Semantically Augmented Documents for Use in Higher Education Institutions

Analyzing the Current State in the Digital Transformation of HEI

Karsten Böhm
FH Kufstein Tirol University of Applied Sciences
Kufstein, Austria
email: karsten.boehm@fh-kufstein.ac.at

Abstract—Higher Education Institutions (HEI) are part of the Digital Transformation of our societies and a higher level of digitization and automatization is already establishing within the sector. The Information and Communication Technology (ICT) landscapes, the data models and the processes are highly individual to the HEI and interoperability and linking of data is a challenge and often creates new digital format discontinuities. This contribution analyses the support level of recent semantic models, namely the European Learning Model and the Educational Verifiable Credentials Model for the application in an Austrian university with respect to graduation documents. The results show the application potential of the new models and relates it to current practices in representing data of academic programs in a meaningful way. An implementation demonstrates use-cases with an immediate effect for HEI and focuses on attractive and lightweight User Experience (UX) to ensure user adoption.

Keywords—Digital Transformation; HEI; Linked Data; European Learning Model; Verifiable Credentials Model.

I. INTRODUCTION

Higher Education Institutions (HEI) are offering education programs with a more or less planned learning journey to reach specified qualification objectives within a qualification framework (e.g., the European Qualification Framework [1]). As a result of the Bologna process of aligning the national education systems in Europe, the education system became more transparent and interoperable – on a national, but also on an European level. While being a positive development for students, this also meant an increasing number of stakeholders for the HEI. They need to be integrated in the internal processes and information systems, e.g., when Recognizing Prior Learning (RPL) for a study program [2] or to include student mobility into a course program. Together with the development of a digital transformation in many organizations, this results for HEI in a higher level of digitization and automatization that is already establishing within the sector. However, the ICT landscapes, the data models and the processes are highly individual to the HEI and interoperability and linking of data is a challenge. Often, this situation creates new digital format discontinuities and requires additional efforts for the organizations, staff and students of a HEI.

At the other hand, there are a number of interesting developments at the European level that try to harmonize and digitize the information exchange within and among HEI. Among the most notable are the new version of the Education Learning Model (ELM) [3] and the W3C standard on the Verifiable Credential Model (VCM) [4] that recently had been extended with a version for the educational sector, the Educational Verifiable Credential Model (EVC) [5]. Both models address the need to formalize and harmonize the unstructured information and document sources that limit the automated processing of relevant documents that maybe already digital in format but very diverse with respect to the semantic representation.

Common ways to document the achievements of academic programs are the two document categories Diploma Supplement (DS) and Transcripts of Records (ToR). While a DS describes the general program aspects and the individual properties of the student, the ToR contains information about the subjects taken and the grade for each subject [6], [7]. Both documents aim at describing the specific aspects of the program and the achieved result of the student in great detail and are mandatory to be generated by Austrian universities for all graduates. However, currently those documents are using a predefined structure, the content is only text based. DS and ToR need to be validated and durable, even when the course program or even the organization is not operational anymore. Therefore, it is important that the document is self-contained and independent from an organization or a technical system.

A ToR is already specifying qualifications, but only as textual data. For automated processing and the support of all stakeholders in a HEI (student, lecturers, managers) it would be useful to have that in a machine-readable way, as the information is often generated from ICT-systems anyway.

This contribution builds on the concept of educational Pre-built Information Spaces (PreBIS-ED) [8] and is researching the current state of the art concerning certification documents with a focus on the situation in Austria and a specific university as an illustrating example. It aligns the features of the aforementioned semantic models with the existing specification documents required in Austria and elsewhere in Europe. By doing so, it assesses the potential of the semantic models to be used in everyday processes within HEI and it identifies the gaps that still exist.

It contributes to the development of the digital excellence of HEI by addressing two important needs: 1) For HEI, the ability to process existing qualifications, e.g., from a first study cycle with the reference to RPL, Recognition of Prior Learning and 2) for companies it would be beneficial to have an opportunity to match qualifications against their job profiles, e.g., when doing a Job Task Analysis (JTA), [9].

While addressing a very specific and practical application area, PreBIS-ED also tries to find new answers for the more general research question on how (existing) semantic models can be put in operational practice by different stakeholders

 ISBN: 978-1-68558-108-4

that are not knowledge engineering experts. The approach focuses on two main aspects 1) putting the model to use with a clear benefit for the user (increased transparency and consistency in this case) and 2) creating an attractive and lightweight UX for areas that are manually crafted anyway (curricular structures and competences in this case).

The rest of the paper is structured as follows: In the Section 2, the current situation in an Austrian university with respect to the ToR and the DS documents is introduced – a situation that is typical for the Austrian sector and very common for universities all across Europe. Afterwards the predefined structure of the ToR and DS documents is matched against the semantic models ELM and EVC to determine the usefulness for those document categories. In Section 4, the implementation of a technical solution is outlined that brings the semantic models together with the existing document renderings. The paper concludes in Section 5 with an outlook on the next steps within these research activities.

II. The Case of an Austrian University

The HEI of the author is already using DS and ToR documents that are supplied automatically to each student that graduates from a bachelor's program or master's program. The documents are generated by the internal campus management system as PDF documents and are digitally signed using the electronic signature supplied for all public administration services in Austria [10]. The organization follows the rules set up by the European Commission and is granted the Diploma Supplement Label and the ECTS Label (ECTS – European Credit Transfer System) for adhering to those standards [11].

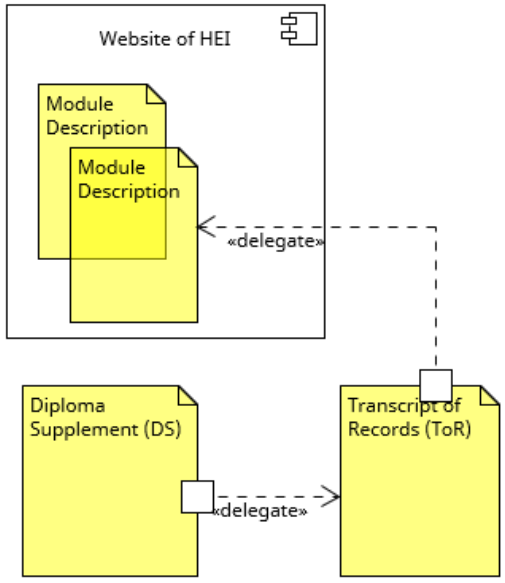

Figure 2. Relation of the different specification documents which are delivered to the graduating student (DS and ToR) and contain references to the module descriptions which are available online at the HEI website.

The documents contain more detailed information about the formal properties of the degree and also overview information about the qualification and learning objectives. In the case of this HEI the respective Section in the Diploma Supplement (Section 4.3) refers to the ToR document that contains more detailed information about the subjects in the degree program and the individual student results.

For competences and learning outcomes, it refers to the website of the HEI that contains detailed information about the curriculum and the learning outcomes in the general module description. Figure 1 illustrates the relation between the different information sources.

Currently, the DS/ToR does not contain any information about specific learning outcomes (but refers the user to the website of the university as an external information source) and is not semantically enriched. Therefore, the documents provide the potential for an extension with semantic information about specifically acquired competences during the program in order to support the role of interfacing between different HEI or the Human Resources departments of companies.

The information on the website is representing a static version of the program as such as it describes the learning outcomes in more general way; its implementation in the different semesters can vary in (within the defined scope).

The current situation with respect to the documentation of competences can be summarized in the following way:

(a) Information about the learning outcomes is stored distributed in three different locations two of which are delivered as documents to the students at the time of graduation and the third is stored online at the HEI website.
(b) The information is human readable only and needs interpretation and research depending on the level of detail needed.
(c) Due to its distributed nature, it is not self-contained and requires the availability of additional online-services to be used.
(d) As the documentation references the module description of a program and not the yearly made syllabus, it might not cover specific details, such as the used technology or software system used in the execution of a lecture in a certain semester, thus lacking some level of detail.

With respect to a more seamless integration of the information into other systems or organizations, it would be beneficial, if those properties could be improved by helping users of this documentation to have a more complete and easier to use data collection at their disposal. As the documentation is generated by an internal campus management system that is being developed by the HEI itself, there is the possibility to develop an improved solution to address this need.

III. Matching the ToR/DS to the Semantic Models

In order to estimate the potential of the concepts of the semantic models for improving the expressiveness of the ToR and DS document categories, the individual properties of each characteristic will be matched against the concepts from the

 ISBN: 978-1-68558-108-4

semantic models. It should be noted that the attributes of a ToR document and a DS document are defined by the European Commission and augmented by National Authorities [12], but the values are mostly free-text form, individual for each HEI and also subject to change. This makes the information less interoperable.

Using semantic models like the ELM and the EVC provide a scaffolding that could be helpful to create more consistent and interoperable representations for the values of the DS/ToR documents. In order to approve the suitability of this approach, the fields of DS and ToR documents are mapped to the semantic models in the next Sections.

A. Analysis of the different features in a DS

To add semantic information to a DS document, both the ELM and the EVC are good candidate models to encode the information in the current documents with an explicit semantic specification. Table I below provides a mapping of the fields defined for a DS document and available in the current DS of the author's HEI with classes (sometimes with properties) from the semantic models ELM and EVC.

TABLE I. ATTRIBUTES OF A DS DOCUMENT (LEFT COLUMN) AND APPROPRIATE SEMANTIC CLASSES AND PROPERTIES (RIGHT COLUMN). ABBREVIATIONS: EVC – EDUCATIONAL VERIFIABLE CREDENTIAL, VC – VERIFIABLE CREDENTIAL, ELM – EUROPEAN LEARNING MODEL

1. INFORMATION IDENTIFYING THE HOLDER OF THE QUALIFICATION	
1.1 Last name(s)	VC: holder
1.2 First name(s)	VC: holder
1.3 Date of Birth	VC: holder
1.4 Student identification number	VC: holder
2. INFORMATION IDENTIFYING THE QUALIFICATION	
2.1 Name of qualification, title conferred	EVC: Credential
2.2 Main field(s) of study for the qualification	EVC: Credential Subject
2.3 Name and status of awarding institution	EVC: Issuer
2.5 Course languages	ELM: language, default language (properties)
3. INFORMATION ON THE LEVEL AND DURATION OF THE QUALIFICATION	
3.1 Level of the qualification	ELM: QF level (property)
3.2 Official duration of program in credits and/or years	ELM: Credit Points
3.3 Access requirement(s)	
4. INFORMATION ON THE PROGRAM COMPLETED AND THE RESULTS OBTAINED	
4.1 Mode of study	
4.2 Program learning outcomes	ELM: Learning Outcome
4.3 Program details, individual credits gained and grades/marks obtained	ELM: Learning Achievement
4.4 Grading scheme, grade translation and grade distribution guidance	ELM: Grading Scheme
4.5 Overall classification of the qualification	ELM: Qualification, Qualification Reference
5. INFORMATION ON THE FUNCTION OF THE QUALIFICATION	
5.1 Access to further study	ELM: Learning Entitlement Specification
5.2 Access to a regulated profession (if applicable)	ELM: Learning Entitlement, Learning Entitlement Specification
6. ADDITIONAL INFORMATION	
6.1 Additional information	(Arbitrary textual information)
6.2 Further information sources	(Arbitrary textual information and URLs to Web Resources)
CERTIFICATION OF THE SUPPLEMENT	
7.1 Date	EVC: Issuance Date
7.2 Signature	EVC: Proof (Digital)
7.3 Capacity	EVC: Proof
7.4 Official stamp or seal	EVC: Credential Proof

It is shown that all the concepts of a DS can be captured in the new semantic models and thus provide a good starting point for an interoperable implementation. The study also shows that EVC focusses more on the administrative information of the DS whereas ELM focusses on the expressiveness of the education related aspects like qualifications and learning outcomes.

For the use case of understanding the (overall) qualification of a DS holder, Sections 3 and 4 of a DS are the most interesting parts and the modelling of ELM is more important for that field of application. Since ELM is developed by the European Commission and the DS document is a mandatory document for Austrian (and probably also for most European) universities, there is a good chance that the semantic encoding of the already defined fields can lead to interoperable and machine-readable specifications. Further evidence for this assumption was collected with a workshop series with members from different European Universities in the CloudEarthI-project [13] held in 2022 and 2023.

Currently, to the best knowledge of the author, no implementations of semantic information on the DS documents have yet been implemented by Austrian universities. This might be due to the fact that version 3 of ELM is just about to be released in its final version, according to [3]. The contribution can be thought as an initial activity to bring the emerging semantic models into operational effectiveness. Once this is achieved with application cases like augmenting the DS and ToR document, additional benefits can be addressed like automated status reports on inconsistencies with learning outcomes or credit point achievement, missing topics and matching qualifications. Machine readable semantic representations will also help to automate processes of validation and verification in a reliable way as it does not depend on ambiguous text representations. This will contribute to accelerate the Digital Transformation of the HEI domain.

B. Analysis of the different features in a ToR

While the DS only holds the overall qualification information, the ToR should document qualifications and learning outcomes in a more detailed level. Usually, a ToR document contains the information represented in Table II, below. This is also the case in the author's HEI. The usual

 ISBN: 978-1-68558-108-4

tabular presentation is ordered by the chronological order of the study program, e.g., into semesters in the case of the author's HEI. The items in Section 2 of Table II will occur for each subject grouped by the semesters in which they are positioned. In addition to the values for the lectures there are also derived information that are calculated per semester, e.g., the amount of credits (usually 30 per semester) and the average grade over all subjects studied in one semester.

TABLE II. ATTRIBUTES OF A TOR DOCUMENT (LEFT COLUMN) AND APPROPRIATE SEMANTIC CLASSES AND PROPERTIES (RIGHT COLUMN). ABBREVIATIONS: EVC – EDUCATIONAL VERIFIABLE CREDENTIAL, VC – VERIFIABLE CREDENTIAL, ELM – EUROPEAN LEARNING MODEL

1. INFORMATION IDENTIFYING THE HOLDER OF THE QUALIFICATION	
Last name(s)	VC: holder
First name(s)	VC: holder
Date of Birth	VC: holder
Student identification number	VC: holder
2. INFORMATION ON THE DIFFERENT LECTURES (GROUPED INTO SEMESTERS)	
Course Title	ELM: Identifier
Course Code	ELM: Identifier
Language of lecture	ELM: language (property)
Contact Hours of lecture	ELM: contact hours (property)
Credit point of lecture	ELM: credit received (property) ELM: volume of learning or workload (property)
Credit points per semester	ELM: credit points (property)
Grade (value and per cent)	(calculated value, information only)
Average grade per semester	(calculated value, information only)
CERTIFICATION OF THE SUPPLEMENT	
Date	EVC: Issuance Date
(Digital) Signature	EVC: Proof
Capacity	EVC: Proof

It is interesting that there are expected similarities in the Sections of the holder information and the certification between DS and ToR, but it is even more noteworthy that the ToR does not contain information about detailed qualifications and learning outcomes in a more specific way.

The only hint on the content of the respective lectures are the names of the lectures that might suggest the content and possible learning opportunities. Relevant information should be found elsewhere, in the case of the author's HEI at the website of the course program. The link between the ToR and the relevant content on the website is made by the name or course code of the lecture only.

This analysis shows that there is a high potential of embedding semantic information into the ToR to make competence information more visible at the level of individual lectures. ELM provides a number of useful classes to model and link this information together such as the classes "Learning Achievement" and the class "Learning Achievement Specification" with the properties "learning outcome" and "learning outcome summary".

Linking could be achieved by using the property "content URL" to refer to the website content with deep linking. This, however could also create the problem of broken links, if the structure of a website changes or if the online resource is not available. In the use case of ToR and DS document the information also need to remain static in the sense that it needs to reflect the information that was current at the time of document creation. Changes in the linked resources could even introduce semantic errors, if the updates content does not fit anymore. It would be better to embed the information into the document directly and use the linked information only as a secondary or supplemental resource. Since the curricular information on the website of the author's HEI are also generated from the same data source as the DS and ToR documents, this is technically feasible to implement.

C. Summary on the potential of semantic models

After mapping the semantic models to the ToR and DS documents it becomes clear that the current application case for the EVC is mostly targeted at modelling the outcomes (certificates) of the program, while the ELM in its new version is more versatile to model important concepts on several levels in the HEI. As illustrated in Figure 2, it can be used at the academic process level to support the operational processes and – most prominently – support the certification process by providing documents that are easier to process by interfacing stakeholders (e.g., other HEI or companies that are

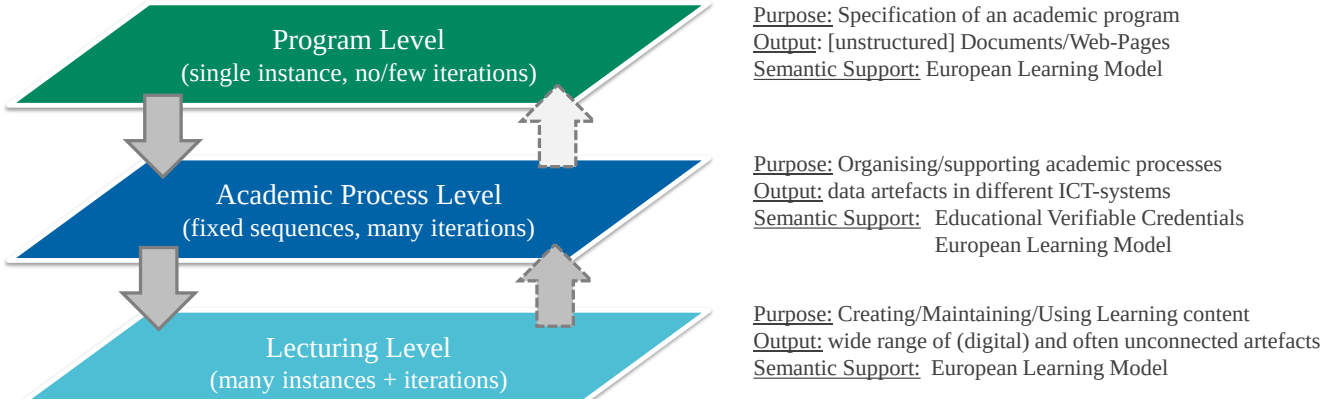

Figure 2. Overview on the different levels in the planning and the execution of academic programs with their semantic support levels by existing semantic models

 ISBN: 978-1-68558-108-4

hiring the graduates). Apart from this middle level, the semantic model could also support the design process (top-level) for aligning the learning objectives among the different modules and towards the qualification profile. Likewise, the operational activities of the lecturers can be supported by using the ELM in tools that help them to create course structure and materials that are linked to each other using learning outcomes.

IV. IMPLEMENTATION

The implementation of a DS and ToR documents with extended semantic information could be provided by different means and does also differ depending on the document format. The most important formats for this application case being HTML for the embedding in web-oriented applications (e.g., the HEI website, a web-based e-learning-system) and PDF for issuing the documents for students and other stakeholders. The use of metadata, microformats and embedded JSON-LD data could be a suitable approach for HTML, as this information is also recognized by search engines as structured markup and can be used to improve the search results based on facts and not only on text snippets [14] [15]. It seems that JSON-LD will be the preferred structured data format in the future by Google, according to the Search-EngineJournal [16], which is compatible with ELM and EVC, since both provide JSON-LD representations.

Simple meta-data formats (as key-value pairs) are also available for PDF, but another promising way is the use of the attachment feature of the PDF-format [17]. There are a number of different ways to relate existing files to a PDF document such as associating or referencing files, but for the purpose of this application the embedding of files as a file stream into the container PDF is the most suitable way, as it provides (a) the property of self-containment and (b) can be identified, used and even exported by the user. Using a suitable library such as PDFLib [18] makes it possible to augment the already generated document with semantic information, e.g., using the JSON-LD format [19] that is also used with ELM. This way the existing generation process does not need to be changed and the augmentation with the semantic information can be added as an additional step in the process.

Furthermore, the use of the attachment feature also has the benefit that the users can identify the augmented data, if the PDF software used, is supporting it (e.g., Adobe Acrobat or the PDF-viewer in Mozilla Firefox). This way, it can be downloaded and used wherever this might be useful. Since the implementation uses a standard functionality of the PDF specification it is agnostic to the tools used. Figure 3 shows an application in the Adobe Acrobat Reader with embedded semantic information as a JSON-LD-file in form of an attachment. The use of this approach would be the most feasible one for the DS and ToR documents as they are self-contained and could be easily represented in the document without any additional IT-system; they are interoperable and preserve the self-containment feature of the document.

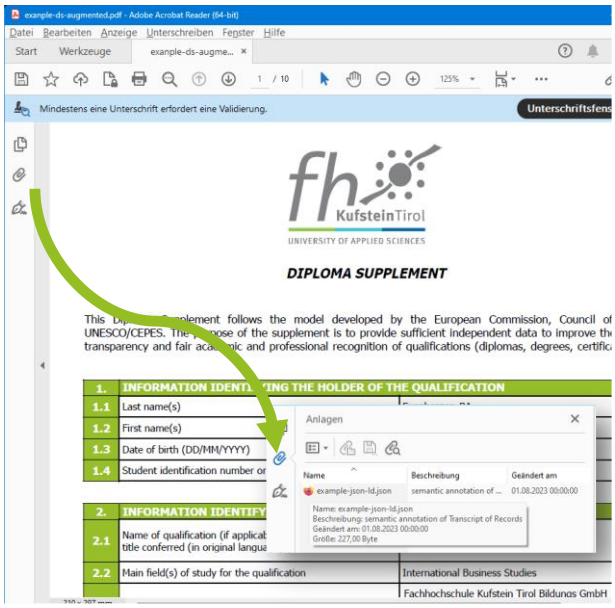

Figure 3. Screenshot of Adobe Acrobat as an example PDF application showing the embedded semantic information as JSON-LD data in a containerized representation.

While the representation of DS and ToR provide a good interfacing to external entities based on the concept of verifiable credentials (EVC) and a detailed domain model for the education domain (ELM), they are not suitable for a more detailed insight into the structure of the learning process that links the qualification objectives (c.f. Section 4.2 in the ToR document) to the individual learning paths that are usually fulfilled by lectures taken by the students. This currently hidden path could contribute to the transparency of the learning process, both during and after the student has completed her study. The ELM could be used to model the path and thus support the upper and the lower level of an academic program specification, as depicted in Figure 2.

It should be noted that the curricular structures are already crafted manually and usually encoded into a textual representation. This approach should help users to model the curricular structures by using an existing model (ELM) with the assistance of an attractive and lightweight user interface that hides the complexity of the semantic modelling by focusing on the instance level and only the needed concepts and properties for the application case. Hence the notion of Pre-Build Information Models, which remove the task of modelling from the user of the implementation. Currently the focus on the implementation is therefore on the visualization and use of the semantic models and less on automated concept extraction from existing textual curricular descriptions using text-mining technologies.

In order to visualize and use such a path though the curriculum the implementation provides two views that should be easy to understand and to navigate. It builds on the concept of Hierarchical Competence Maps (HCM), which are described in [20]. Figure 4 shows a detailed view on a single modular element (e.g., a module, a lecture or a teaching unit) with a few competences and their level of expertise,

 ISBN: 978-1-68558-108-4

visualizing the hierarchy with the metaphor of stacked cards, employing the concept of an HCM and using ELM as a data model for the semantic representation.

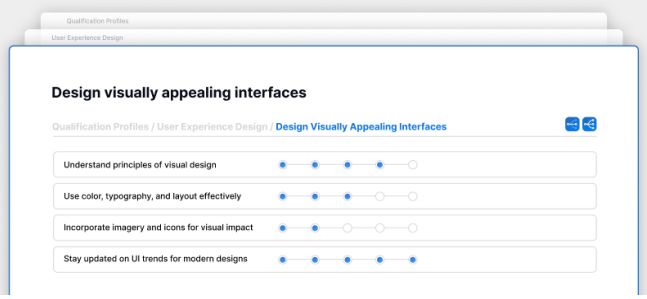

Figure 4. Screenshot of the detailed view of a modular element of an education program, using HCM and ELM.

The intention here is to focus on the immanent competences of the current element providing a detailed view on the learning outcome. As this view is lacking the overview of the connected modular elements this is provided by another view, shown in Figure 5 with the thread from the top-layer to the bottom layer highlighted. This overview should visualize the connection between different elements, helping the user to understand the relation and contributions between different parts of a curriculum. Two variations of the overview are implemented, one that is showing all linked data elements and another one that is emphasizing on a specific chain of competences from the (top-level) learning objective of the program to the learning goals of an individual lecture (at the most detailed level), using connected ELM concepts.

The user can switch between those two visualizations at any time and the implementation is carried out as a web-application that can be embedded in other web-applications, such as the website of a HEI or the e-Learning-system as needed.

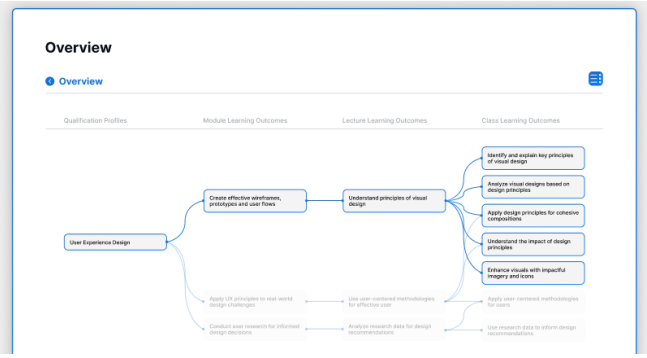

Figure 5. Screenshot of the overview of the connected learning modules.

Both implementations follow the concept of Semantic Specification Documents (SSD) that are self-contained and tool-agnostic [8], which are important properties to ensure an easy adoption in heterogeneous environments with a high variation of tools. PDF and web-based technologies are the document representation used, which are widely adopted and ELM is the backbone for the semantic representation, carrying the meaningful relations among the information elements.

V. Conclusion & Outlook

This contribution analyzed the current situation of the digital transformation within HEI with respect to the digital documentation and certification of learning outcomes (achievements). It demonstrated that a number of European standardization activities are already paving the way towards machine readable semantically enhanced specification documents with the European Learning Model (ELM) developed by the European Commission and with the Educational Verifiable Certificates (EVC) by the W3C. Similar activities can be found in other areas of the world, e.g., the Credential Transparency Description Language (CTDL) [21] and the OpenBadge Specification [22] in the US.

In Austria and many other European countries, the documentation of learning outcomes is already state of the art by supplying a Transcript of Records (ToR) and a Diploma Supplement (DS), which are often generated in an automated fashion, but still in the format of unstructured text. However, the aspect of explicit semantics and precise machine readability is not yet being used widely.

The case analysis and the prototype implementation demonstrated that the use of the standards is technically feasible without harming existing generation processes and that a number of stakeholders could benefit from self-contained, machine-understandable specification documents.

The next steps in this research efforts are a more extensive implementation for a number of course programs at the HEI of the author and the evaluation of the benefits for stakeholders for important use cases, such as the Recognition of Prior Learning (RPL) with different stakeholders. This will be the basis of an ongoing evaluation and improvement process that focusses on the ease of use and a beneficial user experience to ensure a good adoption rate to accelerate the digital transformation of HEI.

Acknowledgment

The author would like to thank the Tyrolean Science Fund (“Tiroler Wissenschaftsförderung”), which supported this research under grant number F.33280/6-2021 and the detailed and helpful comments of the reviewers of this contribution.

References

[1] European Commission, “The European Qualifications Framework (EQF) | Europass.” https://europa.eu/europass/en/europass-tools/european-qualifications-framework (accessed Aug. 01, 2023).

[2] S. Bohlinger, “Comparing Recognition of Prior Learning (RPL) across Countries,” in *Competence-based Vocational and Professional Education: Bridging the Worlds of Work and Education*, M. Mulder, Ed., in Technical and Vocational Education and Training: Issues, Concerns and Prospects. Cham: Springer International Publishing, 2017, pp. 589–606. doi: 10.1007/978-3-319-41713-4_27.

[3] European Commission, “Upcoming launch of the European Learning Model v3 | Europass.” https://europa.eu/europass/en/news/upcoming-launch-european-learning-model-v3 (accessed Jun. 15, 2023).

[4] “Verifiable Credentials Data Model 1.0.” https://www.w3.org/TR/vc-data-model/ (accessed Apr. 19, 2021).

 ISBN: 978-1-68558-108-4

[5] "Modeling Educational Verifiable Credentials." https://w3c-ccg.github.io/vc-ed-models/ (accessed Jun. 15, 2023).
[6] European Commission, "ECTS and Diploma Supplement Labels," *EACEA - European Commission*, Mar. 13, 2018. https://eacea.ec.europa.eu/sites/2007-2013/lifelong-learning-programme/ects-and-diploma-supplement-labels_en (accessed Apr. 09, 2021).
[7] European Commission, "Diploma Supplement | European Education Area." https://education.ec.europa.eu/education-levels/higher-education/inclusive-and-connected-higher-education/diploma-supplement (accessed Aug. 01, 2023).
[8] K. Böhm, "Towards Semantically Enriched Curricula as pre-built Information Spaces in Higher Education Institutions," in *ECKM 2021 22nd European Conference on Knowledge Management*, Academic Conferences limited, 2021, p. 71.
[9] "Seven Steps to a Solid Job Task Analysis," *EDSI*. https://www.edsi.com/blog/seven-steps-to-a-solid-job-task-analysis (accessed Aug. 01, 2023).
[10] BMBWF, "Electronic Signature." https://bmf.gv.at/en/topics/digitalisation/Digitised-Austria/Electronic-Signature.html (accessed Jun. 18, 2023).
[11] European Commission, "Diploma Supplement for stakeholders | Europass." https://europa.eu/europass/en/diploma-supplement-stakeholders (accessed Apr. 09, 2021).
[12] BMBWF, "Diploma Supplement." https://www.bmbwf.gv.at/en/Topics/Higher-education---universities/Recognition-of-qualifications/Diploma-Supplement.html (accessed Aug. 01, 2023).
[13] K. Böhm, "CloudEARTHi workshop on Agility and Semantic Structures for Scaffolding Academic Education, Part I and Part II." https://cloudearthi.com/asss-academic-education/ (accessed Sep. 05, 2023).
[14] L. Recalde, R. Navarrete, and L. Correa, *A Framework for data mining of structured semantic markup extracted from educational resources on University websites*. 2022. doi: 10.54941/ahfe1001745.
[15] Google, "Intro to How Structured Data Markup Works | Google Search Central | Documentation," *Google for Developers*. https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data (accessed Aug. 01, 2023).
[16] R. Montti, "Google On Which Structured Data it Prefers: JSON-LD or Microdata?," *Search Engine Journal*, Mar. 11, 2019. https://www.searchenginejournal.com/google-structured-data-preference/297479/ (accessed Aug. 01, 2023).
[17] P. Wyatt, "PDF 2.0 Application Note 002: Associated Files," Oct. 26, 2018. https://pdfa.org/resource/pdf-2-0-application-note-002-associated-files/ (accessed Jun. 22, 2023).
[18] "PDF-LIB · Create and modify PDF documents in any JavaScript environment." https://Hopding.github.io/ (accessed Jun. 22, 2023).
[19] W3C, "JSON-LD 1.1," 2020. https://www.w3.org/TR/json-ld11/#embedding-json-ld-in-html-documents (accessed Apr. 13, 2021).
[20] K. Böhm, *Agility and Semantic Structures to Scaffold Modern Academic Education Supporting the Digital Transformation in HEI*. 2022.
[21] "Credential Engine Registry | Credential Transparency Description Language Schema Metadata." https://credreg.net/ctdl/terms (accessed Jun. 28, 2023).
[22] "Open Badges Specification 3.0 Candidate Final Public." https://1edtech.github.io/openbadges-specification/ob_v3p0.html (accessed Jun. 19, 2023).

 ISBN: 978-1-68558-108-4