LINKED OPEN DATA
in
Administration
at the
Vienna University of Technology

Stefan Gamerith[a], *Author*
Kevin Haller[b], *Co-Author*
Lukas Baronyai[c], *Co-Author*
{e0925081[a],kevin.haller[b],e1326526[c]}@student.tuwien.ac.at

Abstract

Although many universities have already published their public resources according to the Linked Data principles there have not been any efforts at the Vienna University of Technology.

This paper shows international success stories and best practices regarding Linked Open Data principles in an educational environment. To investigate major benefits and challenges concerning application development as well as the adoption of Linked Data principles and possible use cases for end-user applications, we interviewed two administrative workers at the Vienna University of Technology. Furthermore, to support practitioners and developers a prototypical implementation of a publication framework for educational resources as well as a general architecture of an application implementing the Linked Data principles is given.

Using Linked Open Data is a promising way of enhancing and harmonizing processes at universities, thus creating application scenarios with a large and supportive target audience is recommended. To limit integration costs, using existing, therefore maintained and open data sets is preferred over creating new ones.

Contents

1 Introduction

While the pressure on governments and public organizations to release *Open Data (OD)* has significantly grown with the spread of information systems there has been also a need for *linking* these data from various sources to understand the information in a contextual sense.

As OD includes non-confidential data any restrictions in distribution are prohibited and information is funded only by public money [27]. The application domain for OD providers is not restricted by its nature in any way including from traffic, weather and public transport to name a few. However, exclusively exposing public assets is not enough, in addition establishing a feedback loop facilitates an ongoing adaption to the stakeholders concerns.

The World Wide Web has proven great success in spreading knowledge of various data sources all over the world. The building blocks of the Web are documents and links building a shared, global and connected information source. This can be seen as the key success factor in its nearly unconstrained growth [26]. *Linked Data (LD)* has adopted these principles of publishing and connecting data realized as machine-readable, structured data connected to various data sets which in turn are linked to different data sets.

Berners-Lee [4] developed a five star deployment scheme classifying OD. The scheme ranges from a one star rating covering proprietary data formats (e.g. Portable Document Format (PDF)) to a five rating including machine-readable formats using open standards with links to other data sets.

Although LOD offers universities new opportunities for providing unprecedented insight into its core activities and ease application development, a major **problem** is that **LOD has not been widely adopted by universities yet**. Even tough there are a few examples [10] of publishing university related data sets as LOD there has been little knowledge of its usage for publishing university related information.

To answer the research questions stated in the next sub-section we conducted interviews at the Vienna University of Technology. In particular **the context of this work is Administration**, though there exist papers investigating similar research questions but with different context [2, 17].

The remainder of this section states the addressed research questions, describes the contributions of this work and gives an overview of the structure of this paper.

1.1 Research Questions

The fundamental research question we investigate is:

> *How can Linked Open Data help to improve processes in a university context and how can it be successfully applied?*

More concrete, this paper concentrates on the following four research questions:

RQ1: What are best practices regarding the applicability of Linked Open Data in university settings? At the time of writing this paper there are no established best practices for the use of LOD due to its little adoption in university settings. For this very reason it is crucial to identify strengths and limitations from previous experiences of using LOD as the foundation of information systems [10].

RQ2: What are major benefits and barriers for each stakeholder and what are useful use cases? We identified three different stakeholders *Students*, *Researchers* and *Administration staff*. However, **focus of this work are administrative workers**. Since the success of any new technology highly depends on their acceptance concerns of each target group needs examination. Furthermore, use cases are important to showcase profits and shortcomings.

RQ3: What are major challenges for the implementation of a Linked Open Data solution?
As the implementation of a LOD solution is a time consuming task the knowledge of probable challenges from a technical perspective as well as from a management perspective is key to a successful adoption.

RQ4: How would a prototypical implementation of a publication framework based on Linked Open Data look like? Among the definition of building blocks for a publication framework of university related assets a high level picture of the overall architecture needs to be drawn to give implementers and LOD experts a common understanding of the system. Next, defining sample ontologies representing selected assets of the university domain draws concrete examples of how LOD data might look like.

1.2 Methodology

Finding an answer for the research questions above has lead to the following three methodologies:

A coordinated set of semi-structured interviews To answer research questions (RQ) two to four, we interviewed a selected set of *Administration staff*. Semi-structured interviews were selected as the means for data collection because they are well suited for exploring the impressions and interests of the interviewees like in a discussion while still following a defined structure.

Literature Review Undertaking a literature review to justify scientific contributions and making sound conclusions is an established practice in any scientific community. Since our scientific work targets to the Semantic Web community we made some pre-assumptions about a basic understanding of the technologies and concepts used in that respective area. More specifically, we assume a basic understanding of the concept of an ontology and some pre-knowledge about ontology description languages.

Conceptual System Design The development of applications based on LOD requires a methodology facilitating a common understanding of the overall system infrastructure. Therefore we designed a conceptual model of a prototypical implementation of a publication framework based on LOD.

1.3 Contributions

The work in this paper mainly contributes to different aspects which need to be considered when designing and implementing an application driven by LOD. More precisely, our contributions can be categorized into the following four areas:

1. Identifying best practices for Linked Open Data in university settings. Information systems needs to cope with vast amount of data nowadays growing the need to efficiently handle Linked Data as well. We gave a brief overview of existing research work in that area, in particular, we compared the benefits and shortcomings in existing LOD solutions.

2. Finding benefits/barriers including additional use cases for stakeholders. As with every software project the very first phase of the Software Development Lifecycle (SDLC) is *evaluating of requirements*. As LOD driven software has additional requirements to the accessibility of data and their organization we investigated if the overhead compared to an established technology (e.g. a database based solution) is feasible. A set of selected *administrative workers* are interviewed at the Vienna University of Technology. Additionally we proposed several use cases emphasizing their point of view.

3. Discovering possible obstacles for implementers of a Linked Open Data solution. As the application domain for a Linked Data is limited to the university context our work includes LOD driven applications resulting from our conducted interviews.

4. Sketching a prototypical implementation of a publication framework driven by Linked Open Data. The proposed publication framework covers the whole process of data provision, requirement analysis and application development designed for but not limited to university related assets.

1.4   Structure of this Paper

This paper is structured as follows:
Section 2 provides a summary of existing efforts regarding Linked Data principles in educational contexts. Section 3 discusses the used methodology and results concerning RQ2 and RQ3 from selected administrative workers. While section 4.1 shows a prototypical implementation for a publication framework of educational assets, section 4.2 proposes a technical architecture for the application of Linked Open Data principles in educational environments. Finally, conclusions and future investigations are made in sections 5.1 and 5.2.
Whereas section 4.2 is particularly intended for readers familiar with software architectures, the remaining sections do not necessarily require a deeper understanding. However, a general understanding of Semantic Web technologies for the whole paper is recommended.

2   Related Work

This section gives a brief overview of existing work relating to Linked Open Data (LOD) in educational contexts. In particular, we investigate experiences and established practices in that area. Fortunately there has been plenty of work done especially around *Linked Universities* (section 2.1), an alliance of european universities. Next, selected participants of that project are briefly discussed who had gained great experiences at publishing educational resources. Although beneficial, but many contributions lead to divergent data sets and concepts. The project *LinkedUp* (section 2.2) tries to combine several of those, facilitating a uniform view of these concepts.

2.1   Linked Universities

With the ongoing trend of providing unrestricted access to educational resources several universities participated in *Linked Universities*[1], an alliance of european universities aiming at exposing their data as Linked Data. Although some universities had already published their educational resources, data formats often diverge, raising third party integration costs. For this very reason *Linked Universities* facilitates data aggregation and integration among universities. Cross-university collaboration and sharing of common resources helps in establishing a common goal, which is the creation of a global Web of university data.
More specifically, the goals and visions are:

- Definition of common *vocabularies* used for defining university related resources

- Sharing Linked Data related *recipes* and *tools*

- Collecting *experiences* from various university projects

The most prominent members among their main contributors are:

- `Mathieu d'Aquin` - The Open University, UK

- `Carsten Kessler` - University of Mnster, Germany

- `Tomi Kauppinen` - Aalto University, Finland

- `Stefan Dietze` - The Open University, UK

- `Christopher Gutteridge` - University of Southampton

- `Hannes Ebner` - Royal Institute of Technology (KTH) / MetaSolutions AB

- `Charalampos Bratsas` - Aristotle University of Thessaloniki, Greece

- `Oguz Dikenelli` - Ege University, Turkey

- `Jakub Klmek` - Charles University in Prague

- `Jorge Pantoja` - Universitat Pompeu Fabra

---

[1]http://linkeduniversities.org/

Shared vocabularies and data sets   The list below shows common vocabularies[2] used for describing educational and university related resources together with the involved concepts and recommendations for usage.

- `AIISO - an Academic Institution Internal Structure Ontology`
  The Academic Institution Internal Structure Ontology (AIISO)[3] offers a schema to describe the organizational structure of a university. Its intended usage is together with FOAF describing human relations, Participation Schema [4] describing the roles that people play within groups, and AIISO-Roles[5] describing common roles found in academic institutions and higher education.

- `BIBO - The Bibliographic Ontology`
  The original idea behind the adoption of the Bibliographic Ontology (BIBO)[6] was to express citations and bibliographic relations using a general concept like an ontology. Its usage is not limited to bibliography, but any kind of document describable by RDF. The ontology was originally created by Fredercick Giasson and Bruce D'Arcus, though its specification is in constant change and evolving.

- `LSC - Linked Science Core Vocabulary`
  The Linked Science Core Ontology (LSC)[7] is targeted at relating objects in science to time, space and themes. In particular its design facilitates the definition of scientific assets including elements of research, their context and inter-relations. However, LSC is a collection of basic concepts in scientific publishing, hence the word *core* in its name.

- `TEACH - Teaching Core Vocabulary`
  The Research Core Vocabulary[8] is a simple ontology describing educational courses and staff. The main entry point is represented by the namespace prefix *http://linkedscience.org/teach/ns*. Mixing with other ontologies extends its applicability, though basic concepts of courses or seminars are covered. Some these related ontologies are `MLO` - an ontology for learning opportunities, `FOAF` - a friend-of-a-friend ontology, the `Time Ontology`[9] and `AIISO` - an Academic Institution Internal Structure Ontology.

Listing 1 shows a real world example[10] of the ontologies listed above. It shows a course taught at the Aalto University[11] organized in autumn 2014.

Listing 1: RDF/XML representation of an educational course for industrial symbiosis and industrial ecology concepts taught at the Aalto University using ontologies proposed at Linked Universities

```
<aiiso:Course rdf:about="http://data.aalto.fi/id/courses/noppa/course_Puu-0.3310">
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_48700"/>
  <teach:academicTerm>I (NB! The course will be organized for the last time in autumn 2014.)</teach:
  ↪ academicTerm>
  <teach:hasAssignment rdf:resource="http://data.aalto.fi/id/courses/noppa/assignment_Puu-0.3310_0"/>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_48461"/>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_46879"/>
  <foaf:homepage rdf:resource="https://noppa.aalto.fi/noppa/kurssi/Puu-0.3310"/>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_47564"/>
  <teach:hasAssignment rdf:resource="http://data.aalto.fi/id/courses/noppa/assignment_Puu-0.3310_1"/>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_45862"/>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_46880"/>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_45861"/>
  <teach:grading></teach:grading>
  <dc:language>en</dc:language>
  <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_47020"/>
```

---

[2] http://linkeduniversities.org/lu/index.php/vocabularies/index.html
[3] http://vocab.org/aiiso/
[4] http://vocab.org/participation/schema
[5] http://vocab.org/aiiso-roles/schema
[6] http://bibliontology.com/
[7] http://linkedscience.org/lsc/ns-20111129/
[8] http://linkedscience.org/teach/ns-20130425/
[9] https://www.w3.org/TR/2006/WD-owl-time-20060927/
[10] http://data.aalto.fi/page/id/courses/noppa/course_Puu-0.3310
[11] http://www.aalto.fi/

```
    <teach:courseTitle xml:lang="en">Industrial Symbiosis</teach:courseTitle>
    <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_46680"/>
    <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_46219"/>
    <aiiso:code>Puu-0.3310</aiiso:code>
    <teach:courseDescription xml:lang="en">Introduction to industrial symbiosis ...</teach:
    ↪ courseDescription>
    <teach:arrangedAt rdf:resource="http://data.aalto.fi/id/courses/noppa/lecture_88150"/>
    <teach:ects>5</teach:ects>
</aiiso:Course>
```

Data conversion and transformation tools   There has been put quite a lot of effort into collecting tools and software platforms commonly used to expose Linked Data at Linked Universities. These software provide generic tools for data conversion and transformation. RDF has been chosen as the main target format, allowing further processing by general purpose XML manipulation software (e.g. XPath evaluators and parsers).

The list below shows conversation and transformation tools inspired by the LUCERO Project[12], headed by Mathieu d'Aquin at the *Open University* [14].

- `Triplify`
  Still, a large amount of persistent data is stored in a relational database. Together with the SQL language these technologies dominated the way how data is structured and organized. A strict database schema makes data transformation difficult, however Triplify[13] tackles these problems by converting relational data into *subjects*, *predicates* and *objects*, hence its name.

- `GRDDL`
  Traditionally Extensible Stylesheet Language Transformations (XSLT) are used for altering or converting XML documents. As XML and RDF share a common base (e.g. RDF documents use XML based syntax) XSLT seems to be the first choice, however as RDF descriptions become complex transformation, is often infeasible. The *Gleaning Resource Descriptions from Dialects of Languages* (GRDDL)[9] serves exactly that purpose by making it possible to declare that XML documents include data compatible with RDF.

- `Open Refine`
  The data formats of choice for tabular organized data are CSV and Excel files. Transformation of such formats targeted at facilitating sharing between people can become quite challenging, especially for proprietary data formats. Open Refine[14] (formerly Google Refine) is a tool which eases data conversion and exploration from many different sources including Excel, CSV and Google Spreadsheet. Although it was not originally intended to support RDF export there has been created a tool[15] which adds RDF export as well as including some tools to connect tabular content to external linked datasets.

Data sets and Endpoints   All universities participating in Linked Universities provide endpoints for all different kinds of Linked Data.

The Linked Universities project provides a list of participating universities together with endpoints offering Linked Open Data. Unfortunately up to the time of writing this paper not all endpoints listed at the member page[16] are online.

The table below shows selected universities together with online SPARQL endpoints and data catalogs:

---

[12]http://lucero-project.info/

[13]http://triplify.org/

[14]http://openrefine.org/

[15]http://refine.deri.ie/

[16]http://linkeduniversities.org/lu/index.php/datasets-and-endpoints/index.html

University	Data Set	SPARQL Endpoint
Open University	http://data.open.ac.uk	http://data.open.ac.uk/query
University of Southampton	http://data.southampton.ac.uk	http://sparql.data.southampton.ac.uk
Aalto University	http://data.aalto.fi	http://data.aalto.fi/endpoint
University of Muenster	http://data.uni-muenster.de	http://data.uni-muenster.de/php/sparql
University of Pompeu Fabra	http://data.upf.edu/en/linked_data	http://data.upf.edu/en/sparql

Table 1: SPARQL endpoints and data catalogs of selected members of Linked Universities

While the University of Pompeu Fabra provides access to just academic publications, others offer additional data sets covering academic courses, publications, research projects, places, buildings, researchers, staff, organizational structure, news and events. The most complete data is provided by the Open University, a university located in the UK aimed at offering students a unique learning experience with the support of information systems. In addition to the data sets described above the Open University provides audio/video recording of academic courses and social media related data sets as channels, playlists and posts from http://audioboo.fm and playlists as well as videos from https://www.youtube.com.

2.2 LinkedUp project - Linking Web Data for Education

The LinkedUp project[17] is a project funded by the European Commission aimed at pushing forward the exploitation and adoption of Linked Open Data in educational organizations and institutions. This project started on the 1st November 2012 and lasted about 2 years.

More specifically, the projects main objectives are:

- `Open Web Data Success Stories`
  The identification of innovative success stories using tools and data sets in the education sector helps at spreading knowledge and awareness of Linked Data principles.

- `Web Data Curation`
  The collection of relevant educational data sets facilitate the integration of third party data and applications.

- `Evaluation Framework for Open Web Data Applications`
  Evaluation of successful large scale applications driven by Linked Open Data promotes the adoption of Linked Open Data especially in educational contexts.

- `Technology Transfer in the Education Sector`
  The promotion of Linked Open Data technologies is an explicit goal of the LinkedUp project.

The outcome of the LinkedUp project were several so-called *"deliverables"*[18].

Below, there is a list of the most relevant deliverables:

State of the art and data assessment
In this report [21] a comprehensive study on the developments in the field of Linked Data and related fields of educational data mining and learning analytics is given. The report resulted from a collaboration between the Leibniz University of Hannover, the Open Knowledge Foundation, the Open Universiteit Nederland, the Open University, Elsevier and Exact Learning Solutions. The authors first briefly describe fundamental technologies and concepts, namely Linked Data, educational data mining and learning analytics used throughout the paper. Several candidates representing each of the two basic data mining principles (*data harvesting* and *distributed search*) for heterogeneous data sets are presented. The field of learning analytics focuses on techniques targeted at analyzing and understanding educational learning processes and environments. A complete discussion on this topic though, was out of scope since this is a relatively new research area. Next, challenges and barriers concerning the heterogeneity of open educational resources are investigated. Techniques trying to solve isolated data sets including *schema mapping* and *classification and clustering* are proposed. Finally, data sets from the LinkedUp repository[19] and possible legal obstacles and solutions to these are discussed.

---

[17]http://linkedup-project.eu/
[18]http://linkedup-project.eu/resources/deliverables/
[19]http://data.linkededucation.org/linkedup/catalog/browse/

The LinkedUp Challenge(s) The LinkedUp Challenge[20] was organized as three separate challenges *Veni*, *Vidi* and *Vici* ending on October 2013, April 2014 and November 2014 respectively. While the goal of the first competition (Veni) was developing a *"prototype or demo that uses linked and/or open data for educational purposes"*, the Vidi challenge was targeted at finding *"innovative and robust prototypes and demos for tools, which still may contain bugs, as long as it has a stable set of features and there is some proof that us can be deployed on a realistic scale"*. The rationale of the latter, yet more advanced challenge was to build *"advanced prototypes and tools that should be mature and stable; it should be used or have been used by a fair amount of users on a realistic scale"* [20]. The combined goal though was to promote creativity and innovation in ways to mash-up and link educational resources and services. In addition, companies, universities and government agencies were encouraged to share and (cross-)link to educational and non-educational assets.

The winner for Veni was `Polimedia`[21], an application facilitating large-scale, cross-media analysis of the coverage of political events. `TuvaLabs`[22] won the first place of the Vidi competition aimed at transforming open data into opportunities for meaningful teaching and learning, equipping teachers and students with high quality, consolidated data sets. Finally, `Flax` (Flexible Language Acquisition)[23], the winner of the Vici competition, was designed to automate the production and delivery of interactive digital language collections and targeted to non-expert users (e.g. language teachers, language learners, subject specialists, instructional design and e-learning support teams).

Evaluation Framework Drachsler et al. [13] proposed an evaluation framework throughout the three open educational data competitions Veni, Vidi and Vici, described in detail in the paragraph above. The aim of the evaluation framework was to evaluate the usefulness, usability, acceptance and appropriateness of the contributions to the LinkedUp challenges. Therefore experts had to define and/or refine assessment criteria and indicators for measuring the quality of the framework. Practically the frameworks usefulness and ease of use was tested through a questionnaire and interviews. A detailed questionnaire for each of the three challenges as well as the final version of the evaluation framework are accessible at [13].

2.3 Summary

In this section a brief outline of the relevant research fields for this paper was presented.

In the beginning of this section the project *Linked Universities*, an alliance of several universities, was introduced. The purpose of that collaboration was spreading the knowledge and adoption of Linked Data principles across universities borders. We quickly discussed shared vocabularies and data sets, introduced useful data conversion and transformation tools and finally listed selected SPARQL endpoints and data catalogs.

In the remainder of this section an interesting project called *LinkedUp* and supported by the European Commission was briefly introduced. While being in some sense related to Linked Universities (e.g. sharing similar goals), the LinkedUp project also covers practitioners and implementers viewpoints. That was specifically addressed by the LinkedUp challenge(s), three competitions intended at pushing the exploitation and adoption of Linked Open Data in educational organizations and institutions from a researchers as well as practitioners point of view.

3 Benefits and Challenges

In this chapter we figure out the driving factors of publishing university related assets as Linked Open Data and introduce sample applications built on top of it. First, we describe the used methodology (section 3.1) to show benefits and challenges using LOD at the Vienna University of Technology. Second, we discuss concrete topics of the interview, hence, potential needs, benefits and challenges are examined. We interviewed **2 representatives** of **administration** from the Vienna University of Technology. Next, we compare benefits against shortcomings by giving concrete use-cases to show

---

[20]http://linkedup-challenge.org/

[21]http://www.polimedia.nl/

[22]https://tuvalabs.com/

[23]http://flax.nzdl.org/

under which circumstances LOD solutions are preferable to traditional ones (e.g. a database based implementation) (section 3.2).

3.1 Methodology

In this sub-section we describe the Vienna University of Technology and interviewed people more thoroughly, elaborate on the interview design, analysis and threats to validity.

3.1.1 Interview context

All interviews were taken at the Vienna University of Technology [40]. The university was founded in 1815 as Imperial and Royal Polytechnical Institute. 160 years later (1975) the university was renamed to its current name *University of Technology*. At the time of writing the majority of the faculties are located at the fourth district in Vienna with nearly 9,000 rooms spreading over 276,000 square meters. The Vienna University of Technology provides a website of all organizational units falling into the category Administration & Services [25]. For the evaluation of this work several departments were contacted by email. However, the majority of them declined participating because either they did not feel comfortable answering questions because of the technical concentration of the interview (e.g. Semantic Web in general and Linked Data in particular) or they were too busy in the restricted time frame (4 months) the interviews were planned. However, we were forwarded to representatives of the organizational units *Research and Transfer Support* [34] and *Department for Studies and Examinations* [12].

3.1.2 Interview design

As a solid preparation phase is the foundation of high quality results a considerable amount of time should be invested. Conceptually there exist three different types of interviews: *Unstructured interviewing*, *Semi-structured interviewing* and *Fully-structured interviewing* [3]. Due to the nature of unstructured interviews they are comparable to discussions, not following specific guidelines. In contrast, data collection in semi-structured and fully-structured interviews is typically documented by questionnaires. Harris and Brown [18] describe the difference between these two interview types as, while in fully-structured interviews the participant responds to predefined answers (e.g. choosing from options) semi-structured interviews also focus more on testing interviewees responses motivating them to provide more background information and clarification. Due to the small number of interviews this paper primarily focus on understanding interviewees needs [32].
The questionnaire is composed of three major parts:

1. General questions targeted at getting a better understanding of the participants. It includes exploring prior knowledge about LOD in particular and information systems in general.

2. Analyzing stakeholders thoughts and opinions about presented example applications built using LOD motivating them to provide use cases for new applications.

3. Exploring existing and additional sources for data provision.

The list represents in which order the parts of the questionnaire were addressed during the interview, though there is no need for keeping this order in semi-structured interviewing. A more detailed and complete image of the questionnaire can be taken from the appendix.
In the next few lines the interview guideline is described briefly. First, the research team was introduced to the interviewee as well as the research objective. Afterwards, the moderator asked the interviewee about their background. Due to the level of expertise in the field of LOD the moderator was able to decide whether an introduction to LOD is necessary or not. Then the last two parts were covered discussing open questions concerning specific research goals.
The interviews were conducted in teams of two researchers. One researcher in the role of the moderator had the task to collect data by writing valuable contributions onto a logging sheet. This way, the interviewer can fully concentrate on listening and asking questions. An alternative approach is recording the whole conversation but in this case a single researcher has to spend more time (in worst

case two times more) to extract important contributions. However, no other person would be required and recording can be used for later analysis. Since we had enough human resources we decided for the former approach.

3.1.3 Data Analysis and Validity

Semi-structured interviews mix open and closed questions allowing qualitative and quantitative analyzing methods for examining the interview outcome [35]. Quantitative data analysis is measurable and objective typically generating numerical values. Qualitative data analysis is descriptive and subjective focusing on deriving results and drawing conclusions from collected data [41]. Due to the exploratory nature of the interviews a detailed description of qualitative and quantitative analysis methods were omitted here.
A common threat to the validity of collected data is a biased view on the interview answers. Two countermeasures were chosen to reduce this risk. First, the interviews were conducted by multiple researchers. Second, the interview was recorded for later analysis.

3.1.4 Interviewee Background

At the beginning of the interviews participants answered some questions regarding their background including daily work tasks and responsibilities at the university. Furthermore, interviewees estimated their level of expertise in the areas Information & Communication Technologies and LOD.
The concrete results of expertise are twofold. On the one hand one participant has excellent expertise with information systems and some experience with Linked Data. Expressed in numerical form of a scale from 1 representing little expertise to 5 representing expert knowledge the interviewee estimated his competence with information system systems 5 and Linked Data 3. As a side note, he commented to the latter estimation that he had already worked with LOD applications but never developed one on his own. He justifies his estimation with his background in computer science.
The other participant has some experience with information systems but little with Linked Data. More formally, she estimated her competence with information system systems 3 and Linked Data 2, meaning that she classified herself an ordinary office worker with average skills in information systems and eventually heard but never used applications driven by Linked Data.

3.2 Results

After giving a detailed description of the used methodology in the previous sub-section this sub-section contains interview outcome. Due to the reasons given in section 3.1.1 and the background of a participant this section contains rather general, applicable not only to administration but research and student affairs as well. Concrete, one participant is responsible for administration and research. Even tough more participant were contacted, interviews were refused because of limited time.
First, potential use cases for building applications and publishing data as Linked Open Data are identified including benefits and barriers. Second, benefits and challenges are compared with one another resulting in guidelines for the use of LOD.

3.2.1 Building Map

At the time of writing this paper there exist a website[24] for viewing lecture rooms, offices and other geographical related data. These website includes content of all Austrian universities hiding detailed information as elevators, stairways and other points of interest as buildings near-by, pubs and coffee shops. This data is particularly relevant for external visitors who are new to the campus. Furthermore, maps are published as a PDF-file or image making it impossible to include any real time data such as the utilization of computer rooms and defect elevators.
Figure 1 shows a map application built for the University of Economy in Vienna.
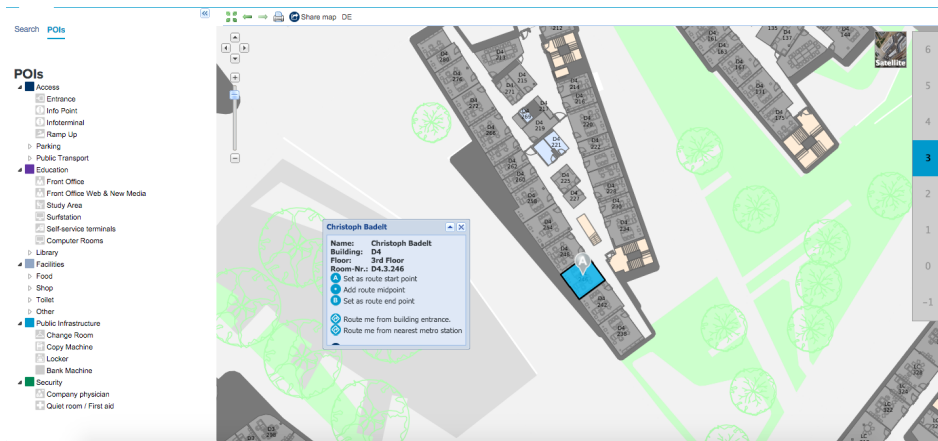
---

[24]http://www.wegweiser.ac.at/

Figure 1: A map application of the Vienna University of Economy

Although there is no indication if the map uses Linked Data the user interface shows everything of what is achievable with LOD. In figure 1 a search for the rector of the Vienna University of Economy has been initiated which results in highlighting his office and selecting the appropriate floor on the left side. In addition, a guided route is initiated by clicking an adequate link on the turquoise overlay box.

Benefits   All interviewed stakeholders responded positive to the idea of a building application as described above. In particular they liked the ability of highlighting specific rooms and navigating through different floors. In general, there are advantages for disabled people and external visitors as outlined in the beginning of this sub-section.

Challenges   Digital data is the foundation of any information system requiring approaches for converting non-machine-readable data to machine-readable formats. Unfortunately the Department of Construction and Technology[25] at the Vienna University of Technology hosts exclusively building plans created by Computer Aided Drawing (CAD) software prohibiting the addition of metadata for further processing.

Special attention needs data ownership and copyright issues. Despite the fact that the Austrian government passed a law[26] regulating how and which governmental data could be published there are still legal factors which need to be clarified, notably if data ownership is exclusively hold by public institutions.

3.2.2   Tool for statistical evaluation of student related data

Unistats[27] has published statistical course data of universities and colleges across United Kingdom. These data is exposed under the Open Government License[28] allowing copying, publishing, distributing and transmitting data without any restrictions. Besides offering data about undergraduate courses, post-graduate courses, students satisfaction scores, post-graduate salaries and other similar information a course assistant supports prospective students in choosing the appropriate study.

Figure 2 shows search results originating from a sample search query which selects all part time, computer science courses in England ending in a Bachelor degree with a sandwich year.

---

[25]http://www.gut.tuwien.ac.at/

[26]Informationsweiterverwendungsgesetz (IWG)

[27]https://unistats.direct.gov.uk/

[28]http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/
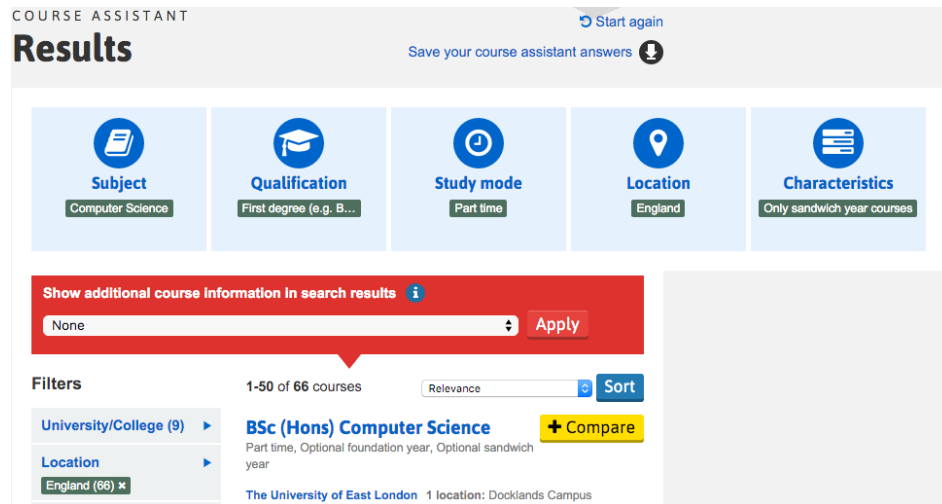
Figure 2: Course assistant application of Unistats UK

Benefits All interviewees found the course assistant useful. The application is particularly attractive for future students who need help in selecting certain courses for participation. However, participants had concerns about major challenges as described below.

Challenges The primary data source is evaluation data calculated from online feedback data. Unfortunately there is often not enough data available to draw solid conclusions. However, this is a rather general problem of feedback related data analysis.

Despite the fact that the Vienna University of Technology do offer mechanisms for rating courses feedback related data is not accessible from outside of the universities backbone net limiting the options of third-party data analysis.

3.2.3 Researcher Portfolio

The Vienna University of Technology provides researchers a platform[29] for presenting their research portfolio. The intention is give external stakeholders an insight into a researchers portfolio. However, searching for specific topics or cooperations from a non-technical perspective is not feasible. In fact, the searching algorithm selects candidates by exactly matching one of the given search parameters with some predefined tags. These tags are words provided by the researcher which are in most cases domain specific and only known by a limited research community.

Hoogerwerf et al. [23] provide showcases of a successful application of a publication databased driven by LOD. Even tough a publication database built using Linked Data has advantages over ones powered by traditional technologies there are some challenges with need to be taken into account.

Benefits Scientific research often involves cooperation with third parties. A Linked Data solution easily brings researchers and partners together as similar competencies are merged being understandable by both.

Linked Data relies on open standards facilitating the creation of third party applications (e.g. statistical evaluation software). Participants responded that extensive data visualization eases the identification of cooperations between research and external stakeholders.

Challenges Keeping data in sync is almost in any case challenging. In particular if there is no way of a (semi-) automatic data transformation process. Even though a central database would remove the burden of duplicate data sets on different places the problem of manual data feeding still remains.

---

[29]TISS Forschungsprofil https://tiss.tuwien.ac.at/fpl/

Legal privacy regulations are different from country to country making the maintenance of a global publication database difficult. Privacy regulations often prohibit data distribution even in restricted areas.

3.2.4 Summary

In this section we described the conducted interviews as the primary methodology as well as their results. We used a technique called semi-structured interviewing for data collection. Furthermore, we gave arguments for data validity and discussed different data analyzation methods. Since our aim is to better understand the stakeholders needs and challenges the majority of questions are of open nature qualifying for qualitative data analysis.

In the remainder of this section we introduced use cases for Linked Data applications and identified benefits and challenges in their design and implementation. In addition to the shown use cases we listed suggested additional use cases. Among the individual benefits for each use case we found challenges common to all proposed applications. Privacy and legal issues are major challenges not only to Linked Data but information systems in general. This is particularly true for Open Data. Thus, thoroughly clarifying data ownership and possible legal obstacles are key to a successful LOD solution.

4 Proposed technical Architecture

A university has to manage a significant amount of knowledge, adding new information on a daily basis. Such an environment of data provision is complex and includes areas like academic data and educational resources which have to be conform to stakeholders requirements. Traditionally *Service Oriented Architectures (SOA)* have been used to meet these needs. However, as the application domain grows many small and similar services tend to emerge. That phenomena can not only be observed at the Vienna University of Technology, but also at the Open University[30] [43].

A major problem of evolving similar, independent services are diverging data formats and different service owners. Thus, knowledge and administrative information that has been collected by multiple services can not be easily interlinked. An example for such isolated services is the e-learning platform called TUWEL[31], a combination of moodle[32] and the central information system called TISS[33]. These services offer course information and material, but are intended for different purposes. Whereas TISS focuses mainly on administrative functionality TUWEL supports the interaction between teachers and students. Adding additional services which, for example, synchronize deadlines and registration dates is costly due to the fact that the information is spread over different isolated sources and not easily accessible.

In this section we give a short overview of a prototypical publication framework for exposing university related assets as Linked Open Data. At first, a high level view of a publication framework is given (section 4.1). Afterwards, a prototypical architecture facilitating publishing university related assets is introduced. Finally, two ontologies are introduced covering human interactions and the university domain as a whole (section 4.3).

4.1 A publication framework

In this sub-section we provide an overview of a publication framework for exposing university related assets as Linked Open Data, as indicated in figure 3, including:

- *Requirements for Applications*
  These requirements emerge from existing resources (e.g. course material, contact data) and stakeholders input. After gathering requirements there has to be investigated whether a solution based on Linked Data is necessary.

---

[30]http://www.open.ac.uk/
[31]https://tuwel.tuwien.ac.at/
[32]https://moodle.org/
[33]https://tiss.tuwien.ac.at/

- *Data Provision*
  Unfortunately information is often not available in a machine-readable format complicating LOD provision. Conceptionally the generation of Linked Data is either done manually or (semi-) automatic. While the latter is preferable for large data sets manual transformation is the only option if information is present in non digital formats (e.g. printed copies of building plans).

- *Development of an Application*
  The generation of requirements originating from stakeholders input and existing resources is done by the development of applications driven by LOD. Since the domain of Linked Data in particular and Semantic Web in general is relatively new to most developers experts are needed to support the development process.

- *Application Use Case*
  Applications run in a particular environment or domain forming the so-called application use case.
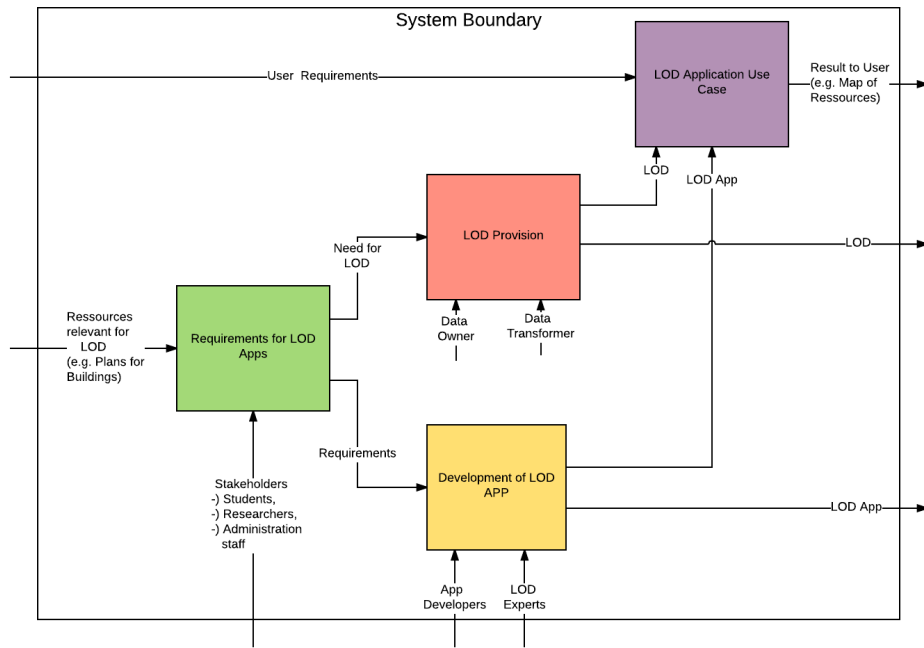
Figure 3: A publication framework driven by Linked Open Data

The publication framework in figure 3 consists of several inbound interfaces shown as arrows pointing from outside of the system and outbound interfaces shown as arrows pointing from inside the of system. Inbound interfaces are listed below:

- *User Requirements*
  Understanding the user requirements is an essential part of software development. However, specifying requirements is not an easy task due to evolving requirements. Not only adapting to stakeholders can become challenging novel technologies and tools are critical factors to the success of a software product. Privacy concerns and legal issues should already be considered in the requirements as misunderstandings are hard to fix in later phases.

- *Existing Resources*
  In the early days of information systems there were certainly any interoperability concerns as huge mainframes work in isolation. With the rise of the World Wide Web distributed, heterogeneous systems emerged, raising integration costs with legacy systems. At the Vienna University of Technology there are a few existing data sets (e.g. contact information of the university staff) which need to be considered.

- *Stakeholders*
  Stakeholders are groups, organizations or individuals who take part in the software development process. Stakeholders have an interest in the success of a software product. Prominent examples are customers, (project) manager and testers.

- *Application Developers*
  Developers create new software products or modify existing ones by implementing (end-user) requirements.

- *LOD Experts*
  Creating remarkable software products in specific domains require people with knowledge superior to those of application developers and collaboration between LOD experts, application developers and customers.

Outbound interfaces are as follows:

- *Linked Open Data*
  A fundamental concept of our proposed model is data connected by links. However, the model does not make any assumptions about the actual data representation format. The data is generated by a data provision process. There exist (semi-) automatic and manual transformation methods depending on the usage of open data formats. Data enrichment and interlinking procedures are applied to data sets to generate metadata properties [42].

- *Application*
  A program or application is the outcome of a comprehensive development process. The user interface is integrated into the application making it the only component where users directly interact with. In modern information systems the application is not a monolithic block installed on a client computer, but rather composed of multiple instances running on different physical machines.

- *End-User Result*
  After a successful launch users interact with the application in many ways. They click on buttons, navigate though menus and fill in forms. All these actions are supported by the user interface.

4.2 Proposal of a prototypical architecture

In order to overcome the environment of data silos and to support the evolvement of a university-wide data space for public resources, a technical architecture following Linked Data principles is proposed and discussed in this section. Berners-Lee suggested four principles for publishing data on the web, which can be considered as best practice. The compliance with this best practice leads to Linked Data: *"the idea is that the more people follow these principles, the more their data will be usable by others."*[36].
The four principles stated by Berners-Lee (2006)[4] are:

1. Use URIs as names for things

2. Use HTTP URIs so that people can look up those names.

3. When someone looks up a URI, provide useful information, using the standards (RDF, RDFS, SPARQL)

4. Include links to other URIs. so that they can discover more things.

Every thing mentioned in one or more datasets of the university shall be identified uniquely by an URI. The Audimax (main lecture hall) of the Vienna University of Technology with the room code BAU176A may, for example, be identified by http://lod.tuwien.ac.at/room/BAU176A and a visit of this HTTP location reveals all the knowledge about this room like capacity, event schedule, . . . in a machine-readable format. This can include links to the hosted events using the same approach with HTTP URIs. In this case, the four principles would be fulfilled and a computer program can easily explore this machine-readable information following those links like humans would naturally would do. As mentioned above,
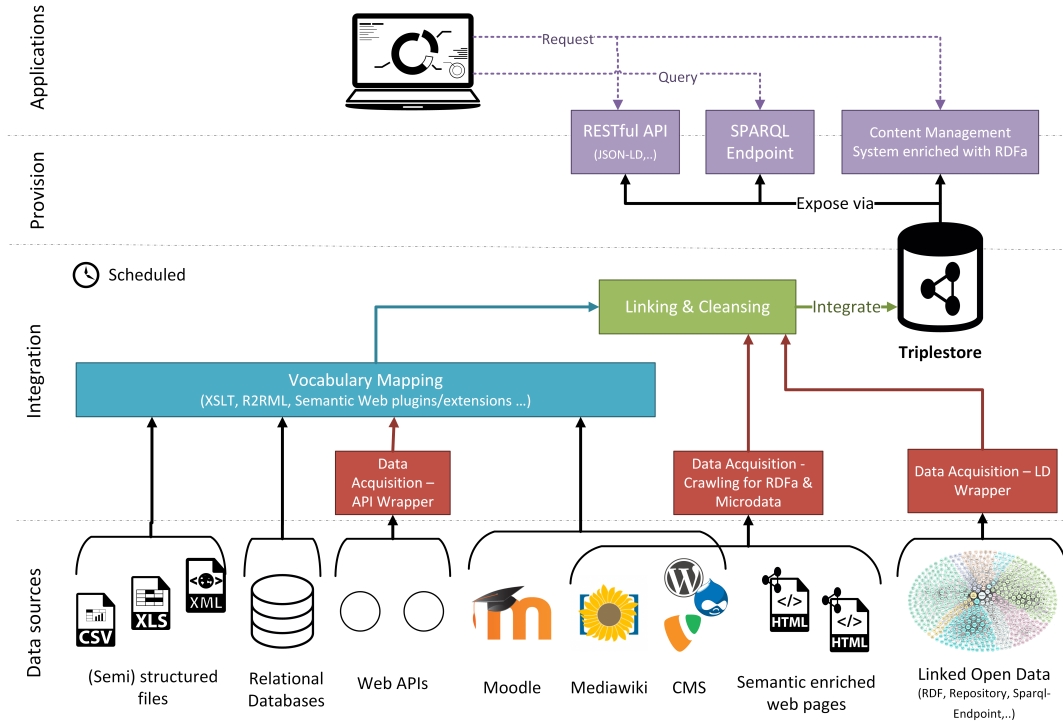
Figure 4: High level architecture for providing Linked Open Data (adapted from EUCLID [36])

following the principles leads to Linked Data. Linked Open Data extends Linked Data with the idea of publishing it in an open way. *"Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)."* [39]

In order to transform demanded datasets of the university into Linked Data and publish it in an open way, a system architecture is required that faces the issues of diverging data sources and data freshness. *"An important architectural pattern used in system development is the multi-tier architecture. It logically separates the functionality of the system in a number of layers and specifies the communication between those layers."* [36]. The proposed architecture (see figure 4) is composed of four tiers. The first tier represents diverging forms of **data sources** that can occur in an information environment at a university (see 4.2.1). The second tier shows the **integration** of such data sources into a central data storage for Linked Data, also called triple store (see 4.2.2). The third tier shows the **provision** of the stored Linked Data (see 4.2.3) so that developers can implement **applications** based on it, represented in the fourth tier.

4.2.1 Data sources

As the information environment of a university may consist of multiple independent services that evolved over time powered by data sources that do not met the principles of Linked Data like relational databases. In order to encourage data owners to publish their data as Linked Open Data, Berners-Lee developed the following star rating system: [4]

★	Available on the web (whatever format) but with an open license (essentially Open Data)
★★	Available as machine-readable structured data (e.g. Excel instead of image scan of a table)
★★★	as (★★) plus non-proprietary format (e.g. CSV instead of Excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can link to
★★★★★	All the above, plus: Link your data to other people's data to provide context

15

There are techniques to integrate these diverging data sources into a five-star Linked Data environment without reconstructing legacy services and work-flows, which is described in the ***integration*** tier. In this section though, different forms of data sources that may occur in the information environment of a university are listed.

Not machine-readable: Not machine-processable sources of knowledge like the expertise of a human (e.g. which rooms are in which way accessible) or not machine-readable data (e.g. images, PDF files) which have to be transformed by humans into data formats that can be processed and interpreted by computers in order to be manageable by the proposed technical architecture. Non machine-readable data like images or PDF files have a rating of one-star (★), if they are exposed on a public accessible web page. These data is essentially considered *Open Data*.

(Semi-)structured data: A quick look at a data repository like http://www.datahub.io shows that major parts of datasets are released as CSV files or proprietary Excel files. The same phenomena can be observed at the Austrian open data portals http://www.opendataportal.at and http://www.data.gv.at. On one side, in this way, it is usual to exchange data at offices and on the other side an easy and inexpensive way to publish data. Universities are no exceptions, the Vienna University of Economics for example releases course information as CSV files on their open data portal. Data published in form of CSV files have a rating of three-stars (★★★), whereas released Excel files have only a rating of two-stars (★★) because of lacking openness of the format. RSS feeds, if used by the university, can also be a source for valuable structured information.

Relational databases: "*Data is often stored in a relational database, perhaps powering an important legacy application. In such cases it is generally advisable to retain the existing data management infrastructure and software, so as not to disrupt legacy applications*".[19] Instead the ***integration tier*** shows how to integrate data originating from relational databases into systems capable of Linked Data. Data sets of relational databases could be presented to the end-user through public web pages in a human readable or in a structured format. The data may also be accessible by a Web API.

Web APIs: Web APIs are one way to give access to particular data of a application in a structured format so that application developers on their hand can use them for their application. It is common to provide results in formats such as XML and JSON. At the Vienna University of Technology, the central information system TISS has a RESTful API, which gives access to some data like course information or the address book. Web APIs achieve a rating of ★★★, if the API uses open and structured formats. However, this Web APIs have some limitations in the perspective of Linked Data. "*Data returned from Web APIs typically exists as isolated fragments, lacking reliable onward links signposting the way to related data. Therefore, while Web APIs make data accessible on the Web, they do not place it truly in the Web, making it linkable and therefore discoverable.*"[19] In either way, Web APIs are a good starting point and the ***integration tier*** shows how to integrate data of Web APIs into a Linked Data system using wrappers.

Moodle: "*Moodle is an open-source e-learning application, written in PHP and widely used by teachers all over the world.*"[30] Such an application is the source of a lot of unstructured data concerning course information and learning materials. Lukichev, Diaconescu, and Giurca [30] described preliminary ideas for empowering the e-learning platform Moodle with rules and semantics, providing tutors a useful application for flexible course management and reporting. At the moment of writing, Moodle does not have a plugin providing this capability. However, Holtham, Rich, and Norris [22] outline the development of a next-generation learning environment based on semantic web-based technologies and pedagogies of the mid-21st century in a position paper for 2020.

MediaWiki: "*Wikis have become popular tools for collaboration on the web and many vibrant online communities employ wikis to exchange knowledge*" [29] At the Vienna University of Technology Mediawiki was employed for sharing experiences and other knowledge about courses as well as about the university itself. Mediawiki is a widely-used open-source wiki written in PHP, originally for use on

Wikipedia. The knowledge is usually exchanged in form of plain text and uploaded files like images, PDF files or similar multimedia content, so that it is not machine-readable. Semantic MediaWiki[34] "*is a semantically enhanced wiki engine that enables users to annotate the wikis contents with explicit, machine-readable information.*"[29] This extension can also be installed for already running MediaWikis. It supports semantically enriched wiki pages and provides a Web API for gathering the machine-readable information.

Content Management Systems: Content Management Systems (CMS) are popular tools for publishing knowledge on the web, "*common used for blogging to express personal or professional views on the world or on observed items that may be of interest to others.*"[8] CMS' are not only used for blogging, but also for more general purposes. The Vienna University of Technology has set up a Typo3[35] CMS to serve valuable information about the university and about some of its organizational units (e.g. interest group for accessibility). As with MediaWikis, the content published by these CMS' is valuable, but often not machine-readable. Particular CMS' provide modules that simplify the annotation of CMS' content with explicit, machine-readable information. Drupal[36] does for example automatically provide meta-data (title, author, topic, etc.) about the published content. In this case, the proposed *integration tier* could integrate those meta-data into the system.

Semantically enriched web pages: Solutions like annotating content published by Content Management Systems and MediaWikis lead to semantically enriched web pages. RDFa is often used alongside microdata and microformats. RDFa is a extension to HTML5 that makes it possible to annotate human-readable information contained in HTML5 pages to provide semantics for machines. "*RDFa is popular in contexts where data publishers are able to modify HTML templates but have relatively little additional control over the publishing infrastructure.*" [19] A snippet of a HTML page using RDFa is shown in Listing 2, where the fact that Heinz Zemanek is an alumni of the Vienna University of Technology is stated.

Listing 2: A snippet of a HTML page using RDFa

```
<div vocab="http://schema.org/" typeof="CollegeOrUniversity">
  <span property="name">Vienna University of Technology</span>
  <link property="sameAs"
        href="http://en.wikipedia.org/wiki/Vienna_University_of_Technology"/>
  <div property="alumni" typeof="OrganizationRole">
    <div property="alumni" typeof="Person">
      <span property="name">Heinz Zemanek</span>
      <link property="sameAs" href="http://en.wikipedia.org/wiki/Heinz_Zemanek"/>
    </div>
  </div>
</div>
```

Microdata and microformats work in a similar way. Not only CMS' employed by the university, also course teams who are employing their own information systems could make use of this technique. A convenient way for data owners to serve data in a human- and machine-readable format. The *integration tier* shows how this data can be extracted and integrated into the system.

Linked Open Data: The Linked Open Data cloud (see figure 5) visualizes datasets of various domains that have been published in Linked Data formats[38]. Berners-Lee emphasizes in his principles for publishing Linked Data that the data shall be interlinked with external Linked Data to give context. Hence, gathering of particular linked datasets must also be considered by the technical architecture for the purpose of interlinking. The ways in which Linked Data is exposed can diverge. It can be done via SPARQL endpoints (see 4.2.3), Web APIs, or RDF dumps, possibly published on a repository like `datahub.io`. For example, the European Library[37] manages the collections of national libraries as well as research libraries of Europe and exports them regularly into a single RDF file. The academic

---

[34]https://semantic-mediawiki.org/
[35]https://typo3.org/
[36]https://drupal.org/
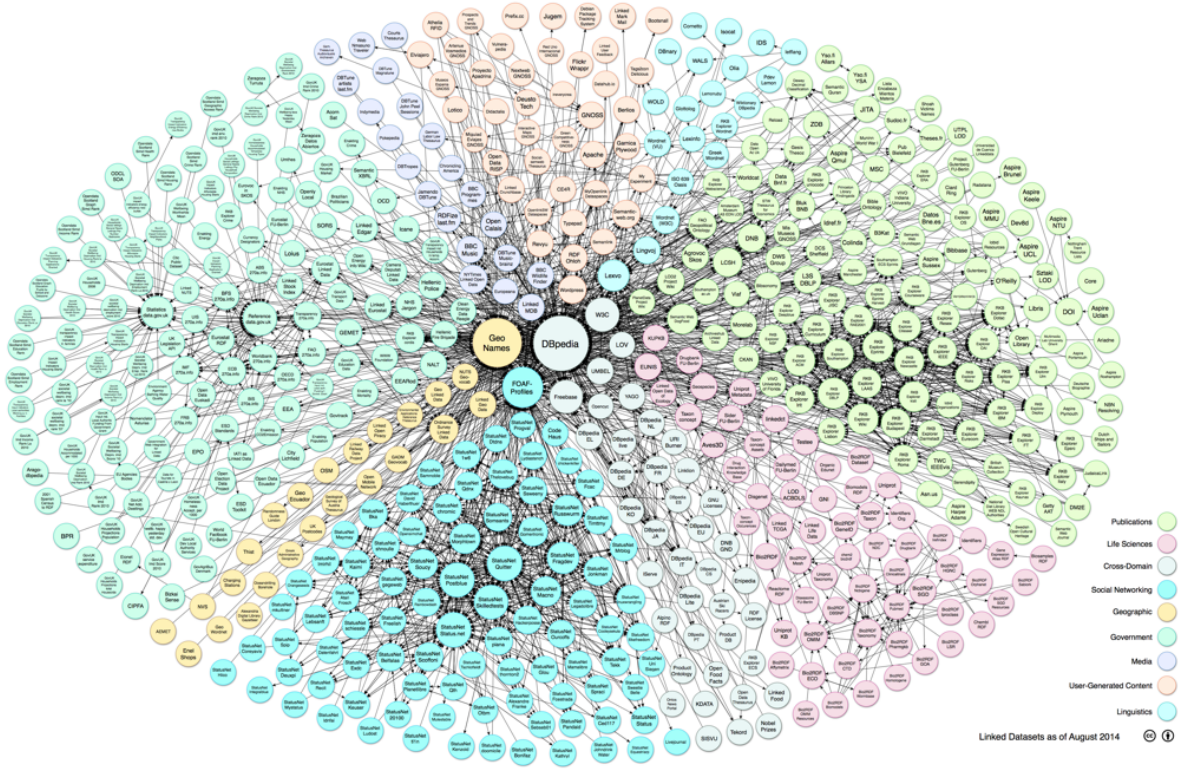[37]http://theeuropeanlibrary.org

Figure 5: LOD cloud diagram of the year 2014

library of the university could interlink items of its own collection with matching items of the European Library project. As a result, the academic library enriches its own knowledge about this item with collected knowledge of the European Library project (e.g. contributors).

4.2.2 Integration

The ***data source tier*** represents diverging data sources, which may be part of the information environment of a university. The ***integration tier*** shows how to integrate the demanded and available datasets of those data sources into the system. The integration process consists of multiple steps, which are described next.

Data acquisition is the first step of the integration process and is required for acquiring the data that shall be integrated into the system. According to Simperl et al. [36] and Heath and Bizer [19] there are a number of ways, in which consumed and integrated data within the system may be accessed. The approach used for this proposed architecture is the *crawling pattern*, where the data is acquired in advance and afterwards integrated and cleansed. The disadvantage of this pattern is that the data will be replicated and might not be up to date. However, "*the crawling pattern is suitable for implementing applications on top of an open, growing set of sources*"[19], which is beneficial for systems that shall integrate public resources of a university. An alternative approach is the *On-The-Fly Dereferencing pattern*, "*where the URIs are dereferenced at the moment that the system requires the data*", hence no local storage is needed. The advantage is up-to-date data, but the performance may be very slow compared to gathering and transforming costs of demanded data.

Vocabulary mapping is the second step of the integration process. In this step the acquired data is translated into the systems target schema (in section 4.3 a collection of vocabularies that may be of interest for a universities domain is listed). The translation is definitely necessary for acquired data that has not already be transformed into Linked Data. For the proposed technical architecture, this would be the case for (semi-)structured datasets like CSV or XML. In the following, some tools are listed for transforming (semi-)structured datasets into Linked Data.

From tables and spreadsheets:

- **Open Refine**[38] is a graphical tool for semi-automatically transforming and cleaning messy data including among others CSV, XML and JSON. A RDF Extension [39] can be used to map the data to ones vocabulary and to export the result into a RDF file.
- `Any23` [40] or `csv2rdf4lod` [41] are Java APIs that can be used for simple transformations from CSV into RDF.

From RSS and XML:

- A common used approach to transform XML into other representations is `XSLT`. W3C recommends the `GRDDL` [42] language for transforming XML into RDF. RSS feeds use a standard XML file format, so the same tools can be applied to it.

From relational databases that can be accessed directly by the system:

- `R2RML` is a W3C recommendation for mapping relational databases to linked data, which can be used in two ways. First, the data can be transformed into a RDF dump by using a specified R2RML mapping. Afterwards, the RDF dump can be integrated. Secondly, SPARQL queries could be translated on-the-fly into relational queries, which is not suitable for the proposed architecture.

From other formats:

- W3C made a document available named `ConverterToRdf`[43] in which tools for transforming several formats into RDF are exposed.

Furthermore, different Linked Data sources often use different vocabularies to represent data about the same type of entity.[5] In this case a vocabulary mapping may also be necessary in order to give the ***applications*** an integrated view of the stored Linked Data.

Interlinking is the third step of the integration process. As mentioned earlier, one Linked Data principle is the interlinking with other Linked Data sets to give context. Interlinking entities of the local triple store with matching entities of other datasets enriches the knowledge about these entities by giving the end user the ability to explore its contents. `Silk` [44] is a Java framework that supports the interlinking of datasets, providing its own specification language (SILK-LSL). It can handle different data sources like SPARQL endpoints and RDF dumps to interlink with. With `Silk`s workbench the linking can be configured graphically.

Cleansing is the fourth step of the integration process. In this step the data that has been already transformed and may be interlinked with other datasets before it is stored (e.g. removal of ambiguities).

Storing is the final step in which the Linked Data is stored in a central triple store. There are a lot of implementations of such a triple store, like `Sesame` [45], `Jena TDB` [46], `Virtuoso Universal Server` [47], `4Store` [48] or `SwiftOWLIM` [49]. The LUCERO project settled for the `SwiftOWLIM`, because it is "*free,*

---

[38]http://openrefine.org/
[39]http://lab.linkeddata.deri.ie/2010/grefine-rdf-extension/
[40]http://lab.linkeddata.deri.ie/2010/grefine-rdf-extension/
[41]https://github.com/timrdf/csv2rdf4lod-automation/wiki
[42]http://www.w3.org/TR/grddl/
[43]https://www.w3.org/wiki/ConverterToRdf
[44]http://silkframework.org/
[45]http://rdf4j.org/
[46]http://openjena.org/TDB/
[47]http://virtuoso.openlinksw.com/
[48]http://4store.org/
[49]http://www.ontotext.com/owlim/

scalable and efficient, and includes limited reasoning capabilities" [11].

4.2.3 Provision

In the ***integration tier*** demanded and available datasets were integrated into a local triple store. The ***provision tier*** gives now an overview of different ways to grant the end user access to the Linked Data stored in the local triple store.

SPARQL endpoint: SPARQL is a language that enables querying of a RDF dataset or in this case the whole triple store. A SPARQL endpoint is a web service that can process SPARQL requests and returns the corresponding results. The advantage of SPARQL endpoints is that the end user can flexibly request a particular subset of the data stored in the local triple store.

RESTful API: "*RESTful APIs are a programming interface implemented using HTTP and the principles of REST (Representational State Transfer) to allow actions on Web resources.*"[24] In order to not violate the principles of LOD the API must be open. `Linked Data API`[50] can be used to provide a RESTful interface to the end user. It makes also SPARQL requests possible so that end users do not have to give up their flexibility.

CMS enriched with RDFa: The previous mentioned techniques for exposing stored Linked Data are not suitable for human readers. They are targeting the application developer as end user. However, *resolvable URIs* are an approach to provide knowledge about an entity contained in the triple store by visiting its unique identifier (e.g. URI). One advantage is the possibility of applying *content negotiation*, a technique where the end user can specify the desired format of the response by modifying the `Accept` entry of the requests header. The default entry for web browsers is one of `application/rdf+xml` - machine-readable or `text/html` - human-readable.
In general, the human-readable response will be a web page, which leads to an additional advantage. These retrieved web pages can be annotated using RDFa to provide semantics readable by machines. The reason why it seems useful, meanwhile the machine-readable knowledge could simply be requested by *content negotiation*, is the ability of nowadays search engines to extract knowledge provided by RDFa or microdata. "*The embedded data is crawled together with the HTML pages by search engines, such as Google, Yahoo, and Bing, which use the data to enrich their search results.*"[6] Schema.org[51] is a movement of major search engines to create, maintain and promote schemas for structured data on the Internet. The idea is to improve the search experience for users.
Whereas *resolvable URIs* are a good approach to serve knowledge in a human- and machine-readable format, it can not be used to query the stored Linked Data. Thus one of the previous mentioned techniques (SPARQL endpoint or RESTful API) must be employed simultaneously to provide this ability. A service that uses *resolvable URIs* and *content negotiation* together with a SPARQL endpoint is DBPedia, which additionally annotates the knowledge published on a web page using RDFa (e.g. http://dbpedia.org/resource/Vienna_University_of_Technology).

4.3 Stakeholder specific concerns

Dealing with stakeholder related concerns involves the definition of domain specific ontologies. In particular ontologies covering the field of organizational data at the Vienna University of Technology. First, this sub-section introduces two ontologies intended but not limited for supporting the generation and evaluation of statistical data of the universities organization.
Gruber [15] defines an *ontology* as:
"*An ontology is an explicit specification of a conceptualization.*"
Conceptualizations are objects, entities or concepts that may or may not exist in the universe. In addition to that, the *vocabulary* defines the relationships between those objects. In other words, the vocabulary defines the conceptual model of what can be represented. Bock et al. [7] describe ontologies from a more practical point of view defining the three conceptual components of an ontology - *classes*,

---

[50]http://www.epimorphics.com/web/projects/linked-data-api
[51]http://schema.org/

instances and *properties*. Regardless of what concrete implementation of an ontology is used graphical representations (e.g. graph) are preferred over textual ones to give a high level overview of the ontology inherent concepts.

4.3.1 University ontology

Figure 6 shows an excerpt of an ontology covering the whole university as an organization.
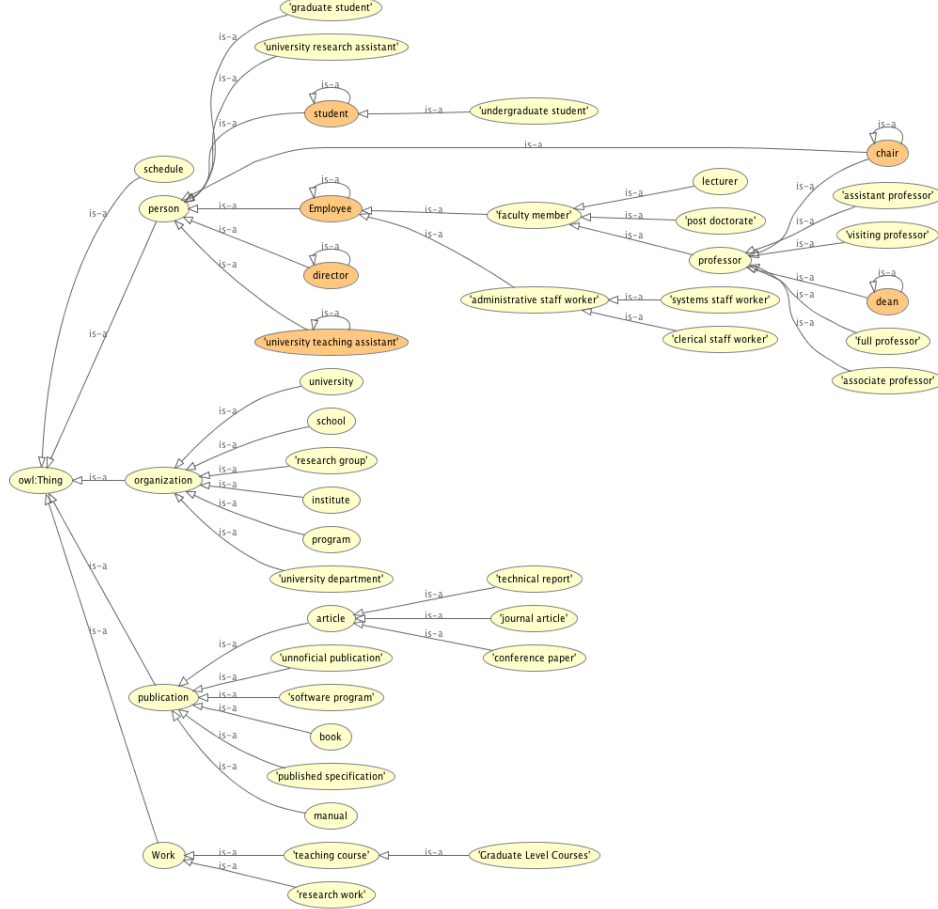
Figure 6: A graph showing an ontology representing the university as an organization

Unfortunately the graph in figure 6 shows just a subset of the overall ontology as the *Protege*[52] plugin *OWLViz*[53] only supports classes and sub-class relationships, omitting object and datatype properties. Its entire definition is retrievable under http://swat.cse.lehigh.edu/onto/univ-bench.owl.
The ontology was originally published by Guo, Pan, and Heflin [16] proposing a benchmark for measuring a variety of properties and several performance metrics. However, this ontology covers the whole university domain extending its areas of application (e.g. for statistical purposes).

Relevant data (sets) at the Vienna University of Technology   At the time of writing this paper the Vienna University of Technology publishes statistical data about active enrollments, beginners and active students. These data is currently publicly accessible[54] through TISS. Figure 7 shows all accessible data.

---

[52] http://protege.stanford.edu/
[53] http://protegewiki.stanford.edu/wiki/OWLViz
[54] https://tiss.tuwien.ac.at/statistik/public_lehre

Öffentliche Statistiken zur Lehre der TU Wien

Allgemeines

- Aktive Studierende (fortgemeldet; ÖH-/Studiengebühren bezahlt) im Vergleich zu den vorangehenden Semestern, inklusive Mitbeleger
  - pro Geschlecht und Staatsangehörigkeit
  - pro Nationalitaet
  - pro Geschlecht und Bundesland
  - pro Geschlecht und Studierendenkategorie

Beginnerstatistik (neu zugelassene Studien)

- Neu zugelassene Studien im Vergleich zu den vorangehenden Semestern, exklusive Mitbeleger
  - pro Studienart und Geschlecht
  - pro Studienart und Staatsangehörigkeit
  - pro Studienrichtung und Geschlecht
  - pro Studienrichtung und Staatsangehörigkeit
  - pro Studienkennzahl und Geschlecht
  - pro Studienkennzahl und Staatsangehörigkeit

Inskriptionsstatistik

- Inskribierte Studien im Vergleich zu den vorangehenden Semestern, inklusive nicht-fortgemeldeter Studien, exklusive Mitbeleger
  - pro Studienart und Geschlecht
  - pro Studienrichtung und Geschlecht
  - pro Studienkennzahl und Geschlecht

Absolventenstatistik

- Erfolgreich abgeschlossene und abgebrochene Studien, exklusive Mitbeleger
  - Absolventen pro Studienart und Geschlecht
  - Absolventen pro Studienrichtung und Geschlecht
  - Absolventen pro Studienkennzahl und Geschlecht
  - Studienabbrecher pro Studienart und Geschlecht
  - Studienabbrecher pro Studienrichtung und Geschlecht
  - Studienabbrecher pro Studienkennzahl und Geschlecht

Figure 7: A screenshot of the publicly accessible data set for statistical evaluation

Technically this data is stored in a relational database as part of the TISS ecosystem. Transformation of these data set would require adding additional metadata to be conform to the ontology presented in figure 6. Furthermore, as that ontology is of rather high level, fitting for almost any educational organization, some adoptions are needed in order to keep the current amount of data. For example gender specific data or any other student related data other than *graduate student* are not covered.

4.3.2 Friend-of-a-friend ontology

Probably the most prominent ontology in the field of describing people and their relations is the *Friend-of-a-Friend (FOAF)*[55] ontology. Despite its simple use case FOAF implements three different concepts: *social networks*, *representational networks* and *information networks*. While social networks represent human collaboration and relationships, representational networks reveal communication relationships [1], e.g. people mentioned together in a communication channel enable a representation of network among them. Last, information networks connect independently published descriptions constituting an inter-connected graph.

Figure 8 shows the generated ontology graph after its import from http://xmlns.com/foaf/spec/20140114.html. The graph below was automatically generated by *Protege*[56] with the help of the plugin *OWLViz*[57] neglecting object and datatype properties.

---

[55] http://xmlns.com/foaf/spec/
[56] http://protege.stanford.edu/
[57] http://protegewiki.stanford.edu/wiki/OWLViz

Figure 8: The FOAF core ontology describing people and their relationships

The FOAF ontology is simple, yet expressive enabling the adoption of many other ontologies in similar contexts. In fact, it is the most widely used ontology on the semantic web [31]. Listing 3 shows a simple example drawn from the FOAF-specification and encoded using the Resource Description Framework (RDF)[28].

Listing 3: RDF/XML representation of a person using the FOAF vocabulary

```
<foaf:Person>
        <foaf:name>Dan Brickley</foaf:name>
        <foaf:mbox_sha1sum>241021fb0e6289f92815fc210f9e9137262c252e</foaf:mbox_sha1sum>
        <foaf:homepage rdf:resource="http://rdfweb.org/people/danbri/"/>
        <foaf:img rdf:resource="http://rdfweb.org/people/danbri/mugshot/danbri-small.jpeg"/>
</foaf:Person .
```

The example in listing 3 shows a person whose name is Dan Brickley. He has a homepage (http://rdfweb.org/people/danbri/) and uploaded an image (http://rdfweb.org/people/danbri/mugshot/danbri-small.jpeg). A randomly generated identifier has been added (mbox_sha1sum) to distinguish it from records with identical data.

Relevant data (sets) at the Vienna University of Technology   Due to the broad applicability and its vast incarnations a single use case can not be spotted. Rather, more complex adoptions have been proposed to describe a comprehensive set of more domain specific properties. Concrete, the project *Semantic Campus* [37] aimed at defining an ontology for describing a universities personnel in more detail.

4.4   Summary

In this section we initially described the idea of a prototypical publication framework for exposing university related information as Linked Open Data. The framework was described in a rather general manner, consisting of *Requirements for Applications*, *Data Provision*, *Application Development* and an *Application Use Case*. Second, an architecture for a general purpose application driven by Linked Data principles was introduced. Third, two selected ontologies representing candidates for modeling a

universities organization at high level and human relations respectively were presented. Last, relevant data sets at the Vienna University of Technology were introduced, though the FOAF-ontology is mostly extended by more sophisticated ontologies rather than used exclusively.

5 Conclusions and Future Work

The high-level goal of this work is to enhance and improve processes at the Vienna University of Technology. We have broken down this goal into four distinct research questions:

- *RQ1: What are best practices regarding the applicability of Linked Open Data in university settings?*

- *RQ2: What are major benefits and barriers for each stakeholder and what are useful use cases?*

- *RQ3: What are major challenges for the implementation of a Linked Open Data solution?*

- *RQ4: How would a prototypical implementation of a publication framework based on Linked Open Data look like?*

In the first part of this paper (section 2) we investigated the first research question and provided international best practices and success stories in educational environments. In the second part (section 3) we focused on benefits and challenges regarding the applicability as well as the implementation of Linked Data principles for administrative workers at the Vienna University of Technology by conducting interviews, addressing the second and the third research question. In the final part (section 4) we addressed research question four by giving a high level overview of a publication framework for educational resources at first (section 4.1) and then providing a general architecture for applications following Linked Data principles (section 4.2).

In this section we discuss conclusions and contributions related to the four research questions we investigated (section 5.1) and present ideas for future work (section 5.2).

5.1 Conclusions and Contributions

In this section we focus on major conclusions and describe contributions we made. This sub-section is organized around our four major research questions.

As this work is highly influenced by the conducted interviews the global conclusion is that the most benefits will be achieved if Linked Open Data principles are implemented right from the beginning, meaning that every process is implemented with these principles in mind. Unfortunately this is almost never the case, therefore requirements and expectations for (end-user) applications and supported processes have to be cut down.

5.1.1 Best practices regarding the applicability of Linked Open Data in educational contexts

In the first part of this paper (section 2) we presented two international projects targeted at spreading knowledge and success stories regarding Linked Open Data principles in educational environments. We hereby conclude:

Using Linked Open Data support is a promising way of enhancing and harmonizing processes at universities. In particular, these achievements have let to several contributions. One major contribution was the creation of ontologies (e.g. AIISO) and tools (e.g. Triplify) applicable in general university settings. Hence, a general guideline should be *using already existing ontologies and tools instead of creating new ones* which cover similar concepts and domains.

Our contribution is to identify and describe these (international) best practices.

5.1.2 Major benefits, challenges and use uses for stakeholders

In the second part (section 3) we guided our attention on benefits and challenges regarding the applicability as well as potential use cases for administrative workers at the Vienna University of Technology. Our work resulted in the following conclusion:

Use application scenarios with large and supportive stakeholder groups. It is much easier and profitable to build applications in a preferably large and supportive environment. Nothing hinders progress more than a skeptical target audience. As we have seen from interviews the group of administrative workers and researchers are small and manageable. Rather, to take advantage and spreading knowledge of Linked Open Data principles we suggest choosing students as target stakeholders. In addition, offering courses, provide students the ability to develop applications driven by LOD and helps in enhancing and maintaining LOD data sets.

Our contribution is to get a better understanding of the stakeholders needs at first and then identify potential use cases and aspects in realizing these.

5.1.3 Major challenges for implementing applications following Linked Open Data principles

Also in the second part (section 3) we additionally provided challenges and obstacles in application development. Therefore, the following conclusion were drawn:

We recommend using existing, therefore maintained and open data sets as a foundation of new applications. The cost of integrating large data sets can become quite time consuming if independent domains are mixed. Even for smaller, more limited domains using existing data sets outweighs the benefits of creating new, fully customized data sets. If additional requirements arise modifying and/or adding to existing data is preferable. A cost minimizing strategy for keeping the data clean and manageable is to clean existing data sets before adopting them to Linked Data principles. Rahm and Do [33] gives a more in-depth introduction into data cleaning approaches and challenges.

Our contribution is to identify potential challenges and obstacles in implementing applications driven by LOD.

5.1.4 Prototypical implementation of a publication framework

In the final part (section 4) we addressed research question four by introducing a publication framework as well as a general application architecture using Linked Data principles for university related resources. We come to the following conclusions:

Applications making use of LOD should be built following a general architecture. There have already been efforts at building LOD applications at enterprise level, sharing a similar foundation. An agreed picture of a successful architecture facilitates establishing common vocabularies and understandings under developers, customers and management.

Our contribution is to provide a general architecture for LOD applications, hence supporting developers in creating these kinds of applications.

5.2 Future Work

Contributions and methods of this work is intended as an initial step to investigate the adoption of Linked Data principles and spread awareness and knowledge at the Vienna University of Technology. The following list summarizes planned future work:

- `(Educational) courses aimed at spreading Linked Data principles`
  Educational courses are great opportunities to facilitate building of LOD applications and maintaining related data sets.

- `Building (end-user) LOD applications`
  Application development is not limited to resulting from educational courses. Rather, due to the

heterogeneity of universities data sets, a collaboration between external parties and university representatives is required.

- `Consolidated report to executive board`
  There will be a followup report, summarizing our contributions for students, researchers and administrative workers. This report is addressed to the executive board of the Vienna University of Technology. A side-product of this report will be a brief presentation about the overall result and achievements, planned in the near future.

6 Acknowledgements

The authors would like to thank `Assoc. Prof. Dr. Stefan Biffl`[58] and `Dr. Marta Sabou`[59] who had the initial idea. They provided great feedback and supported us in writing this report and finding participants for the conducted interviews.
Special thanks goes the interviewees (`Mag.iur. Dr.iur. Jasmin Gruendling-Riener`[60] and `Dipl.-Ing. Wolfgang Spreicer`[61]) for their time and thoughts.

---

[58]stefan.biffl@tuwien.ac.at
[59]marta.sabou@ifs.tuwien.ac.at
[60]jasmin.gruendling-riener@tuwien.ac.at
[61]wolfgang.spreicer@tuwien.ac.at

References to refereed scientific work

[1]   George A Barnett. *Encyclopedia of social networks*. 2011.

[3]   Harvey Russell Bernard. *Research methods in cultural anthropology*. 1988.

[5]   Tim Berners-Lee and Lalana Kagal. "The fractal nature of the semantic web". In: *AI magazine* 29.3 (2008), p. 29.

[6]   Christian Bizer et al. "Deployment of RDFa, Microdata, and Microformats on the Web — A Quantitative Analysis". In: *The Semantic Web — ISWC 2013: 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part II*. Ed. by Harith Alani et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 17–32. ISBN: 978-3-642-41338-4. URL: http://dx.doi.org/10.1007/978-3-642-41338-4_2.

[8]   Uldis Bojrs, John G. Breslin, and Knud Mller. "Using semantics to enhance the blogging experience". In: *The Semantic Web: Research and Applications, 3rd European Semantic Web Conference, ESWC 2006, Budva*. Springer, 2006, pp. 679–696.

[14]   Mathieu DAquin. "Putting linked data to use in a large higher-education organisation". In: *Proceedings of the Interacting with Linked Data (ILD) workshop at Extended Semantic Web Conference (ESWC)*. Citeseer. 2012.

[15]   Thomas R. Gruber. "A translation approach to portable ontology specifications". In: *Knowl. Acquis.* 5.2 (1993), pp. 199–220.

[16]   Yuanbo Guo, Zhengxiang Pan, and Jeff Heflin. "LUBM: A benchmark for OWL knowledge base systems". In: *Web Semantics: Science, Services and Agents on the World Wide Web* 3.2 (2005), pp. 158–182.

[18]   Lois Ruth Harris and Gavin Thomas Lumsden Brown. "Mixing interview and questionnaire methods: Practical problems in aligning data". In: (2010).

[19]   Tom Heath and Christian Bizer. *Linked Data: Evolving the Web into a Global Data Space (1st edition)*. Vol. Synthesis Lectures on the Semantic Web: Theory and Technology. Morgan & Claypool, 2011.

[23]   Maarten Hoogerwerf et al. "Linking data and publications: towards a cross-disciplinary approach". In: *International Journal of Digital Curation* 8.1 (2013), pp. 244–254.

[29]   Markus Krtzsch, Denny Vrandei, and Max Vlkel. "Semantic MediaWiki". In: *The Semantic Web - ISWC 2006: 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006. Proceedings*. Ed. by Isabel Cruz et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 935–942. ISBN: 978-3-540-49055-5. URL: http://dx.doi.org/10.1007/11926078_68.

[30]   Sergey Lukichev, Mircea Diaconescu, and Adrian Giurca. "Empowering Moodle with Rules and Semantics". In: ().

[31]   Miltiadis D Lytras et al. "Social networking on the semantic web". In: *The Learning Organization* 12.5 (2005), pp. 418–435.

[32]   Jeremy Miles and Paul Gilbert. *A handbook of research methods for clinical and health psychology*. Oxford University Press on Demand, 2005.

[33]   Erhard Rahm and Hong Hai Do. "Data cleaning: Problems and current approaches". In: *IEEE Data Eng. Bull.* 23.4 (2000), pp. 3–13.

[35]   Per Runeson and Martin Höst. "Guidelines for conducting and reporting case study research in software engineering". In: *Empirical software engineering* 14.2 (2009), pp. 131–164.

[36]   E. Simperl et al. *Using Linked Data Effectively*. The Open University, 2013. ISBN: 978-1-78007-926-4. URL: http://www.euclid-project.eu/.

[37]   Natenapa Sriharee and Ravikarn Punnarut. "Constructing Semantic Campus for Academic Collaboration." In: *FEWS*. 2007, pp. 23–32.

[41]   Robert K Yin. *Case study research: Design and methods*. Sage publications, 2013.

[42] HQ Yu et al. "A Linked Data-driven & Service-oriented Architecture for Sharing Educational Resources; Linked Learning 2011". In: *1st International Workshop on eLearning Approaches for the Linked Data Age, Heraklion, Greece.* 2011.

[43] Fouad Zablith et al. "Consuming Linked Data within a Large Educational Organization". In: Bonn, Germany, 2011.

References to non-refereed work

[2] Lukas Baronyai. "Publishing TUWien Data as Linked Open Data". In: (2016).

[4] Tim Berners-Lee. "Linked Data - Design Issues". In: *W3C* 09/20 (2006). URL: http://www.w3.org/DesignIssues/LinkedData.html.

[17] Kevin Haller. "Linked Open Data at the Vienna University of Technology - A case study concerning students". In: (2016).

[22] Clive Holtham, Martin Rich, and Leona Norris. "Moodle 2020: A Position Paper". In: 2012.

[24] Bernadette Hyland et al. *Linked Data Glossary.* URL: https://www.w3.org/TR/ld-glossary/#rest-api (visited on 03/30/2016).

[26] Ian Jacobs and Norman Walsh. *Architecture of the World Wide Web, Volume One.* W3C:rec. http://www.w3.org/TR/2004/REC-webarch-20041215/. W3C, 2004. URL: http://www.w3.org/TR/2004/REC-webarch-20041215/.

[27] Marijn Janssen, Yannis Charalabidis, and Anneke Zuiderwijk. "Benefits, adoption barriers and myths of open data and open government". In: *Information Systems Management* 29.4 (2012), pp. 258–268.

[28] Graham Klyne and Jeremy J Carroll. "Resource description framework (RDF): Concepts and abstract syntax". In: (2006).

References to websites

[7] Conrad Bock et al. *OWL 2 web ontology language structural specification and functional-style syntax.* [accessed 19-February-2016]. 2008.

[9] World Wide Web Consortium. *Gleaning Resource Descriptions from Dialects of Languages (GRDDL).* [accessed 19-February-2016]. 2007.

[10] Mathieu D'Aquin, Carsten Kessler, and Tomi Kauppinen. *Members of Linked Universities.* http://linkeduniversities.org/lu/index.php/members/index.html. [accessed 8-February-2016]. 2014.

[11] Mathieu D'Aquin et al. *The LUCERO Project: Tabloid.* http://lucero-project.info/lb/tabloid/index.html. [accessed 07-March-2016]. 2010.

[12] Department for Studies and Examinations of the University of Technology in Vienna. *Homepage.* http://www.tuwien.ac.at/de/dle/studienabteilung. [accessed 19-February-2016]. 2015.

[13] Hendrik Drachsler et al. *Final Version of the LinkedUp Evaluation Framework.* http://linkedup-project.eu/files/2014/11/LinkedUp_D2.2.2.pdf. [accessed 28-March-2016]. 2014.

[20] Eelco Herder, Marieke Guy, and Hendrik Drachsler. *LinkedUp Challenge Results.* http://linkedup-project.eu/files/2014/11/LinkedUp_D1.3.pdf. [accessed 27-March-2016]. 2014.

[21] Eelco Herder et al. *State of the art and data assessment.* http://linkedup-project.eu/files/2013/09/LinkedUp-D1.1-SotA.pdf. [accessed 25-March-2016]. 2013.

[24] Bernadette Hyland et al. *Linked Data Glossary.* URL: https://www.w3.org/TR/ld-glossary/#rest-api (visited on 03/30/2016).

[25] Information Technology Services (ZID) of the University of Technology in Vienna. *Address book - Administration & Services.* https://tiss.tuwien.ac.at/adressbuch/adressbuch/administration. [accessed 19-February-2016]. 2016.

[34]   Research and Transfer Support of the University of Technology in Vienna. *Homepage*. http://www.
       tuwien.ac.at/dle/forschungs_und_transfersupport. [accessed 19-February-2016]. 2015.

[38]   *The Linking Open Data cloud diagram*. URL: http://lod-cloud.net/ (visited on 03/30/2016).

[39]   *The Open Definition - Open Definition - Defining Open in Open Data, Open Content and Open
       Knowledge*. URL: http://opendefinition.org/ (visited on 03/27/2016).

[40]   University of Technology in Vienna. *Homepage*. https://www.tuwien.ac.at/. [accessed 18-February-
       2016]. 2016.

A Questionnaire

General Questions

Which organizational area would you associate your work with?

Research	☐
Student Affairs	☐
Administration	☐
Other	_____

How would you describe your area of responsibility? How would you characterize your *daily* work tasks?

_____

_____

_____

How would you say classify your level of experience with Information Systems?

Novice[1]	Advanced Beginner[2]	Competent[3]	Proficient[4]	Expert[5]
☐	☐	☐	☐	☐

How would you classify your level of expertise with Linked Open Data?

never heard	heard but never used	used in small example	used in practice	Expert in LOD
☐	☐	☐	☐	☐

---

[1]No experience with the situations in which they are expected to perform tasks.

[2]Has coped with enough real situations to note the recurrent meaningful situational components.

[3]Has already gained a considerable amount of routine to do tasks by deliberate planning to reach long-term goals. (e.g. do things in a organized way)

[4]Perceives situations as wholes, rather than in terms of aspects, and performance is guided by maxims.

[5]Has a vision of what is possible and uses analytical approaches in new situations or in case of problems.

LOD in Administration

How useful do you find the improvements with Linked Data in building maps?

not useful		useful		extremely useful
☐	☐	☐	☐	☐
1	2	3	4	5

Why?

_____

_____

How useful do you find publishing Linked Data for statistical evaluation?

not useful		useful		extremely useful
☐	☐	☐	☐	☐
1	2	3	4	5

Why?

_____

_____

Can you imagine similar projects at the Vienna University of Technology? Yes, in particular:

_____

_____

No, because:

_____

_____

What benefits do you see?

_____

_____

What disadvantages or barriers do you see?

_____

_____

Can you imagine/recommend other roles/persons?

_____

_____

LOD Applications

Do you know any useful application of LOD which will help you in your daily work or will make internal procedures at the university more efficient?

_____

_____

_____

_____

What benefits do you see?

_____

_____

What disadvantages or barriers do you see?

_____

_____

Can you imagine/recommend other roles/persons?

_____

_____

LOD Data

Do you know any other data regarding administration or in general which can be published as LOD?

_____

_____

_____

_____

What benefits do you see?

_____

_____

What disadvantages or barriers do you see?

_____

_____

Can you imagine/recommend other roles/persons?

_____

_____

LOD in general

Do you have any other kinds of suggestions regarding to LOD in general or with focus on administration?

_____

_____

_____

_____