Supporting Learning Mobility With Student Data Harmonisation: A European Perspective

Thashmee Karunaratne and Efstratios Kontopoulos
Stockholm University, Sweden
thasmee@dsv.su.se
skontopo2009@gmail.com

Abstract: Digitalisation promotes online education, internationalisation and student mobility. Based on the Bologna process and the European higher education area, learning mobility has been successful under Erasmus and other similar initiatives. However, a key issue for students and universities is that a significant amount of time is spent on the manual labour involved in the process of applying to degree programs overseas. It is therefore essential for higher education institutions to better exploit the potential of technology and Web 2.0 to enable a secure exchange of evidence during application for degree programs and academic courses in foreign Higher Education Institutions, as well as applying for study grants and obtaining recognition for academic and other types of studies. Harmonisation of the student data is a key initial step for enabling such exchange. In this study, an approach to a secured exchange of education evidence is instrumented under the H2020 project Digital Europe for All (DE4A). Existing semantic standards for Web 2.0 applications, core vocabularies for public service data and semantic assets from existing best practices such as W3C, ISA2 core vocabularies, and Europass data model are used to curate data models that allow the exchange of a higher education diploma, secondary education diploma and information of special needs (disability, large family), which is required by students when requesting study grants (waive of tuition fees). The semantic interoperability agreements are established cross-border through these data models called canonical evidences. The canonical evidences are tested with the national data services of three countries, Portugal, Slovenia, and Spain. The final data models are implemented in XML Schema format that could be used by any educational organisation intending to use trusted public service databases within Europe to automatically retrieve information on students’ degrees. The validity of the canonical evidences is tested on two pilot occasions within the DE4A project. The outcome of this study summarises the procedural requirements for evidences when applying for a higher degree program and seeking grants. Furthermore, it resulted in verified canonical evidence data models that fulfil the procedural requirements for applying for studying abroad.

Keywords: learning mobility, semantic interoperability, data harmonisation, higher education, cross-border

1. Introduction

Student internal (within the resident country) and external (outside the resident county) mobility has become pervasive under flexible education provisions in the digital age (Choudaha and Chang 2012; Joint Research Centre (European Commission) et al. 2015; Rivza and Teichler 2007). Digitalisation of the education sector paves the way forward for universities to introduce student exchange programs and mix-and-match education allowing personalisation and individualised education (Borghetti and Beaven 2017; Hefler and Steinheimer 2020). For instance (Rivza and Teichler 2007) describe four varieties of student mobility; 1) In search of unique specialisation programs offered by reputed universities and scholars, 2) Program/s in a neighbouring country to experience a change with relatively low risk of cultural or language differences, 3) Taking a term or two of a program in another country that has a similar level of education/reputation, and, 4) University motives for attracting international students for diversity and financial reasons. The trends of students expanding their education within and among their higher education programs can only be increased in the future (Choudaha and Chang 2012; Hefler and Steinheimer 2020; Joint Research Centre (European Commission) et al. 2015). To deal with this emerging trend and its consequences, and to allow smoother and more efficient student mobilisation, many global and regional programs have been introduced in the past. For example, EU’s ERASMUS and ERASMUS+ programs support students financially and academically to study abroad by providing education exchange platforms (‘Home | Erasmus+’ n.d.). This possibility of mixing and matching education programs and courses from different universities in different countries is enhanced by standardising the educational qualifications by the Bologna process, which is “a process aimed at ensuring comparability in the standards and quality of higher-education qualifications” (‘Bologna Process - University of Bologna’ n.d.). This initiative has successfully harmonised the academic qualifications at least across Europe. The European Higher Education Area (EHEA) brings together “49 countries with different political, cultural and academic traditions, which, step by step during the last twenty years, built an area implementing a common set of commitments: structural reforms and shared tools”, with educational reforms that support the higher education systems, increasing student and staff mobility (‘Higher Education | European Education Area’ n.d.). These initiatives justify the EU’s growth strategy for the coming decade that highlights higher education as a key policy area where collaboration
Thashmee Karunaratne and Efstratios Kontopoulos

between the EU and the Member States can deliver positive results for jobs and economic development through quality education (‘Bologna Process - University of Bologna’ n.d.; ‘Higher Education | European Education Area’ n.d.).

Typical higher education provision has two stages, application to programs by students and selection process by the university. From this, digitalisation of the application process is something that is considered under the digital Europe programs (‘Higher Education | European Education Area’ n.d.). The reason is, however, that a major issue for students and universities in this regard is that a significant amount of time and cost is spent on the manual labour involved in the process of applying to degree programs overseas. It is a key policy issue for the Member States and higher education institutions to “better exploit the potential of ICTs to enable secure exchange of evidence for the registration in foreign Higher Education Institutions”, as well as applying for study grants and obtaining recognition for academic and other types of studies.

A key requirement for cross-border education services (registration for foreign degree programs) is the interoperability of the services led by the harmonisation of student information (data) (DE4A Consortium 2022). The process of cross-border information exchange entails an absolute necessity, which is the semantic agreements of the data to be exchanged across-borders that each EU member state can agree on in connection with the welcoming of foreign students in their countries. It is therefore essential for higher education institutions to better exploit the potential of technology and Web 2.0 to enable secure exchange of evidence to enable their current and former students to register for education activities in foreign Higher Education Institutions, as well as to apply for study grants and obtain recognition for academic and other types of studies. Harmonization of the student data (the semantics) is a key initial step for the exchange. With the aim to enable higher educational institutions to offer the possibility of their students to apply for degree programs in foreign public higher educational institutions, this study instruments an artefact addressing the following research questions: 1) What are the procedural requirements for degree programs when students apply from a different country of residence in the EU?, and, 2) How can harmonised and standardised data models enabling the secure student information exchange across borders be developed based on the existing semantic standards for Web 2.0 applications (W3C n.d.), core vocabularies for public service data (‘Homepage | ISA2’ n.d.) and semantic assets from existing best practices such as the Europass data model (Europass n.d.)? In answering the research questions, the partnerships under a European commission Horizon 2020 project named Digital Europe for All (DE4A) (DE4A 2020) are utilized for increasing the impact of the research outcomes and ensured validation and applicability of the developed artefacts.

The rest of the article is organized as follows. The next section provides the context and background of this research work. Section 3 briefly about the methodological approach followed. Section 4 illustrates how the data models are developed based on the elicited requirements, and finally, the concluding remarks section summarizes the funding of this work and some promising further work.

2. Background

Interoperability of education services depends heavily on the similarity of education data both in usage and in concept or meaning (Berberidos et al. 2020). In other words, respective public authorities that are responsible for education offering (the foreign university student wish to register for) on one end, as well as credentialization of education (the university student received pre-requisite qualifications) on the other end, should speak the same language for effective information (data) exchange. The prerequisites for registering for a degree program may depend on the subject area, university, country and the type of the qualifications, which may be granular questions to be overseen by a qualified administrator in the university. For example, when referencing data of student’s domicile, Country A, where the student resides, may use a concept of “address” as a “structured object” while Country B, where the student is going to apply for a university, may just allow presenting the address as a “text field”. When such different formats of data are being exchanged between countries A and B, a prior understanding of the data type is extremely important, viz, the data should be interoperable. Such an understanding – mapping – of data is demanding under both technical and legal grounds of educational services provided through public procedures. Data harmonization, i.e., one-to-one mapping of data at both sending and receiving countries has many bottlenecks, of which, the agreement between the two authorities (in two countries in the cross-border case), is the narrowest among them (DE4A Consortium 2022). When referring to interoperability, the first step is to identify, which pieces of data are required by the procedure for obtaining foreign higher education. According to the single digital gateway regulation (EC n.d.), this data set should be the
minimal information needed to prove the previous education and other data required to execute the said public procedure. In this article, such a data set is defined as “A collection of data, published or curated by a single agent (Data Providing authority), and available for access in one or more representations.” (DE4A Consortium 2022).

2.1 Canonical evidence for interoperability

In the European Union Horizon 2020 project of Digital Europe for all (DE4A 2020), interoperability is achieved through an “Evidence, which is a Dataset required to prove a Requirement or Criterion” (DE4A Consortium 2022). In the public administration context, evidence is the information to legally prove that procedural requirements are met. This could be “any document or data, including text or sound, visual or audio-visual recording, irrespective of the medium used, required by a competent authority to prove facts or compliance with procedural requirements referred to Article 2.2.b (SDGR)” (Berberidos et al. 2020). In the traditional form of registering for degree programs at universities, the students manually produce the required documents (evidence) when they fill in the application forms. Digitalised public service intends to reduce the administrative burden of this inefficient process and increase the efficiency by providing all the evidence required by a procedure automatically from the accredited public agent, lawfully issued, even if it is cross-border. In this context, it involves electronic evidence which is “Lawfully issued evidence by competent authorities ranging from completely unstructured formats (such as pdf or picture formats) to structured databases” (Berberidos et al. 2020).

As mentioned before, the agreement between the issuing and the receiving authorities on which data are required and can be provided by the agencies involved is the primary concern, which needs to be already solved prior to the execution of the procedure and enabling the online registration service. This agreement is established in the DE4A context by the concept of canonical evidence. By definition, canonical evidence is a “Structured data model that includes a common set of attributes associated with the evidence type that can be provided by the corresponding lawfully issued evidences”. Hence it is essential to define canonical evidences for every piece of information required by public service procedures if they intend to provide the public service in Europe under the Single Digital Gateway Regulation (SDGR) (EC n.d.).

2.2 Web 2.0 for student data exchange

Defining the concepts and meaning for pieces of information is a major part of semantics. With the evolution of the Semantic Web, the technologies around it were gradually developed and reached a hype with the establishment of the World Wide Web Consortium - W3C (W3C n.d.), aiming at providing harmonization in the construction of the World Wide Web services. To date, the baseline for any resource, service, or action that is hosted through the web follows the W3C guidelines, definitions, vocabularies, and other respective technical resources required to set up any web service. Under the EU initiatives for digitalization of public services, and the once-only principle that entails “The public administrations should ensure that citizens and business can supply the same information only once to a public administration and administrations should be able to retrieve and share this data to serve the user, in accordance with data protection rules.” (Berberidos et al. 2020), an administrative web portal service has become the main component (e-Delivery), that “helps public administrations to exchange electronic data and documents with other public administrations, businesses and citizens, in an interoperable, secure, reliable and trusted way”.

According to SDGR, it is every EU member state’s responsibility to contribute to a single digital gateway by establishing their procedure portals and thereby connecting their public services with the rest of the EU. In dealing with such a requirement, semantic interoperability of the evidence exchange system is a must. W3C provides the baseline for this semantic interoperability. Through many initiatives, the EU has also been trying to establish the semantic interoperability for the once only technical system, such as ISA and ISA² (“Homepage | ISA” n.d.), to help its member states create this “common language for the European public administrations, so that information can be exchanged easily” (DE4A Consortium 2022). For the academic domain, the most interesting development in this line would be the Europass data model (Europass n.d.). The Europass Digital Credentials Infrastructure (EDCI) is a set of standards, services and software that permits institutions to issue digital, tamper-proof qualifications and other learning credentials within the European Education Area. This data model is an extension of the W3C Verifiable Credentials Data Model (DE4A Consortium 2020a). EDCI streamlines the issuing, viewing and automatic verification of credentials and benefits a wide range of stakeholders, including individuals, universities, vocational training providers and employers. The infrastructure can guarantee a common understanding of qualifications and types of certifications across and beyond the European Union.
Thashmee Karunaratne and Efstratios Kontopoulos

The infrastructure can also enhance transparency and portability of qualifications and skills between countries, reduce the administrative burden for citizens, learning providers and employers, make tampering and credential fraud easily identifiable, empower people to own and control their own credentials, and contribute towards digitisation of government processes. The Europass data model can constitute the basis for describing the concepts of higher and secondary education while providing the necessary credentials for certifying the evidences related to the domain of education. It should be noted that this paper describes semantic models for lawfully providing evidences for proving the educational qualifications at secondary and tertiary levels. Harmonising the pre-requisites for each and every higher education program in the universities in Europe is hence out of the scope of the EU Digital Europe program and this study.

3. Design and development methodology

The research questions urge a deeper investigation of the requirements for enrolment of students in a university in a foreign country, that complies with both the business situation and the legal agreement. Hence, a case is carefully chosen to allow specified scenarios to be used to study the case. A design science approach (Peffers et al. 2007) is followed as the overarching methodology of this research, which binds each of the steps in the artefact development process, i.e., the steps of the stages of the scenarios in the use case.

3.1 Use cases

In this study the use cases (UC1 &2) that are interconnected in applying for studying abroad were considered. These use cases were piloted in Portugal, Slovenia, and Spain. The pilots aimed to “prove the optimal process/procedure for students from the three participating Member States for registration to higher education and eventually applying for a student grant as well as for studies recognition” (DE4A Consortium 2020a).

UC1 - Application to Public Higher Education focuses on the Administrative Procedure of submitting an initial application for admission to a public tertiary education institution. This UC counts on the facilitation of electronic and physical mobility of (Erasmus/Erasmus+) students between Higher Education Institutions across Europe (’Home | Erasmus+’ n.d.). The goal is to enable foreign students to get access, in a digital-by-default and once-only manner (Piswanger, Helger, and John 2019), with secure and GDPR-compliant consent-based exchange of electronic evidence, to register, enrol, or admit fully online and remotely using a trusted cross-border authentication such as CED eID (EUR-Lex 2014). The administrative procedures are traditionally burdensome in terms of time and human resources and students can be discouraged to apply by the complexity of procedures (DE4A Consortium 2020a). The solution for this use case requires the specification of semantically interoperable electronic evidence information structures including information on ECTS credits and exchange of simple academic attributes as needed by the participating universities in the role of the data providers, e.g., considering Europass (Europass n.d.) data model standard, to maximize replicability and scalability of results to other European countries, with due consideration of existing national workflows (involving directly the Higher Education Institutions and, where relevant, also regional or national Higher Education IT systems relevant for the enrolment of students), as well as on-going initiatives seeking interoperability or inter-linking of government and education identities (DE4A Consortium 2020a).

UC2 - Applying for Study Grant: Oftentimes students apply for a study grant from a public institution in conjunction with the application for degree programs. If the applicant wishes to obtain a study grant, the admissions office/competent authority would likely ask for proof of their identity for cross-border remote authentication, proof of their grades, and their financial situation. If, having checked this information, the admissions office or the responsible competent authority finds it suitable, they might issue a certificate attesting that the person qualifies for a study grant (DE4A Consortium 2020a).

3.2 The design science research methodology for modelling the process

Based on the underlying need in the three UCs, a systematic process is required for the development of the required artefacts, which, as mentioned above, could be found in the premise of the DSRM process (Peffers et al. 2007). Figure 1 illustrates the design science application in this study.

Accordingly, the problem explication and the requirement elicitation are conducted together with the education stakeholders, i.e., university administration (student registration authorities), policy makers, semantic experts, legal advisers, officials in public service digitalization bureaus of selected countries, and students. The initial
Thashmee Karunaratne and Efstratios Kontopoulos

requirement set drawn is co-created together with the stakeholders. The designed artefacts are regularly tested with the participants of the use cases described in section 3.1. These tests are conducted to canonical evidence models are evaluated within the public administration (Universities) of the said Member states using real students.

Figure 1: Design science application on a case study for artefact development

3.3 Ethical implications of the design and evaluation of the artefacts

The cross-border application for university studies requires the exchange of data that may include sensitive information as well (DE4A Consortium 2020b). However, in this research, we focus on developing only the metadata models (canonical evidence) that can be agreed upon between the university student registering themselves and the responsible competent authority of students’ prior education certificates. However, the privacy of the stakeholders involved in the DSRM process is preserved according to the General Data Protection Regulation – GDPR (‘Data Protection Impact Assessment (DPIA)’ 2018).

4. Results and design outcomes

Based on the above, this section presents the implemented canonical evidence schemas for representing (a) a higher education diploma, (b) a secondary education diploma, and (c) information on special needs (disability, large family), which are required when requesting study grants (waive of tuition fees). The functional requirements underlying the implemented schemas were suggested by the DE4A pilot partners during regular consulting sessions in a co-creative and “agile” fashion that involved frequent iterations. An overarching horizontal requirement was that all canonical evidence models should rely on existing data exchange standards and should reuse existing models and controlled lists (i.e., code lists) as much as possible. All the schemas are available at the respective DE4A GitHub repository: https://github.com/de4a-wp3/study_abroad.

4.1 Higher education diploma evidence

This type of evidence proves that an individual has acquired a higher education diploma and encompasses all pertinent aspects for submitting an initial application for admission to a public tertiary education institution. This model largely adopted EDCI (Europass n.d.) for representing education-related concepts, and CBC (Common Basic Components) of the OASIS specification for the Universal Business Language v2.0 (UBL 2019) for representing temporal aspects, like, e.g., dates and durations. A diagrammatic overview of the higher education diploma model is illustrated in Figure 2 while Table 1 gives a more detailed specification of the attributes.
Figure 2: Higher education diploma evidence diagrammatic overview

Table 1: Higher education diploma evidence attributes specification

<table>
<thead>
<tr>
<th>Field</th>
<th>Data type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>HolderOfAchievement</td>
<td>edci:PersonType</td>
<td>Person that has obtained the academic title or degree.</td>
</tr>
<tr>
<td>Country</td>
<td>edci:MDRCountryCodeEnumType</td>
<td>Country where the study programme was completed by the student. Assumes values from the Country Authority Table codelist (EU publication office n.d.).</td>
</tr>
<tr>
<td>InstitutionName</td>
<td>edci:TextType</td>
<td>The name of the higher education institution where the student obtained the degree.</td>
</tr>
<tr>
<td>MainFieldOfStudy</td>
<td>edci:IscedFOetCodeType</td>
<td>Field of finished higher education. Assumes values from the ISCED-F controlled list (EU publication office n.d.)</td>
</tr>
<tr>
<td>StudyProgramme</td>
<td>edci:TextType</td>
<td>Name of a study programme that the student finished at the higher education institution in order to obtain the degree.</td>
</tr>
<tr>
<td>Degree</td>
<td>edci:TextType</td>
<td>An academic title or degree obtained by the student and proven by this diploma or certificate.</td>
</tr>
<tr>
<td>DateOfIssue</td>
<td>cbc:DateType</td>
<td>Date of issue of the certificate or diploma.</td>
</tr>
<tr>
<td>PlaceOfIssue</td>
<td>edci:LocationType</td>
<td>Place of issue (location) of the certificate or diploma.</td>
</tr>
<tr>
<td>DurationOfEducation</td>
<td>xsd:duration</td>
<td>Official duration of education.</td>
</tr>
<tr>
<td>ModeOfStudy</td>
<td>edci:LearningScheduleTypeEnumType</td>
<td>Mode of study, e.g., full-time, part-time etc. Assumes values from the Europass Standard List of Learning Schedule Types (EU publication office n.d.)</td>
</tr>
<tr>
<td>Scope</td>
<td>edci:ECTSCreditPointsType</td>
<td>The official workload of the study programme in ECTS credit points.</td>
</tr>
<tr>
<td>AverageGrade</td>
<td>Custom type</td>
<td>The average grade awarded. Extension to edci:TextScoreType, allowing representing the actual score grade, the grading scheme, along with a short textual explanation.</td>
</tr>
</tbody>
</table>

4.2 Secondary education completion evidence

As its name implies, the second type of evidence proves that an individual has completed their secondary education. Similar to the Higher Education Diploma Evidence, this type also largely relies on EDCI and CBC/OASIS, as also illustrated in the diagrammatic overview (Figure 3) and the attributes specification (Table 2).
Table 2: Secondary education completion evidence attributes specification

<table>
<thead>
<tr>
<th>Field</th>
<th>Data type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>edci:MDRCountryCodeEnumType</td>
<td>Country of completed secondary education. Assumes values from the Country Authority Table codelist.</td>
</tr>
<tr>
<td>Degree</td>
<td>edci:TextType</td>
<td>Degree previously obtained (e.g. General upper secondary education).</td>
</tr>
<tr>
<td>NameOfSchool</td>
<td>edci:TextType</td>
<td>Name of a secondary school that the person finished.</td>
</tr>
<tr>
<td>NameOfProgram</td>
<td>edci:TextType</td>
<td>Name of a secondary school programme that a person successfully finished (e.g. General secondary programme).</td>
</tr>
<tr>
<td>IssuingDate</td>
<td>cbc:DateType</td>
<td>Issuing date of the accomplishment.</td>
</tr>
<tr>
<td>Grade</td>
<td>Custom type</td>
<td>Mark indicating a degree of accomplishment. Custom type, same as AverageGrade above.</td>
</tr>
</tbody>
</table>

4.3 Non-Academic information evidence

In order to represent non-academic information (e.g., for the purposes of awarding a scholarship or grant), we introduce two canonical evidence types, DisabilityCertificateType and LargeFamilyCertificateType, which correspond to the respective certificates for proving a disability and a large family, respectively. As seen in Figure 4 and Table 3, both certificate types share a similar structure.

Table 3: Disability and large family certificate attribute specification

<table>
<thead>
<tr>
<th>Field</th>
<th>Data type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CertificateID (both types)</td>
<td>xsd:string</td>
<td>ID of the certificate.</td>
</tr>
<tr>
<td>EffectiveDate (Disability)</td>
<td>cbc:DateType</td>
<td>Date from which the disability certificate is valid.</td>
</tr>
<tr>
<td>RevisionDate (Disability)</td>
<td>cbc:DateType</td>
<td>Date on which the disability certificate will be reviewed.</td>
</tr>
<tr>
<td>DisabilityPercentage (Disability)</td>
<td>xsd:int</td>
<td>Specifies the degree of disability of the beneficiary. Indicated as a percentage.</td>
</tr>
<tr>
<td>ValidCertificate (LargeFamily)</td>
<td>xsd:boolean</td>
<td>True/false depending on whether the large family certificate is valid or not.</td>
</tr>
<tr>
<td>IssuingDate (LargeFamily)</td>
<td>cbc:DateType</td>
<td>Date of issue of the large family certificate.</td>
</tr>
<tr>
<td>ExpiryDate (LargeFamily)</td>
<td>cbc:DateType</td>
<td>Expiry date of the large family certificate.</td>
</tr>
<tr>
<td>NumberOfChildren (LargeFamily)</td>
<td>xsd:int</td>
<td>Indicates the number of children the holder of the large family title has.</td>
</tr>
</tbody>
</table>
5. Discussion and concluding remarks

The verified data models (canonical evidences) created in this study can serve as semantic models when exchanging the evidence for student’s higher and secondary education qualifications, which is needed to be provided to the academic institution during the process of applying for studies and study grants. Through the pilots set up with three countries, the definitions of the concepts in the data models, their meaning and viability are tested. These final data models (canonical evidences), hence, can be used by any public institution when they implement procedure portals for exchanging electronic data cross-border, allowing students from other countries to apply remotely. Such a provision enables universities to open up their education efficiently and effectively outside the borders of their country. Such an approach will also help minimize the fraud and possibility to present wrong information, since the evidence are being received not from a citizen but from an accredited trusted public agency. This will also allow autonomy of the process since the whole process is electronic and in real time.

This study achieved the goal of creating harmonized data models (canonical evidences) that are technically, legally and semantically verified together with the respective competent authorities in public (education) service. However, the study is limited to developing only the canonical evidence (the information needed for cross-border services enhancing student mobility). The metadata and process models for information exchange, which is the other absolute necessity, do not fall under the scope of this work, but would be the essential next step towards achieving a complete cross-border information exchange.

Acknowledgements

This study is conducted as a part of the Digital Europe for all (DE4A) project, which has received funding under the European Union’s Horizon 2020 research and innovation programme, under G.A. No 870635, in the area of SOCIETAL CHALLENGES - Europe In A Changing World - Inclusive, Innovative And Reflective Societies. The authors would like to acknowledge the contributions from the project consortium, with special mention of the work package 3 - Semantic interoperability solutions.

References


