications to the existing platform. The second is by processing the data to present new unprecedented views of the curriculum that can potentially support the program review process at the business school.

This chapter is organized as follows. We first present an overview of existing work in the field in Section 2. The data model is then presented and discussed in Section 3. Then in Section 4 we present the process of generating the linked data graph through the semantic Mediawiki that we implemented. After representing the curriculum through the linked data layer, we present two use-cases of the data in Section 5, and conclude in Section 6 with potential future research directions.

2 Related Work

The web has been witnessing tremendous changes recently. When the term Semantic Web was first coined by Tim Berners-Lee, it was made clear that the move from links between documents to links between objects will "unleash a revolution of new possibilities" [1]. The objective is to have a web of connections that computers can "understand". In other words entities will be represented and connected through explicit meanings using well-defined vocabularies. Now after around fifteen years, the impact of having "semantics" added to the web

[1] This chapter is an extended version of the paper published in the WWW2015 companion proceedings [11].

is obvious. We have seen governments creating and opening up their data on the web with explicit semantics. For example the "data.gov" platform of the US government that today includes around 164,112 accessible and processable datasets. Similarly the UK government initiated their "data.gov.uk" platform for similar purposes, and other governments are now following this trend.

Part of the Semantic Web community effort was the creation of linked data publishing principles [4] that can be summarized as follows:

- To use unique resource identifiers (URI) to uniquely identify abstract concepts and real world objects
- To generate URIs that can be dereferenced on the web through the HTTP protocol
- To publish the data using the Resource Description Framework (RDF) model [6], which is a graph-based data representation that can be easily queried and retrieved
- To have hyperlinks in the form RDF links between entities represented as linked data

The simplicity of such principles contributed to the continuous expansion of the available linked open data cloud [9]. In addition to the government examples mentioned earlier, linked data is getting increasingly adopted in various contexts. Higher education is no exception. In their survey, Dietze et al. [10] highlight the growing adoption of linked data by various universities. They identify the challenges that lie ahead of the technology enhanced learning community for a wider embrace of linked data. This includes (1) the integration of various heterogeneous repositories; (2) the adaptation to the changes underneath services provided such as wed APIs; (3) the mediation and mapping of meta data across educational resources; and (4) interlinking and enriching unstructured data coming for example from text documents. While various research efforts are being invested in the above challenges, educational institutions cannot but benefit from the added value of this linked data layer. In the educational field, various platforms have emerged where linked data is made available for direct consumption and reuse. This includes for example the OU's linked open data platform (`http://data.open.ac.uk`), the University of Muenster (`http://data.uni-muenster.de`), the University of Oxford (`http://data.ox.ac.uk`), the University of Southampton (`http://data.southampton.ac.uk`), among others.

Methodologies and frameworks have been proposed to transform existing data sources into linked data [3,12,13]. In this context, available organizational data was transformed, using pre-programmed transformation patterns, into linked data. A LUCERO framework was proposed to perform the following main tasks [13]: *collect* data from various sources of organizational data, where a scheduler automatically checks for data updates through for example Really Simple Syndication (RSS) feeds; *extract* the data and transform it into linked data following predefined URI creation rules; *link* the generated data with internal and possibly external data; *store* the data in a triple store and *expose* it for consumption through a SPARQL endpoint.

Following the trend and success of social and knowledge graphs functionalities provided by Facebook[2] and Google[3], there are discussions around the value of having an education graph[4] [5]. Heath et al. [5] proposed in their work to create an education graph by processing courses information and learning material from various universities in the UK. In their approach, they mainly rely on bibliographical data of material repositories to identify links to course resources [5].

In our context, we focus in this chapter on interlinking courses within the same institution at the level of concepts covered in course topics. To our knowledge, there is no existing vocabulary that covers this granular information about courses information. To achieve our objective, we had to go at lower (i.e. more specific) conceptual levels by enabling users to have a direct impact on reshaping how courses interconnect among them and with learning resources. This controlled environment is necessary for aligning the conceptual coverage of courses delivered by more than one instructor. We also aimed to have direct input from students to connect to and expand the graph around the education program. For example when students find an interesting material online (e.g. video or article) relevant to a specific course, we wanted to enable them to connect it back to the course by extending the graph and creating the appropriate links to the course.

3 Linked Data Graph

We present in this section our proposed data model to capture course related information. We visualize the data model in Figure 1. The sources of information can be grouped in two main parts: the first is coming mainly from the courses syllabi, aggregated in the upper part (A) of Figure 1, while the lower part (B) of Figure 1 captures information mostly from learning material.

We used the information included in the courses' syllabi template as a starting point to identify the course elements to represent. A typical course syllabus follows a predefined template, which includes the course details such as course number, description, prerequisites, textbook, topical coverage and others. Furthermore, each course has well-defined learning objectives.

We focus in this part on the description of specific aspects that guided the development of the elements that are not covered in the syllabi content. Based on the need for conceptually connecting courses (i.e. beyond the analysis of topics in common), we captured in the graph concepts that are taught at the level of every course topic. This design offers many advantages. First we are getting a more granular view of what is covered in each course. Second, such concepts can be used as anchors between learning material and courses. This will enable course designers to know where exactly each piece of material fits in the course, and this enables learning material to float around not only course topics, but also around the program as a whole. In other words an interesting article that

[2] http://newsroom.fb.com/News/562/Introducing-Graph-Search-Beta
[3] http://www.google.com/insidesearch/features/search/knowledge.html
[4] http://hackeducation.com/2011/11/10/is-there-an-education-graph

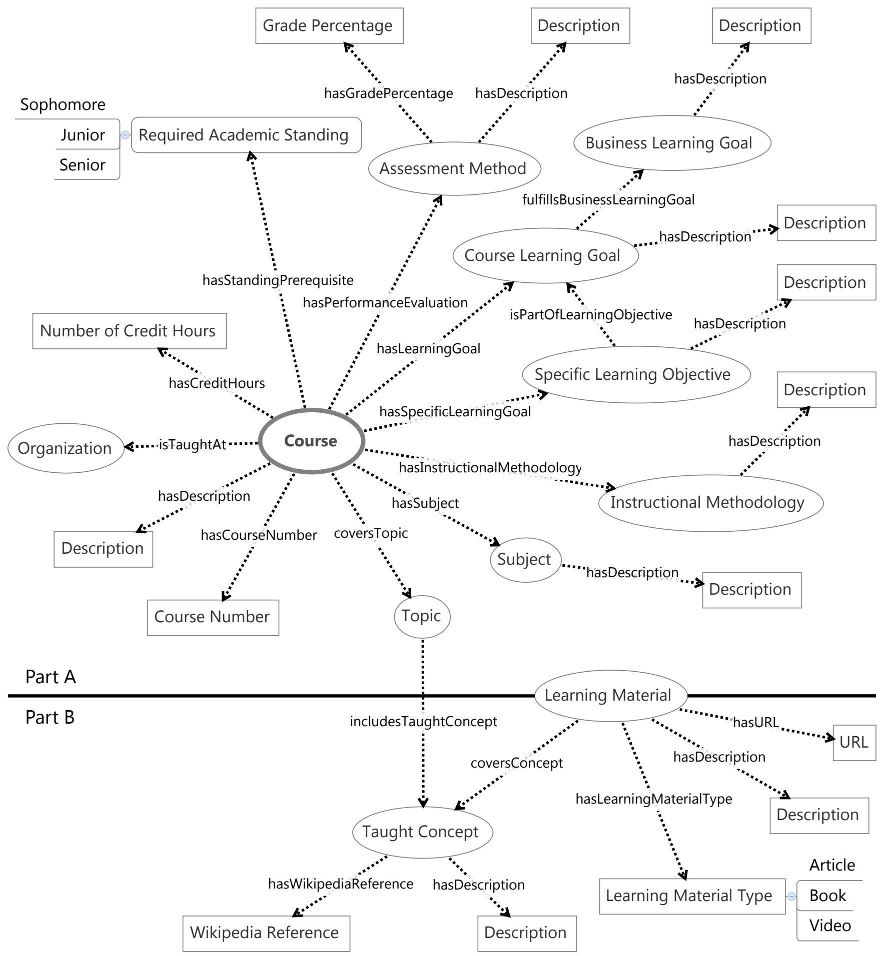


Fig. 1. The Linked Data Graph.

is used in one course, can also be relevant in other courses if the concepts are shared between the courses.

Following the linked data principles [4], we aimed to reuse the available ontologies that are relevant to our context. For example the CourseWare ontology[5] was reused to represent information related to course types, student interaction types, number of credits, assessment methods, and others. The AIISO[6] was adopted to capture the courses' unique codes. We have also used the Dublin

[5] `http://courseware.rkbexplorer.com`
[6] `http://vocab.org/aiiso/schema`

Core Terms[7] to represent generic properties such as descriptions. In cases where no vocabularies were found, we created a local vocabulary used in our context.

4 Linked Data Generation

After defining our linked data model, we discuss in this part our effort on creating linked data around the higher education program of the School of Business of the American University of Beirut. We followed a collaborative process to build the linked data graph around the program. This involved professors, students and program coordinators. As courses are delivered by different professors, we required to have a system that supports collaboration, through which each professor or involved person can see, expand and modify content as seen appropriate. Another requirement was to enable students to also enrich and add to the graph in a quick and easy way online material that they find interesting. In addition to the above, another needed feature was the ability to control linked data vocabularies. We discuss next our semantic Mediawiki implementation, followed by the steps we took to generate the data using the wiki and a semantic bookmarklet for linking online resources to the courses.

4.1 Semantic Mediawiki Deployment

We have implemented a semantic Mediawiki available at `http://linked.aub.edu.lb/collab`. For controlling the data generated and vocabularies used, we created several forms that are automatically loaded when new content is to be added. In the case of creating new courses for example, the course form[8] is automatically loaded[9]. The advantage of using such forms is that the user is guided around what fields to fill, and the vocabularies used can be predefined in the form when the RDF is generated.

Linking data is done by controlling the field content in the wiki. By adopting the wiki forms, the fields can guide users in reusing existing concepts from the wiki. For example when specifying the course prerequisite relation, the user is prompted with the list of courses available in the knowledge space that can be chosen from. This is a core feature used for the interlinking process. Another example is at the level of concepts covered in course topics. The same concept can be covered in one or more course topics, creating the links needed.

4.2 Steps Followed for Building the Linked Data Graph

Building the linked data graph was done in three phases. In the first phase, course syllabi are used as entry points, where the high level course information is

[7] `http://purl.org/dc/terms`

[8] `http://linked.aub.edu.lb/collab/index.php/Special:FormEdit/Course/New_Course`

[9] For a filled form example, check the "Foundations of Information Systems" course at the following link: `http://linked.aub.edu.lb/collab/index.php?title=INFO200_-_Foundations_of_Information_Systems&action=formedit`

entered. In the second phase, textbook materials used in the course are processed by the teaching assistants (TAs) to identify the concepts covered in the course. In the third phase, new external materials are added to the graph by instructors and students, using a semantic bookmarklet.

Phase 1: Creating Courses Information. This phase was straight forward, as existing course syllabi follow a predefined structure. This part was mainly extending Part A of Figure 1.

Some difficulties were faced at the level of identifying learning goals, as some of the courses learning goals were described as text, without a clear structure. Another complexity at this level was that course learning goals were at two levels (as depicted in Figure 1). Courses have specific learning goals, which are linked to a broader list of business learning goals identified within our School of Business. This required a two-step data entry to capture linkages among learning objectives.

Another challenge we faced is when the topics of courses change with time, due for example to changes in textbook editions, or in the delivery of content across semesters. For instance, in the Foundations of Information Systems course, the Social Media topic covered in the "Experiencing MIS" [7] textbook has changed from the 3^{rd} to 4^{th} edition. In this case, while the topical coverage has changed, some of the concepts that were covered in the previous edition were still there. We handled such cases by archiving topics and removing the corresponding links to the course. This way we were able to preserve the concepts related to the old topic, and reuse them if needed in the new topic.

The evolution of changes at this level can be better managed in the future. One possible improvement can be done through capturing temporal changes, coupled with the type of changes performed (e.g. adding or removing concepts from the concept graph). While tracing such evolution patterns can be done at the data entry level, a post analysis of changes occurring on the data graph can be possibly performed. Such features are beneficial for analytics applications, and can be further explored as part of our future research. We have processed so far the 19 core courses offered at the School of Business, leaving the elective courses to be represented at a later stage.

Phase 2: Identifying Concepts Taught in Courses. This phase was the most extensive and time consuming phase. The focus at this level was on capturing the concepts covered in the core learning material of courses such as textbooks mentioned in the courses syllabi.

TAs were trained to identify concepts covered in the topics of the courses. Based on the model we created (cf. Part B of Figure 1), concepts are linked to the topics of a course, and not directly to the course itself. This choice of design enables grouping concepts by topic, rather than by course. This has a practical implication in filtering concepts covered in specific topics (as we discuss later in the chapter when we integrate data in moodle). Another implication is at the level of program analytics. Overlap among course topics can be highlighted easily, which proved to be useful for the program review and design exercises.

At this level, the TAs were going into each topic covered in the textbook to identify the main concepts, adding definitions from the book, and linking to a Wikipedia[10] reference when found. When applicable, the TAs were instructed to reuse existing concepts in the repository. The *concepts* field in the *topic* form automatically provides the TAs with the list of existing concepts to choose from.

One challenge at this level is when concepts from different topics share the same name, however are semantically different. For example, the *Optimization* concept covered in the *Managerial Decision Making* (i.e. operations research), is semantically different from the optimization concept in the *Managerial Economics* field. This is where Wikipedia is used as an external reference for disambiguating such cases. While linking the concepts to Wikipedia entries is currently done manually, this task can potentially be performed (semi)-automatically by querying DBpedia [2] and aiming to find overlap between the concepts' definition and the description in the DBpedia page. In addition to text matching, the graph of the concepts can be used as a context to highlight the degree of matching. Then the user can browse the proposed matching concepts to select the most appropriate one.

Currently concepts identified are not related through an explicit relation. We are planning to capture in the future relations such as sub-class and prerequisite relations among concepts. This can also have an impact on the analysis of course sequencing in the program. So far we have identified around 2,680 concepts[11] covered in the core courses of the School of Business.

Phase 3: Semantically Anchoring Learning Material to Courses. While the previous phases were mainly focusing on reorganizing internal knowledge sources, this phase is more about linking new materials to the program, by anchoring them to concepts covered in courses. Here the aim is to enable students and instructors to link interesting online material to the graph.

We developed a simple bookmarklet that can be installed in any browser. When the user links a learning material, clicking the bookmarklet will automatically extract the page link, title and description. The user is then prompted to a page pointing to our wiki platform, where the concepts covered in the material can be entered by reusing existing concepts from the graph[12].

As mentioned earlier, links between a learning material and the program are done through the concepts. This somehow enables educators and students to think more around the relevance of the material around concepts. For example we witnessed a student who bookmarked an article related to Big Data[13], relevant to the Foundations of Information Systems course she was enrolled in. While anchoring this article, she reused concepts from the graph that indirectly spread

[10] www.wikipedia.org

[11] The full list of concepts can be accessed at: `http://linked.aub.edu.lb/collab/index.php/Category:Learning_concepts`

[12] A video tutorial on how to use the bookmarklet can he accessed at: `http://linked.aub.edu.lb/docs/tutorial_material_bookmark`

[13] `http://www.capgemini.com/resources/the-deciding-factor-big-data-decision-making`

to two other courses. Starting from an information systems' related material, the student indirectly linked to a management and operations management courses. The full list of concepts highlighted by the student can be found in Figure 2.

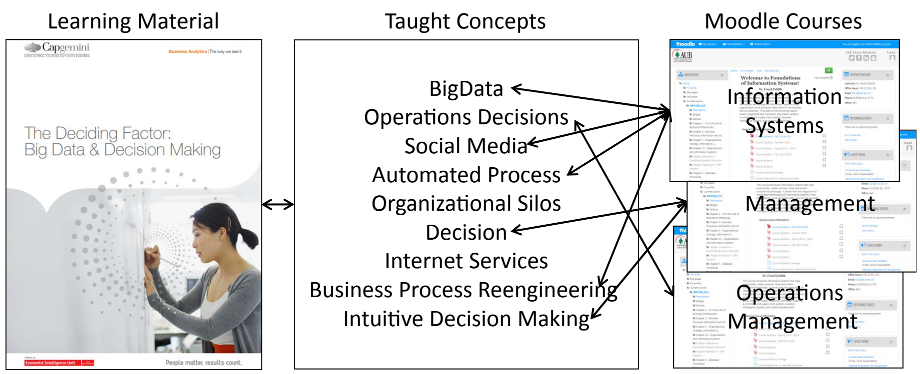


Fig. 2. Learning Material Cross Connecting Courses.

5 Consumption of Linked Data

We present in this section two scenarios where we used the linked data generated. We accessed the data through the semantic search feature of the wiki endpoint that provides querying functionalities with different output formats. The queries were formulated and passed through php to the wiki query endpoint, and results were returned in JSON for processing.

5.1 Interlinking Course Learning Material on Moodle

One challenge that web developers face when customizing existing platforms (e.g. moodle), is the need for understanding the existing code infrastructure to be able to extend the system's functionalities. Having linked data provides several advantages, including data portability and reuse. The ease of data extraction and consumption through endpoints such as SPARQL endpoints or the wiki's semantic search features can somehow break such development barriers. In our context we show that, when the data is decoupled from the application layer, it was possible to enrich moodle pages without having to modify the corresponding source code.

The linked data generated was used to enrich and connect the moodle course pages. Moodle is extensively used at the School of Business as a way to communicate course related information, and to interact around course deliverables. The suggested design of moodle course pages at the American University of Beirut is to subdivide the page using the covered topics in the course, and add topic related material, assignments and other activities under each section. Course instructors design their own page at the beginning of each semester, or reuse an existing one if the course was already taught by the instructor. One trend that is observed is that instructors of the same course tend to share interesting materials that could be used in classrooms. However such insights are not usually captured, and have to be re-shared whenever a new instructor teaches the course. In addition to sharing constraints, another bottleneck observed is that courses are designed (even on moodle) in isolation. However material relevant to some courses, can potentially be relevant to others (as perceived with the Big Data article discussed earlier).

The aim of this application is to break out of the static and isolated nature of content shared within a course topic on moodle. The application offers the functionality to dynamically enrich moodle pages, without leaving it, with material relevant to the topics of the course. Figure 3 highlights the steps performed within the moodle page.

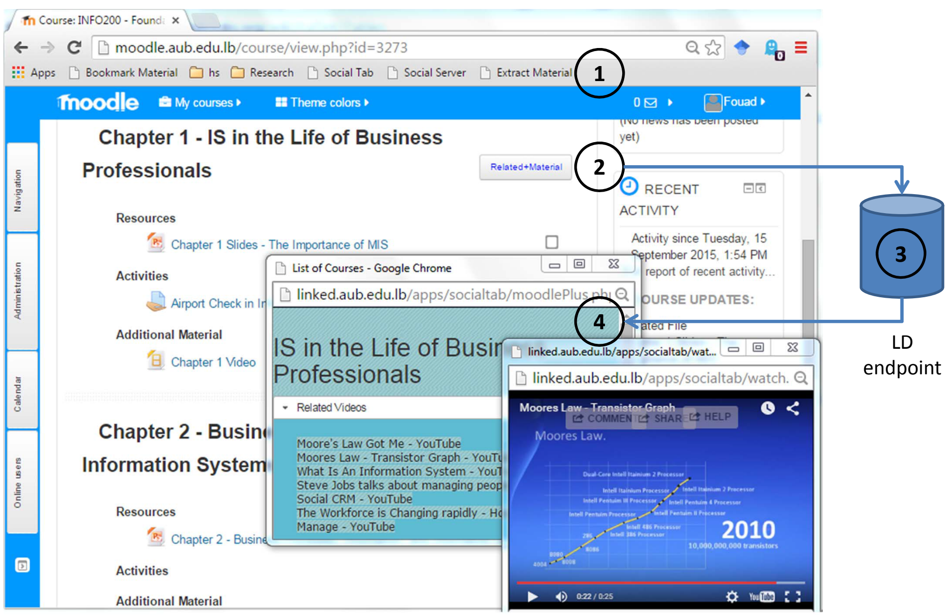


Fig. 3. Dynamically Enriching Moodle Course Pages.

A video tutorial is available as a guide for students and instructors to follow[14]. This application can be launched by pressing a bookmarklet in the browser (Part 1 in Figure 3), and the following sequence of steps is performed:

[14] http://linked.aub.edu.lb/docs/tutorial_extract_material

1. Scan the topics available in the moodle page: the executed code first will launch a javascript to scan the moodle page for the topics header. This code snippet extracts the section names by filtering the HTML class names that match the ones provided by the moodle page.
2. Enable the buttons on the moodle page: the javascript will dynamically inject form buttons next to each course topic on the moodle page (Part 2 in Figure 3), with the corresponding course code and topic embedded inside the button links.
3. Send query to the wiki linked data endpoint: when the user presses the button, a query is passed to the wiki endpoint (Part 3 in Figure 3) with the topic and course code. As per our model in Figure 1, the link between the course and material is done at two levels, through the *topic* and then through the *taught concepts*. Hence the query is built to first fetch the concepts covered in a topic, and then filter the learning material based on the concepts in focus.
4. Parse and visualize query results: the query results are returned in a JSON output, and parsed to identify the different types of related material (so far we have video material, articles, and books). Below is an example of a JSON output linking the topics to concepts:


```
"results": {
"Business Processes Information and Information Systems": {
      "printouts": {
          "Includes learning concept": [
           {
               "fulltext": "Business Process",
               "fullurl": "http:\\linked.aub.edu.lb\collab\
                   index.php\Business_Process"
           },
           {
               "fulltext": "Automated Process",
               "fullurl": "http:\\linked.aub.edu.lb\collab\
                  /index.php\Automated_Process"
           }...
```


The following shows the JSON output from the second query extracting the learning material and their concepts:


```
"results": {
"Amazing mind reader reveals his ‘gift’ - YouTube": {
      "printouts": {
          "Covers concept": [
           {
               "fulltext": "Social Media",
               "fullurl": "http:\\linked.aub.edu.lb\collab\
                   index.php\Social_Media"
           },
```

```
            {
                "fulltext": "Social CRM",
                "fullurl": "http:\\linked.aub.edu.lb\collab\
                    index.php\Social_CRM"
            },
            {
                "fulltext": "Privacy",
                "fullurl": "http:\\linked.aub.edu.lb\collab\
                    index.php\Privacy"
            }
        ]
    },
  "fulltext": "Amazing mind reader reveals his ‘gift’ - YouTube",
  "fullurl": "http:\\linked.aub.edu.lb\collab\index.php\
      /Amazing_mind_reader_reveals_his_%27gift%27_-_YouTube"
    }...
```

Finally the results are used to populate the page where users can read articles or play videos (Part 4 in Figure 3).

Following a concept centric design in our proposed model, coupled with the portability and ease of reuse of linked data, enabled us to implement a solution to dynamically enrich learning environments with relevant material. This design can automatically place learning material under the right topic on the course pages. For example going back to the Big Data article mentioned earlier in Figure 2, this article will dynamically appear in the information systems, management and operations management moodle pages.

5.2 Using Linked Data for Program Review

The second scenario in which we relied on the linked data generated from this work is in the program review process at the School of Business. Every four years, the curriculum has to be reviewed for changes, where courses are studied to be added or removed from the program. Another task that is part of the review process is the course sequencing. To achieve this purpose, traditionally each course is studied on its own, and compared to other courses, and to the learning objectives of the program. However it was not possible to perform an in-depth analysis beyond the syllabi content, and hence it was hard to know exactly what is covered in each course, and what are the concepts that are repeated across different courses.

The ability to have an overview of what is covered in a program and in which course is important for the program review activity. Hence course mapping was highlighted to be one of the major tasks required for adjusting course sequencing and coverage. Currently course sequencing is done at very high level based on the topics covered. However over the years, with some changes that occur in the content and course delivery, such sequencing should be revisited regularly. Our platform provided an unprecedented view of how courses overlap, down to the concept level.

We implemented a visualization showing how courses connect through topics and taught concepts. This visualization is dynamically generated based on the wiki content. A visualization example around the "Foundations of Informations Systems" course is presented in Figure 4.

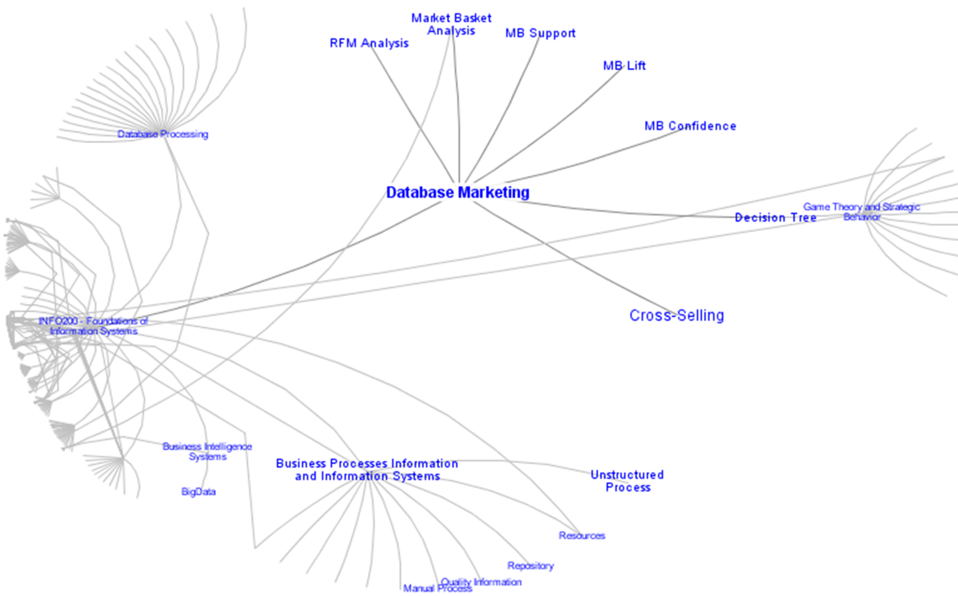


Fig. 4. Concept Map around the Foundations of Information Systems Course with a focus on the "Database Marketing" topic.

It highlights how the "Database Marketing" topic covers the "Decision Tree" concept, which is also repeated in the "Game Theory and Strategic Behavior" topic of the "Managerial Economics" course. This visualization can be accessed online at: `http://linked.aub.edu.lb/collab/index.php/Learning_Concepts_Graph`, and is created using the Semantic Graph extension[15], with the HyperGraph layout[16]. With the presence of semantic relations, this visualization was easy to implement. The extension requires the following to be specified: the wiki resource, the semantic relations to extract, the depth of relations and the layout dimensions. For example to generate this graph, we set the course page in focus, and used the "covers topic" relation to extract the related topics, and the concepts were selected based on the "includes learning concept" relation. Informal feedback from faculty members highlighted the complexity of this view

[15] `http://semanticgraph.sourceforge.net`
[16] `http://hypergraph.sourceforge.net`

in analyzing the concepts covered in a systematic way. A simpler representation was required to easily identify the repetition of concepts across courses.

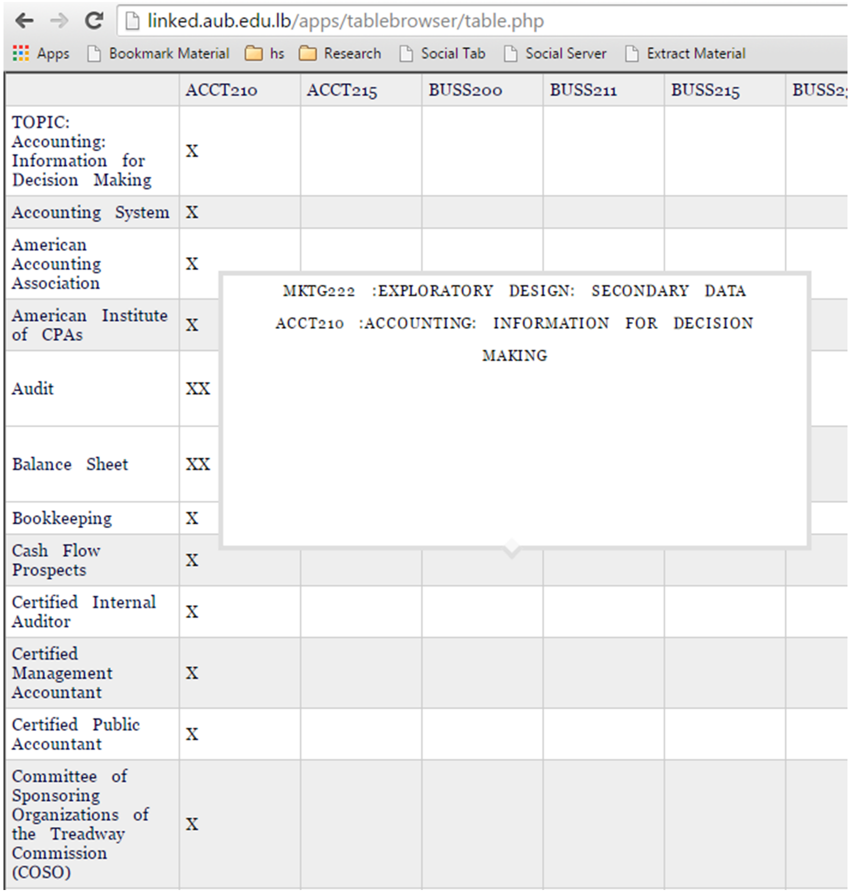
A tabular view was identified as a potential representation of courses, their topics, and corresponding concepts. We implemented a table on top of two queries. One fetches the course to topic relationships, and another extracts the topic to concept relations. We stored the results in JSON format, and loop through the entities to detect overlaps. We built a simple HTML-based table to render the results. The table can be accessed online[17]. We show part of the table in Figure 5. The users were able to see the list of courses (course codes are on the top row of the table), and in each row we present the list of topics and their corresponding concepts. The "X" marks the occurrence of this concept in the whole program. Due to the large size of the table, and when more than one "X" is in the cell, the user can roll the mouse over the table cell to see where this concept occurs in other courses. For example Figure 5 indicates that the concept *Audit* is mentioned in the Accounting 210, and in the Marketing 222 courses. We also fetch the topics where the concepts are covered.

This tool proved to be useful in the program review process, as it enabled the program coordinators and instructors to highlight the parts of the courses that require adjustment. This table implementation can benefit from various improvements. For example we anticipate the need for creating filters for users to specify a specific set of courses to display. For example if one of the departments at the School of Business is interested in how courses in their subjects overlap, a filter can be added based on the subject, and the table will only highlight courses that fall within this subject. Other filters could be based on learning objectives, to highlight how concepts feed into such objectives. In addition to improvements based on the data selection features, we see further improvements that can be made at the interface level. For example with the long list of courses and subjects, some aspects of the table can be made static such as the courses row on the top, or the topics and concepts column on the left of the application page. Another improvement can be made by enabling collapse and expansion features of concepts under the topics. Such features can be easily enhanced by using more sophisticated script-based modules.

6 Conclusion and Future Work

Higher eduction programs are represented in different ways. Traditionally representations were mainly performed through text documents including for example course syllabi and program catalogs. However such text layers do not allow us to exploit and dig deeper in the knowledge exchanged in higher education. As a result courses are mainly designed in isolation, with the aim to fulfill specific learning objectives. Furthermore, semantic connections between course entities such as learning material and topics are hard to systematically infer. We presented in this chapter our effort on connecting higher education program information at a conceptual level using linked data.

[17] `http://linked.aub.edu.lb/apps/tablebrowser/table.php`



	ACCT210	ACCT215	BUSS200	BUSS211	BUSS215	BUSS2
TOPIC: Accounting: Information for Decision Making	X					
Accounting System	X					
American Accounting Association	X					
American Institute of CPAs	X					
Audit	XX					
Balance Sheet	XX					
Bookkeeping	X					
Cash Flow Prospects	X					
Certified Internal Auditor	X					
Certified Management Accountant	X					
Certified Public Accountant	X					
Committee of Sponsoring Organizations of the Treadway Commission (COSO)	X					

Fig. 5. Tabular Representation of Concepts Covered in Courses with Overlap Detection.

The aim was to go beyond the information captured in course syllabi and catalogs. We required to go at a more granular conceptual level with explicit and machine processable semantics. For that we proposed a data model that captures and connects courses information, going down to the topical coverage and concepts taught. The concepts are then used as anchors between learning material and the higher education program. We relied on the existing syllabi structure to create the schema around courses information, and went deeper through the identification of concepts covered in the adopted material in the courses.

We built the linked data graph through a semantic Mediawiki implementation. The aim was to collaboratively expand this knowledge layer by involving various faculty members and students at the School of Business. The wiki offers a platform where instructors can reach a consensus around what is taught in their courses, by having a controlled environment to manage the reuse of existing knowledge and appropriate vocabularies for creating linkages. We created a semantic bookmarklet for learning material through which users can directly bookmark an interesting learning resource and be redirected to the wiki with existing elements automatically extracted from the page.

We also introduced how this interconnected data layer around the curriculum can help in the review and design of higher education curricula. The deep links among courses can be visualized in different formats. We presented two in this work, (1) a concept map that shows the connections around courses, topics and covered concepts, and (2) a table that highlights the occurrence of concepts in the program. This work offered program designers at the School of Business a unique view that was not possible before on how courses conceptually connect. They were able to see how concepts are repeated in courses, enabling them to make better decisions around required changes in the program.

Linked data provides several benefits in the management of curriculum and educational data, we discuss three advantages herein. The first advantage is related to the native graph-based structure of the data model. Such representation enables an easy extension of the model whenever required. This is one of the crucial requirements for keeping up with the dynamic nature of curriculum content. The second benefit lies at the level of portability and ease of data reuse in different contexts as highlighted in our scenarios. When semantics are explicitly encoded in the data, building applications on top of it is made easy. For example we were able to selectively extract parts of the knowledge around courses to visualize course contents in the form of tables or diagrams. We also showed how we can bring learning material in the context of a course, and place them under its corresponding topic within existing online learning environments such as moodle. The use of taught concepts as anchors between the materials and courses gave us a great flexibility in fetching material that stretch the boundaries of courses delivery, which can improve the learning experience of students and highlight how courses cross-connect in the degree program they are enrolled in. The third advantage of the use of linked data is driven by the growing number of the available tools that are supporting the storage, maintenance and extraction of linked data. This trend will engage data publishers in adopting linked data principles as a way to serve their data. Hence in our context, we know that the availability of our conceptual layer will have a longer life span. We also believe that the value of this data and the sophistication of applications built will increase, when the data is combined with external linked data sources.

There are different research directions to follow next. One natural continuation of this work is to evaluate the impact of this linked representation of courses on curriculum changes and reviews. We are currently developing further visualizations, and planning to evaluate the degree of insights that can be generated from such diagrams. We are also interested in evaluating the impact of having such data layer in learning environments through a guided user study coupled with evaluation measures. Another line of work we are currently pursuing is on capturing social interactions around our education program. By merging the social and education graphs, we anticipate that we will be able to granularly analyze how teachers and students interact around concepts delivered during their higher education journeys. This new linked data layer will offer endless opportunities in manipulating curricula related information in new and insightful contexts.

7 Acknowledgments

This work was supported by the University Research Board grant at the American University of Beirut. I would like to thank my colleagues and students who contributed to the extension of the graph through the wiki, and supported me in the implementation of the tools. I would also like to thank the students who have used the bookmarklet to bookmark the interesting materials that triggered invaluable insights and observations.

References

1. T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific american*, 284(5):28–37, 2001.
2. C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. DBpedia-A crystallization point for the Web of Data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3):154–165, 2009.
3. M. d'Aquin. Linked Data for Open and Distance Learning. Commonwealth of Learning report, 2012.
4. T. Heath and C. Bizer. *Linked Data: Evolving the Web into a Global Data Space*. 2011.
5. T. Heath, R. Singer, N. Shabir, C. Clarke, and J. Leavesley. Assembling and applying an education graph based on learning resources in universities. In *Linked Learning (LILE) Workshop*, 2012.
6. G. Klyne and J. J. Carroll. Resource description framework (RDF): Concepts and abstract syntax. Technical report, 2006.
7. D. Kroenke, D. Bunker, and D. N. Wilson. *Experiencing mis*. Prentice Hall, 2010.
8. M. Krotzsch, D. Vrandecic, and M. Volkel. Semantic mediawiki. In *The Semantic Web-ISWC 2006*, pages 935–942. Springer, 2006.
9. M. Schmachtenberg, C. Bizer, A. Jentzsch, and R. Cyganiak. Linking Open Data Cloud Diagram (http://lod-cloud.net/), 2014.
10. Stefan Dietze, Salvador Sanchez-Alonso, Hannes Ebner, Hong Qing Yu, Daniela Giordano, Ivana Marenzi, and Bernardo Pereira Nunes. Interlinking educational resources and the web of data. *Program*, 47(1):60–91, Feb. 2013.
11. F. Zablith. Interconnecting and Enriching Higher Education Programs using Linked Data. In *Proceedings of the 24th International Conference on World Wide Web Companion*, pages 711–716. International World Wide Web Conferences Steering Committee, 2015.
12. F. Zablith, M. d'Aquin, S. Brown, and L. Green-Hughes. Consuming linked data within a large educational organization. In *Proceedings of the Second International Workshop on Consuming Linked Data (COLD) at International Semantic Web Conference (ISWC)*, 2011.
13. F. Zablith, M. Fernandez, and M. Rowe. Production and consumption of university Linked Data. *Interactive Learning Environments*, 23(1):55–78, 2015.