Interoperability of educational data demands standards

Tor Fridell[1], Geir Vangen[2], Janina Mincer-Daszkiewicz[3], Jan-Joost Norder[4], Jukka Kohtanen[5], Igor Drvodelić[6], Guido Bacharach[7]

[1]the Ladok consortium in Sweden
[2]Sikt – Norwegian Agency for Shared Services in Education and Research, Norway
[3]University of Warsaw, Poland
[4]DUO, Dienst Uitvoering Onderwijs, the Netherlands
[5]CSC – IT Center for Science Ltd, Finland
[6]AZVO – Agency for Science and Higher Education, Croatia
[7]Stiftung für Hochschulzulassung, Germany

tor.fridell@liu.se, geir.vangen@sikt.no, jmd@mimuw.edu.pl, janjoost.norder@duo.nl, jukka.kohtanen@csc.fi, igor.drvodelic@azvo.hr, guido.bacharach@freenet.de

Abstract

Student mobility is a keystone in the building of the future Europe but mobility requires transfer of educational data. This interoperability of educational data demands standards. There are basic regulations, like ECTS (European Credit Transfer and Accumulation System), which defines grading rules and transfer of credits, and ISCED (International Standard Classification of Education), which defines educational levels and subjects. There are also EU standards for data transfer, like EuroLMAI (EN 15981) and MLO (EN 15982). However, the implementation of the standards differs slightly and in practice there is more than one standard. The possible implications for the end user are that there will be "silos" with different systems not being able to communicate and there will be no mobility of data.

This can and should be solved by cooperation with the long-term goal to unite standards and until then by building converters between them. This paper describes regulations, different standards, ongoing work and ways to go forward.

1 Introduction

It is undeniable that standards are needed to make the world go around. How could we communicate and interact if we didn't have a common standard for time measurement, mass and length? There are lots of standards that we don't even think about that make life easier.

It is not true that exactly everyone needs to use the same standard. In the case of what side of the road you drive there exist two standards and it seems that if a big enough user base can be found for each standard they can coexist without problems. The same goes for mobile telephones, both Apple and Android have enough users to continue developing their own standard.

What is the case for educational data then? There have been European standards for learning elements for quite some time, some published already in 2011 but there has not been a great general need until recently. The EU-focus on cross-border education, life-long learning, border-less Europe, micro credentials and such have given rise to many projects and initiatives related to sharing educational data. Examples are Erasmus Without Paper, EMREX, Europass but many more exist.

1.1 Historical flashback

It is interesting to look back at different areas and how standards have affected the industry. One example would be the networks for mobile communication where initially no standards were in place and there were several siloed systems for each network supplier. Since the technical development is so rapid and every major leap requires total exchange of essential equipment there has been movement to common standards and more or less worldwide possibility to use mobile devices. The key point here is that each development step requires completely new technology, but the benefits (in increased capability) are so big that it overshadows the cost. The marginal cost for adhering to a new standard is minimal and the perceived use is substantial.

Another example could be delivery of electricity. Also, here it was siloed systems in the beginning, but the standards have not converged completely. There are several standards regarding both voltage and type of outlets. Here we have a situation where legacy matters. The investment in infrastructure is significant, the infrastructure is long-lived and the added benefit for changing to adhere to a standard is not that clearly visible.

In many cases the companies in the industry try to bridge the gaps. Regarding electricity there are travel adapters so you can plug in your stuff and many "travel oriented" devices (like razors) can manage different voltages. The same goes for mobile phones which are built to work with the frequencies used in different countries.

1.2 Conclusion

Standards are indeed important, but it is not imperative that there is only one single standard. Adaptions can be made to bridge the gaps. Seen over time standards are wild blooming in the beginning of new technology advances but then often converge after some time.

2 Standards in the education area

The education area has standards for several aspects of describing educational data. First there are the basic regulations of what a course or degree is and how to compare educational elements originating from different technical systems. Secondly there are standards for the data formats to be used, that is what elements can be included, what data type are they etc. Lastly there are standards for the transfer of data which includes for instance authentication and transfer protocol.

2.1 Basic regulations

ECTS, European Credit Transfer and Accumulation System [1], is not a standard in itself but a tool of the European Higher Education Area (EHEA) for making studies and courses more transparent and thus helping to enhance the quality of higher education. ECTS is the basis for all transfer of educational data since it defines words like credit (what workload should one credit be), courses, programs, levels and diplomas. It has resulted in the ECTS Users guide which has been adopted by Ministers for Higher Education of the European Higher Education Area in 2015 at the Yerevan ministerial conference. It is therefore the official Guide for the use of ECTS.

I**SCED 2011** [2] is a standard maintained by UNESCO describing the level of education ranging from K9 to higher education with levels 0 to 8. It has sublevels but the most commonly used are the levels where 6 corresponds to bachelor studies, 7 to master studies and 8 to doctoral studies, or first, second and third cycle as described in ECTS terms.

I**SCED-F 2013** is a second standard from UNESCO that defines subject matters so that all education elements can be compared and assessed easily. There are main areas like Natural sciences, Arts and Humanities or Business, Administration and Law. Under each main area there are sub categories like Biology, Statistics or Database and network design and administration.

ESCO [3] is the multilingual classification of European Skills, Competences, Qualifications and Occupations. The ESCO classification identifies and categorizes skills, competences, qualifications and occupations relevant for the EU labor market and education and training. It systematically shows the relationships between the different concepts.

2.2 Standards for data format

There are a few different standards regarding data format. Within EU the two standards, MLO and ELM, are the core norm.

European Learner Mobility – Achievement information (EuroLMAI, EN 15981). This European Standard specifies a model for the recording and exchange of learner achievement information among student management information systems, as well as the aggregation of information by third party suppliers. The scope of the standard is restricted to the definition of the electronic representation of official, institutionally attested achievement information for learners engaged in formal learning processes, in order to facilitate its recording and subsequent exchange within the European education area.

Metadata for Learning Objects – Advertising (MLO, EN 15982). This European Standard specifies the characteristics of electronic representation of Learning Opportunities in order to facilitate their advertising and subsequent discovery by prospective learners. The model specifies three resources about which metadata can be stored to facilitate advertising of Learning Opportunities: a) the Learning Opportunity Provider; b) the Learning Opportunity Specification; and c) the Learning Opportunity Instance. This European Standard specifies the characteristics of relations between the three resources and recommends a core set of metadata for each.

PESC, Postsecondary Electronics Standard Council [4] is mostly North American organization with standards for admission, college transcript, course catalog and e-portfolio.

IMS Global [5] is another organization based in North America that issues standards for many entities, like learning objects e-portfolio, micro credentials, open badges etc. The scope is world-wide and since products using these standards are in operation in European countries the standard affects us.

2.3 Standards for transfer of data

It is more difficult to describe standards for transfer of data. Usually each system has its own implementation. To take Erasmus Without Paper (EWP) as an example the communication is done in a way that an information requester goes to a discovery service and finds a data provider. The communication is then done peer to peer, according to specifications in APIs, with little user interaction. This method is used only in EWP and any system in need of data must follow the exact procedure.

One of a few initiatives that tries to standardize the data transfer is the European Blockchain Services Infrastructure (EBSI) [6] based on blockchain paradigm.

3 Implementations of data formats

3.1 ELMO

ELMO [7] is a data format for the exchange of (education) result information. ELMO is an implementation of the European (CEN) standards ELM-AI (European Learner Mobility – Achievement Information, EN 15981) and MLO (Metadata for Learning Objects, EN 15982). It is an XML based format and public domain.

ELMO is built up of three basic elements; Learner, Issuer and Learning Opportunity Specification. The following elements are included in the latest version:

- Diploma
- Diploma Supplement
- Transcript of Records
- EMREX transcript
- Micro Credential
- Letter of Nomination
- Certificate of Training
- Learning Agreement

ELMO has been used in production by the EMREX [8] network since 2017 for international exchanges of educational information and for some members also extensively for exchanges within the country. Apart from national providers of Student Information Systems (SIS) also some companies providing SIS have included support for EMREX, for instance HISinOne. About 50 000 – 100 000 exchanges are done yearly in the EMREX network, with a strong rising trend.

ELMO is also used by the Erasmus Without Paper network (EWP) [9] with a registry in production and plans for rollout 2022-2023 where after that period ALL Erasmus+ student exchanges must be transferred via the network, as a prerequisite to achieve Erasmus charter. Even if the actual data exchanges are not in place yet the work for adapting software has been in progress for a few years and the software providers are heavily committed to completing the development. The registry contains over 2,400 higher education institutions so the user base will soon be vast.

The European Blockchain Services Infrastructure (EBSI) has announced that they plan to use ELMO as data format. In addition to the use of diplomas and other verifiable credentials, there is a need for students to be able to access their current student status and grades (transcript of records)

through the EBSI SSI (Self Sovereign Identity) wallet, in order to be able to have a digital record even before graduation, which can be instantly verified using the same principles. For these needs, ELMO has proven to be the best solution. Given that most higher education institutions will already use ELMO due to EWP, as mentioned above, the additional service they can provide to their students through EBSI should not be technically demanding for higher education institutions and could be very useful for students, especially if they will continue to use the EBSI SSI-compliant wallet even after graduation.

3.2 ELM

European Learning Model, ELM, is used by Europass [10]. The Europass Learning Model aims to capture the results of any non-formal and formal learning across Europe, as well as the validation of non-formal and informal learning. It is designed to provide a single format to describe certificates of attendance, examination results, degrees and diplomas, diploma supplements, professional certifications, employer recommendations and any other kind of claims that are related to learning. The ELM specification is available in GitHub [11].

It has to a large extent the exact same attributes as ELMO but since ELM has a broader scope and includes entities not related to educational data, such as skills and qualifications, it has also other attributes. In 2020 Europass published a model for digital credentials, the EDCI (Europass Digital Credentials Infrastructure) format but the name was later changed to ELM.

The EU project EDSSI [12] has done a comparison between ELMO and ELM. The educational part of ELM corresponds reasonably to the ELMO format with the same basic elements but with different names:

ELMO	**ELM (EDCI)**
Learner	credentialSubject
Issuer	<agentReferences><organization>
learningOpportunitySpecification	<achievements> <learningAchievement>

3.3 Other standards

There are of course also other standards related to educational data. These are not described further in this document. The purpose is not to give a detailed description of every possible standard in the area. Instead, the purpose is to examine the foundations for standards in the area and compare the most commonly used ones.

4 Coexistence of formats

Every organization that develops or adheres to a standard has its own use case. Therefore, one cannot assume that there can be one single standard for everything. Different systems have different purposes and with that come also demands on the data formats. The needs may be identical for some parts, for instance transcript of records which is a well defined object, but may then differ substantially, for instance if you include non-formal education elements defined in a completely other way. Each application must select the data formats and other standards that fits the needs best.

In the comparison report from EDSSI it was concluded that “To build a general converter from ELMO to EDCI should be doable. It is of course not trivial and a number of examples should be available to do this work. The most important items we feel missing in EDCI are academic term and status for each credential. Status is describing uniquely if a result is passed or failed. The other items which are not part of EDCI are of some importance for doing recognition, in particular the information about grade distribution and engagement hours.”

4.1 Example: IDunion

As part of the German IDunion project [13], research assistants and students work in multiple working packages in the context of Self Sovereign Identity (SSI) solutions. The work is done under supervision of Patrick Herbke, from Technische Universität Berlin.

In the education working package it is planned to make existing ELMO documents usable in SSI infrastructures as it provides functionality that the current ELMO structure does not intend to cover. In the SSI context, all documents are stored by the students themselves in full sovereignty and can be verified and combined with proper authentication.

Currently, SSI networks use JSON(-LD) data exchange formats exclusively to exchange verifiable credentials, attestations, and presentations. The need for educational credentials leads to the development of an ELMO to JSON(-LD) converter to lower the barrier of importing the existing and newly issued diplomas and certificates into the SSI ecosystem.

Furthermore, the IDunion converter considers different ELMO schemas for educational certificates. Schemas describe the profile, including data structures and semantic definitions, of the data used in EMREX. The IDunion ELMO converter contains a schema mapping from an existing ELMO schema to a recently published education schema by EBSI. The converter automatically adapts the EBSI schema based on a given ELMO certificate (transcript of records, bachelor/master certificate or upper secondary school certificate) and returns JSON-LD formatted content.

4.2 Discussion - general approach vs. special approach

A major problem on the way to find a common data standard for education data is the need to find a common semantic basic understanding on this way as well. This is particularly evident in an initiative underway since the end of 2021 to convert the digital upper secondary education certificates generated in the German DIGIZ NRW project with data in an ELMO structure to the ELM data structure as part of the EBSI program. In this attempt, it becomes clear that without a semantic understanding of such an upper secondary school diploma, such a transfer into the ELM schema is not possible either from the ELMO schema or from the original diploma.

If we assume that data structures such as ELMO or ELM should generally represent the data of educational documents and degrees (and even more comprehensively the SDG structure currently under development, which intends to include not only educational data), the question is how such a thus also comprehensive semantic basic understanding can be achieved in a reasonable time with a reasonable effort valid throughout Europe. Wouldn't it make more sense to abandon this approach of a generally valid data standard for educational data in favor of a more specialized approach of different, but Europe-wide valid data structures that can always be interpreted in a comprehensible way by the recipient?

For the Transcript-of-Records (and other formats of so-called "Living Transcripts", which represent a current state of education and results), the ELMO format has been standardized productively for years and to a large extent Europe-wide. The idea of the "special approach" would now be to create a separate data structure (which may be based on ELMO, ELM or SDG or also other "standards") for other types of evidence per type of evidence and to use it bindingly throughout

Europe for this type of evidence. The prerequisite for such a specialized data standard would be a semantic understanding coordinated throughout Europe for this type of evidence, which would thus make this digital evidence largely unambiguously interpretable in each country.

Thus, the above-mentioned initiative to implement the upper secondary school diploma in ELM from a German perspective could be a first step towards a standard for upper secondary school diplomas. This approach would also prevent a "battle of the data standards", which would only lead to a delay in the Europe-wide implementation of such standards.

4.3 Cooperation

Dissemination is usually a required and large part of all EU projects and there has always been cooperation between projects. One example is that the EMREX and EWP projects, running almost in parallel in year 2015–2017, decided to use the same ELMO standard originally designed for EMREX. Another example is the EDSSI project that brought together results of EWP, eIDAS, eduGAIN, myAcademID, Online Learning Agreement and Erasmus app into one project to create a greater usefulness than each project could have done by itself.

There are talks at this very moment between the stakeholders of the ELMO and ELM formats. The long-term goal is to unite the standards and in between make the standards interoperable within the context of the special approach mentioned above. The outcomes will probably be known at the time of the EUNIS conference.

4.4 Conclusion

We believe we have reason to state that while one standard would always be the best option different standards can coexist. There are always different scope for the different initiatives going on, and with that also at least partly different needs for data formats. Applications in need of accessing data described with different standards may solve this by converters. In the long run the standards may converge or it may seem so to the user. One example of the last is Microsoft Excel which can read several different formats.

Regardless of the above said we also believe that cooperation between organizations, projects and initiatives always are of benefit for all parties.

Therefore this paper can be seen as an invitation to (continued) cooperation on formats for educational data.

5 References / Citations

All links have been retrieved in February 2022.

[1] ECTS – European Credit Transfer System Users guide: https://education.ec.europa.eu/sites/default/files/document-library-docs/ects-users-guide_en.pdf

[2] UNESCO website: http://uis.unesco.org/en/topic/international-standard-classification-education-isced

[3] EU website on ESCO: https://ec.europa.eu/esco/

[4] PESC website: https://www.pesc.org/pesc-approved-standards-1.html

[5] IMS website: https://www.imsglobal.org/specifications.html

[6] EBSI website: https://ec.europa.eu/cefdigital/wiki/pages/viewpage.action?pageId=380862522

[7] ELMO schemas in GitHub: https://github.com/emrex-eu/elmo-schemas

[8] EMREX website: www.emrex.eu
[9] EWP website. https://www.erasmuswithoutpaper.eu
[10] Europass website: https://europa.eu/europass/en
[11] ELM schemas in GitHub: https://github.com/european-commission-europass/Europass-Learning-Model
[12] EDSSI website: https://edssi.eu
[13] IDunion website: https://idunion.org/

6 Author biographies

Tor Fridell, M. Sc. in Computer Science and Engineering. Currently Head of Student Information System at Linkoping University and also coordinator for international affairs for the national Swedish Ladok Consortium. Previous jobs include Operations manager for the Ladok Consortium, IS manager for Linkoping Institute of Technology, and programmer. Tor has been employed by the university since 1996. Tor has long been involved in international cooperation regarding exchange of student data and is also active in standards work and development of student information systems. Tor has been President of the European Campus Card Association and Chairman of the National Swedish Standards Committee SIS TK450, the national body for CEN TC 353, working with Learning Technologies. Tor has been involved in the EMREX project since start and is also active in development of student information systems. Tor is currently Chair of the EMREX User Group Executive Committee.

Geir Vangen has more than 20 years' experience in developing nationwide systems within higher education. He is head of development at Sikt – the Norwegian Agency for Shared Services in Education and Research. He participates in national and international standardization work. He has been member of national committees appointed by the Ministry of Education and Research, and has led projects on behalf of the Ministry. Geir Vangen graduated from University of Oslo, Institute of Informatics in 1989.

Janina Mincer-Daszkiewicz graduated in computer science in the University of Warsaw, Poland, and obtained a Ph.D. degree in math from the same university. She is an associate professor in Computer Science at the Faculty of Mathematics, Informatics and Mechanics at the University of Warsaw specializing in operating systems, distributed systems, performance evaluation and software engineering. Since 1999, she leads a project for the development of a student management information system USOS, which is used in 80 Polish Higher Education Institutions. Janina takes an active part in many nation-wide projects in Poland. She has been involved in Egracons, EMREX, Erasmus Without Paper, European Digital Student Service Infrastructure European projects.

Jan Joost Norder works at the Dienst Uitvoering Onderwijs, part of the Dutch Ministry of Education, Culture and Science. In his role as Product Owner he is responsible for the Dutch Diplomaregister and also Chair of the Executive Committee of EMREX. He has many years of experience in improving the digital

enrolment process and exchange of student data in higher education. Since 2016 he has been involved in the international projects.

Jukka Kohtanen MSc (econ.) works as a development manager at CSC – IT Center for Science, a non-profit company owned by the Finnish state and Finnish higher education institutes. Jukka is responsible for a variety of digitalization and collaboration services for the Finnish higher education institutes – such as EMREX and VIRTA – the Finnish national data warehouse for higher education institutes. He has been working with digital services and student registers for 10 years.

Guido Bacharach, Former Head of Strategy and Digitization Unit at the Stiftung for Hochschulzulassung in Dortmund. After his study he had managing positions especially in the sales area and in public services. The focus of his work is on strategic digitization, process improvement and project management. He is a member of the Deutsche Gesellschaft für Projektmanagement (GPM e.V.).

Igor Drvodelić, Assistant Director, Agency for Science and Higher Education participated in the introduction of the central admission of study programs and national secondary school's final exams of in Croatia as one of the project managers. He has 16 years of experience in information systems and services in higher education and has been heading the Central Admission Office since its foundation in 2009. Igor has actively participated in launching projects which better link higher education and the labor market (e.g. graduates tracking). He also actively promotes the introduction of new technologies into Croatian higher education.