Metadata for Learning on Distributed Platforms

Martin Christian[1], Sonja Klante[2], Carmen Biel[3]

German Institute for Adult Education, Germany

Abstract

For two decades, learning object metadata standards have facilitated the filtering, cataloguing, and creation of dependencies. Initially, cross-schema compatibility was limited, as learning content was primarily managed within individual systems. However, recent years have seen a shift towards greater interoperability, where the sharing of content across platforms has become crucial. In the light of the growing importance of interoperability, particularly in the context of competing or complementary content across various learning management systems (LMS), this paper will outline the most relevant metadata standards for the educational sector, as well as different scenarios of application. Finally, we present an insight into the TrainSpot2 project and the metadata solution developed for targeting interoperability amongst different platforms and providers in the ACE context.

Keywords: *Metadata, competence oriented learning, learning on distributed platforms, interoperability, adult education*

1. Introduction

The advent of digitalisation has created a plethora of new opportunities in the realm of online learning, encompassing diverse learning environments, recommendations for learning, and the networking of learning opportunities. Metadata, in its capacity as a data structure, can provide information about learning environments and learning objects, including their content, origin, structure, or quality. Additionally, metadata can furnish insights about the learners themselves, such as their learning status, learning activities, preferences, and learning objectives.
The growing availability of learner data makes it increasingly feasible to utilize it as a foundation for providing precise feedback and recommendations for further learning opportunities. The processing of this data is also becoming relevant for cross-platform analysis. However, for this to be possible, it is necessary to have uniformly comprehensible data from the respective systems. Currently, many learning platforms are only inadequately or not at all equipped with metadata, which is one essential building block for functioning recommendation systems and other applications. In addition, there are numerous different metadata schemes and standards that usually only cover certain cases, such as different educational sectors or different formats of offerings. In the following section, we will outline the most relevant metadata standards for the educational sector and the advantages and disadvantages associated with them. Thereafter, we will present three application scenarios for metadata in the educational sector that are all used in our current project, which will then be discussed to trace the process of agreeing on a common metadata scheme.

2. Relevant Metadata Standards for Education – Differentiation and References

The number of existing metadata models and standards is considerable. The models are designed with a view to the specific purpose of the learning content or the educational sector in question. This encompasses digital learning resources, face-to-face offerings, blended learning programs, educational institutions at all levels, vocational training, and adult education.
The most prevalent standard in the field of education is LOM (Learning Object Metadata) [1], which was initially developed by the Institute of Electrical and Electronics Engineers (IEEE) in 2002. The LOM was developed with the intention of facilitating the search for, exchange and use of learning objects. This focus on reusability and interoperability is evident in the way it has been designed [2]. The LOM covers nine categories that address distinct facets of a learning object. The standard begins with general attributes, life-cycle information, meta-metadata, technical aspects, educational descriptions, rights, relations and finally annotations and classification. The LOM standard is characterised by a high degree of complexity, yet simultaneously offers considerable flexibility in the assignment of metadata, thereby facilitating flexible use.
The Dublin Core (DC) is designed to facilitate the discovery and reuse of educational resources. Consequently, the schema is also designed to be relatively concise, comprising only 15 metadata

elements at its core. These include the following elements: Contributor, Topic, Creator, Date, Description, Identifier, Relation, Rights, and others [3]. The minimal number of elements inherent to the standard, for which the Dublin Core Metadata Initiative (DCMI) is responsible, renders it relatively straightforward to utilise. However, this also constrains the potential applications of the metadata.
To address this limitation, the DCMI developed an extension to the DC, the Learning Resource Metadata Initiative (LRMI), which was published in 2013. In addition to the core elements, there are further attributes that can be used to describe a learning object. Such information may include details of the target group, level or workload [4], [5].
In 2020, several MOOC providers initiated the development of a unified scheme, specifically designed for this type of course. This led to the development of the MOOChub platform, which provides a metadata schema that is specifically tailored to the needs of this offer. Meanwhile, the schema has become established in German-speaking countries and has already been revised on several occasions [6]. The MOOChub API schema enables the provision of courses with highly detailed metadata. Since the release of version 3 in 2023, it has enabled the standardised representation of competences based on the values of the three competence models ESCO [7], DigCompEdu [8] and GRETA [9].
The most recent schema to be discussed is the General Metadata Profile for Educational Resources (AMB). Similarly, this schema is based on the standards of scheme.org. This facilitates the international recognition of standardisation. Another distinctive feature of the AMB scheme is the mandatory use of authorised vocabularies. This implies that the values of the individual attributes must also be standardised to ensure optimal interoperability [10].

3. Scenarios of Application

Depending on the scheme that is applied, there are many different attributes that learning content can be equipped with. In practice, however, it is very time-consuming for authors and editorial teams to assign metadata for every available attribute, as this task is sometimes associated with a high level of research effort. In some cases, the usefulness of individual attributes is also questioned, which is why it can be assumed that the willingness to assign metadata is rather low [11].
Using metadata in an educational context happens for a specific purpose. It is very relevant to always have in mind, that there is a didactical need, that has to be addressed by the metadata. Otherwise, the use of metadata becomes an end in itself. There are three areas that profit immensely from well curated metadata sets. First of all, the application of metadata for search purposes will be described, followed by the use within recommendation systems to provide more adaptive learning experiences and finally the use of metadata for providing meaningful micro credentials and learning progress reports.

3.1. Exchange of Catalogues

One of the primary reasons for the labelling of learning opportunities with metadata is to facilitate their discovery. The search can be conducted within the system itself, or alternatively, by an external service, such as when an offer is presented within a repository. For open educational resources (OER), there are numerous examples, including WirLernenOnline (https://wirlernenonline.de) and twillo (https://www.twillo.de), which are especially suitable for German-speaking countries. The WirLernenOnline repository employs the LRMI metadata profile, while the Twillo repository utilises the OERSI search index, which is based on the AMB schema [10]. Both repositories contain a substantial number of learning resources and materials. To facilitate users in identifying suitable materials from the vast array of offerings, it is of paramount importance that the materials have been provided with metadata that has been carefully curated.
To make a sensible selection from the variety of offers, the portals use filters that allow the resources to be narrowed down. One of the most important criterions for this is subject classification. This categorisation is particularly easy in the school [12] and higher education sectors [13], as subjects can usually be easily distinguished from one another based on curricula. However, in the field of ACE, this is more difficult, as the boundaries are usually more fluid. Furthermore, there is no standardised list of subjects that could be made available for a search. A Europe-wide thesaurus, such as that offered by the EuroVoc portal, is probably too coarse for the requirements [14].

3.2. Using Educational Metadata for Recommending and Adaptive Learning

Recommender systems are becoming more and more state-of-the-art for E-Learning systems and the desire to adapt learning to the individual needs and to the specifics of a learner is as old as education itself.
To generate meaningful recommendations for further learning, algorithms need amongst other things information about the learner themselves. This includes for instance the level the learners are at for each competence facet, their targeted learning goal (e.g. acquiring certificates or information on a specific topic) and their preferences how they like to learn (e.g. duration, preparation, and presentation of content). This information requires data from assessments, learning behaviour and learner queries or self-assessments.
However, this information is something that a learning environment does not have when a new user enters the platform (cold-start problem) [15]. To overcome the cold-start problem, the content must be provided with appropriate metadata, such as competence facet, type of preparation, duration, task type and level, so there is a second content-based layer that helps alleviate missing data points in the system. The system must also recognize which content is relevant for which assessments or can serve as additional support, which can be done by tagging of metadata with curriculum-related information [16].
In addition, a recommendation system can furthermore be supported by rules on how two or more data sets should be linked meaningfully; intelligent systems (with machine learning capability) can furthermore learn to adapt those rules over time.

3.3. Generating Meaningful Micro Credentials and Learning Progress Reports

In the context of migration movement and skills shortage the recognition of non-formal and informal learning and making competences visible is more important than ever. To recognize and validate competencies as well as to identify possible entry points for further development, micro credentials as well as learning progress reports can be of great support. Although especially qualitative approaches are well suited to recognize competencies as those are more open to individual differences [17], quantitative procedures like reliable and valid assessments are in context of recent technological developments and the potential that automated processes can offer in reference to the topic considered here of the utmost importance.
These assessments need to be mapped with metadata referring to competencies, levels and curriculum-related information, and their respective outcome data on said competencies, etc. can be used to document and visualise the competencies a person has. Furthermore, it would also be possible to define thresholds within the competence layer and to provide micro credentials once a certain threshold is met.
When learning or concrete development goals are also stored in the persons profiles and aligned with the measured outcomes, these progress reports, micro credentials, and potential visualisation of the status quo can become a reference point for reflection on potential further competence development.

4. One Size Fits All? Insights from the TrainSpot2 Project

In order to gain a deeper insight into the requirements of various ACE providers, we conducted an analysis of the information they wished to make available in their metadata. This analysis also aimed to identify the most crucial information when targeting interoperability between different platforms and providers within the ACE context.

4.1. The TrainSpot2 Project

TrainSpot2 is a project within the so-called initiative "Digital network infrastructure for education" funded by the Ministry of Education and Research (BMBF) in Germany. This initiative aims at developing a digital ecosystem which, among other things, tries to link educational institutions across educational sectors to make it easier for learners to find suitable learning offers and to be able to document educational certificates in a uniform manner from pre-school to ACE.
TrainSpot2 is working with five project partners to provide various digital services to enable the professional development of teachers in ACE via interconnected platforms to provide a Train-the-Trainer Network and Community (TtT-HotSpot). A special focus is on the targeted acquisition and measurement of generic pedagogical competences that are necessary for teaching in ACE. Therefore, overarching assessments are being developed and existing ones reused to help learners determine their current level of expertise. The assessment structure builds on materials from the GRETA

competence model [9] and is as short as possible to be able to provide a quick entrance point for competence assessment and potential competence development.
Utilizing the GRETA competence model, TrainSpot2 follows the path of several other ACE providers. Educational content related to generic pedagogical competences is provided by the project partners' platforms, portals, and infrastructures. In addition, the integration of further providers will be possible and desirable due to open and standardized interfaces.
Based on initial assessments, learners can target learning content from the connected partner platforms. For this purpose, the content of all platforms that are part of the TtT-HotSpot will be expanded with metadata that reference the GRETA competence model. In addition, the competence facets are aligned with the ESCO competence model [7] so that a connection of further learning content and other content areas from third party platforms can be realized. By passing the assessments and recording their competences, learners receive an individual dynamic competence record. This competence record can also be documented in a personal data wallet and made available to third parties if users agree to this.
Networks and exchange formats will be established to connect other TtT offers with a view to agreeing on common standards as well as on generic pedagogical competences across educational sectors. In addition, further TtT providers are to be advised and won over to participate in the hotspot. The focus is also on the question of how the GRETA competence model can be applied in other educational sectors as well as the creation of cross-educational sector standards for awarding certificates and badges.

4.2. Finding a Common Standard

The intended use of the metadata was based on a criteria-based search and the development of a cross-platform recommendation system. A total of five platforms from providers whose aim includes the further training of teachers in ACE were analysed. The existing offerings are both commercial and non-commercial. There are complete web-based self-study programmes as well as online and face-to-face programmes. The first step was to identify commonalities that could be identified from the names of the individual attributes. The initial basis for the analysis was the LOM standard. Subsequently, the attributes were also compared with the MOOChub API schema to ascertain whether this could be used to cover all relevant requirements. A comparison with the AMB schema would also have been beneficial. However, as this was not published until the end of 2023, it could not be included in the study which was already conducted in mid 2023. The aforementioned standards DC and LRMI were also not considered in this analysis. This was because the analysis of the standards had already demonstrated that they were unable to provide the requisite scope of attributes for the purpose of the project.
It subsequently became evident that the MOOChub API schema is capable of mapping all metadata relevant to the project. In addition to the title, description, and numerous technical details, the metadata with the highest relevance was found to be the content type (video, quiz, podcast, text, etc.), duration, educational framework with a description of the competences, keywords, and target group.
It should be noted that a thematic assignment is made based on the GRETA competence model, which distinguishes 23 competences for working as a teacher in ACE. All participating providers have aligned their programmes with this. However, to differentiate it from other offerings, a subject-specific categorisation must also be carried out, which could be done via the "educationalSubject" (MOOChub API) or "about" (AMB scheme) field.
Only one portal expressed a need for attributes that extend beyond the MOOChub API offering. This pertained to the designation of the course phase during which an object can be used and the type of teaching and learning arrangement. However, for the purposes of assigning metadata within the project context, this information can be dispensed with, which is why the lack of information does not represent a disadvantage.

4.3. Learning Success Through Learning Progress Reports

The idea behind the TrainSpot2 project is as mentioned above that learners can see their skills development by a dynamic competence balance and recognize in which areas they are already well positioned and where they could usefully continue their learning and professional development. Depending on their current interests, it should be possible to display directly which content is useful for which area next. This can refer to different providers and content types. This way, learners know

where they can find suitable content and can then decide for themselves which offer they want to make use of.
In the example presented, metadata is assigned to the competences addressed. As the competence metadata relate to the GRETA competence model [9], which has achieved a certain degree of dissemination within ACE, it can also be used to make references between programs from different providers. Furthermore, it is also possible in principle to use mapping to compare content with programs that work with comparable competence models, such as ESCO [7] or DigCompEdu [8].
In this specific case, the metadata schema of the MOOChub API is used. This contains an area that is used to label competence information. Only standardised values are permitted here, which are fed either from ESCO, DigCompEdu or GRETA. In addition to the educational framework, a description of the competence and a corresponding ID are included. It is also possible to store a link to the framework and use an URI, for example, as the ID.
In addition to the designation of the skill, another criterion is the level, i.e. the so-called educational level. In the case of TrainSpot2, four levels are used, which are roughly based on the learning objective taxonomy of Anderson and Krathwohl, but are presented in a somewhat reduced form overall, as the transitions between the individual levels are not particularly clear-cut. [18]
The data schema employed for the generation of a competence balance utilises the attributes provided by the MOOChub API and are subsequently transferred into a novel data scheme, which represents a micro-credential. Only those attributes required for the generation of the visualisation are collected here. The following example illustrates the scheme that is employed in the TrainSpot2 project.

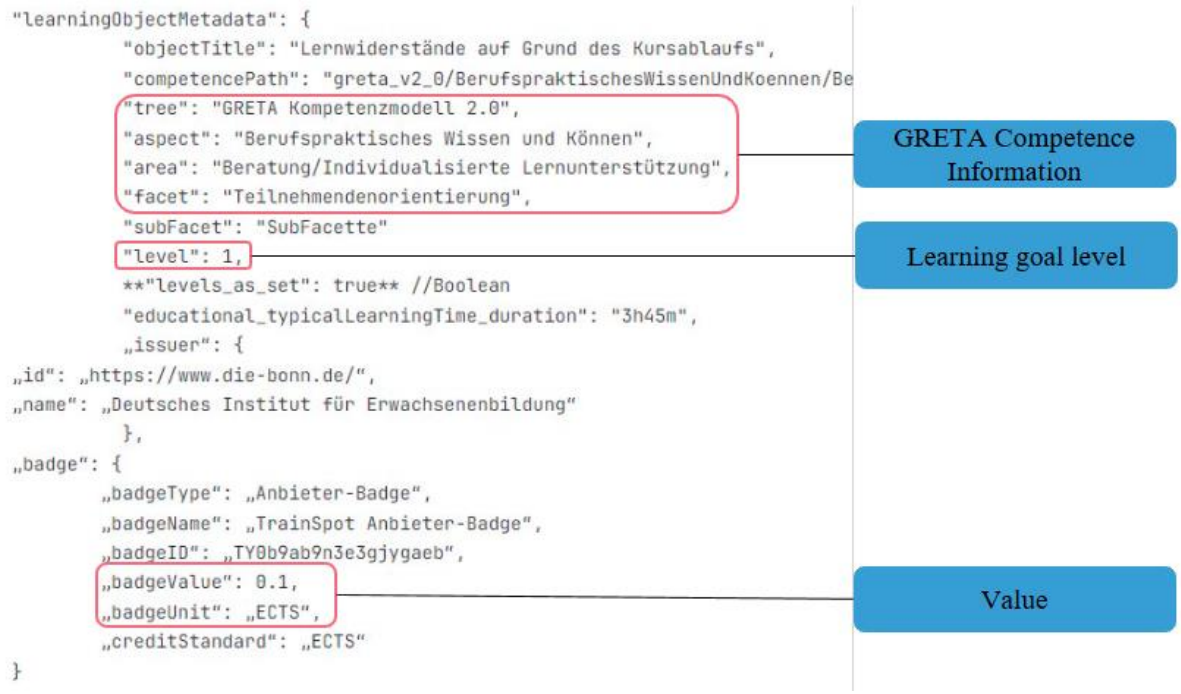

Fig. 1. TrainSpot2 Data Scheme for Microcredentials. Own representation.

If a learner successfully completes a learning unit, the competence information linked to the learning object is first transferred to a data wallet. The learner then has the option of transferring the result to a tool for analysing the skills assessment. The competence information is then made available again via an API and can be interpreted by the tool, as this maps the complete GRETA competence model. If the values of the information sent now match the values of the tool, a graphical evaluation is created from this, which enables the user to see an overview of competence development (Fig. 2). As the tool has an API, it is possible to send the results regardless of which training provider they come from - if the offer is mapped to the GRETA competence model. It is intended that the visualization supports learners in deciding on how to proceed their individual competence development.

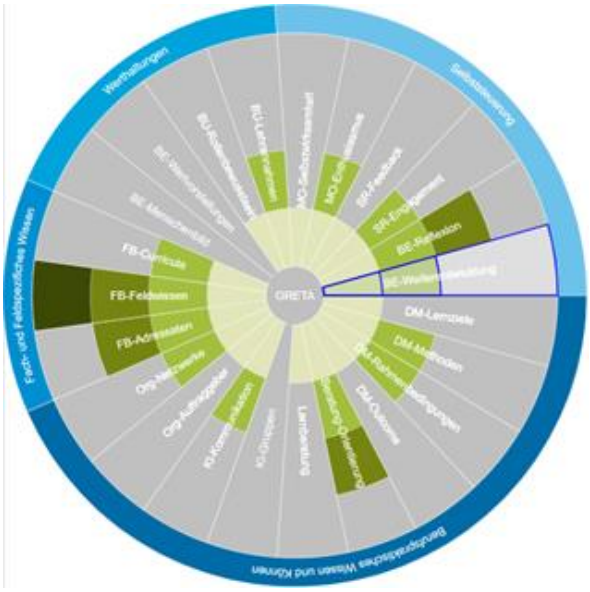

Fig. 2. Competence Balance based on the GRETA Competence model. Own representation.

5. Conclusion

The existing metadata standards are well suited to the implementation of the applications mentioned in this paper. Newer developments, such as the MOOChub API or the AMB schema, combine the advantages of some older standards. The use cases demonstrate that the requisite attributes are available and can be visualised. Nevertheless, a significant obstacle persists, one that has only been partially addressed thus far. In general, each portal operator is at liberty to select the values to be attributed to the attributes. This approach offers considerable flexibility in the creation of metadata, which is beneficial for editors and authors. However, this also creates significant disadvantages in terms of interoperability, as different attribute designations can be misinterpreted. Consequently, the logical subsequent step is to increasingly rely on standardised vocabularies when creating metadata. As the networking of educational programmes progresses, the necessity for this requirement becomes increasingly apparent. Consequently, further studies may be conducted to ascertain the extent to which standardisation would enhance the value proposition.

It will be important to conduct a review of individual projects to ascertain the efficacy of the mapping and translation of metadata. The establishment of rough structures across educational sectors provides an overview of the schemas that are specific to each sector. Depending on the educational sector, there may be more detailed schemas that can also be mapped to higher-level schemas and models. Furthermore, the utilisation of automated mapping tools can facilitate the alignment of learning content with higher-level aggregated and sector-specific schemas. A prototype of this tool is currently being tested. The tool is employed to elucidate the interrelationships between the European competence framework and the GRETA competence model. This would streamline the work of content developers and provide users with a more comprehensive understanding. Furthermore, the implementation of such a system would facilitate the comparison of learning programmes with one another, potentially leading to the development of joint certificates.

REFERENCES

[1] IEEE. (2020a). *1484.12.1-2020—IEEE Standard for Learning Object Metadata*. https://doi.org/10.1109/IEEESTD.2020.9262118

[2] IEEE. (2024). *Learning Metadata (LMT)*. GitLab. https://opensource.ieee.org/lmt/lmt

[3] DCMI (Hrsg.). (2020). *DCMI Metadata Terms*. https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#section-3

[4] DCMI (Hrsg.). (2024). *LRMI Concept Schemes*. https://www.dublincore.org/specifications/lrmi/concept_schemes/

[5] LRMI (2014). *LRMI–Dublin Core Metadata Initiative.* https://www.dublincore.org/specifications/lrmi/

[6] Ebner, M., Koschutnig-Ebner, M., Rampelt, F., Serth, S., Staubitz, T., von Stetten, A., Thomas, M., & Wittke, A. (2023). *Metastandard für den internationalen Austausch von MOOCs – der MOOChub als erster Prototyp*. Zeitschrift für Hochschulentwicklung, 18(1), 17–35.

[7] European Commission, Directorate-General for Employment, Social Affairs and Inclusion. (2018). *Europäische Kompetenzen, Kompetenzen, Qualifikationen und Berufe (ESCO)* (V.1.0.3). http://data.europa.eu/88u/dataset/european-skills-competences-qualifications-and-occupations

[8] Punie, Y. (Hrsg.), (2018). Europäischer Rahmen für die Digitale Kompetenz Lehrender. DigCompEdu. https://mz-hofgeismar.de/flip/digcompedu/files/assets/common/downloads/publication.pdf

[9] Lencer, S. & Strauch, A. (2016). Das GRETA-Kompetenzmodell für Lehrende in der Erwachsenen und Weiterbildung. http://www.die-bonn.de/id/34407.pdf

[10] Pohl, A., Klinger, A., Hartmann, B., Schuurbiers, C., Steeg, F., Kummerländer, M., Oellers, M., Stengel, M., Hoffmann, M., Rörtgen, S., Kulla, S., & Bülte, T. (2023). *Allgemeines Metadatenprofil für Bildungsressourcen (AMB)*. https://dini-ag-kim.github.io/amb/20231019/

[11] Tischler, F., Heck, T., & Rittberger, M. (2022). Nützlichkeit und Nutzbarkeit von Metadaten bei der Suche und Bereitstellung von offenen Bildungsressourcen. *Information – Wissenschaft & Praxis*, *73*(5–6), 253–263. https://doi.org/10.1515/iwp-2022-2238

[12] i.e. vocabs.openeduhub.de/w3id.org/openeduhub/vocabs/discipline/

[13] i.e. skohub.io/dini-ag-kim/hochschulfaechersystematik/heads/master/w3id.org/kim/hochschulfaechersystematik/scheme

[14] https://eur-lex.europa.eu/browse/eurovoc.html?params=32,3211,795#arrow_795

[15] Su, X., & Khoshgoftaar, T. M. (2009). A Survey of Collaborative Filtering Techniques. Advances in Artificial Intelligence, 2009, 1–19. https://doi.org/10.1155/2009/421425

[16] Digel, S., Krause, T., & Biel, C. (2023). Enabling Individualized and Adaptive Learning – The Value of an AI-Based Recommender System for Users of Adult and Continuing Education Platforms. In N. Wang, G. Rebolledo-Mendez, V. Dimitrova, N. Matsuda, & O. C. Santos (Eds.), Communications in Computer and Information Science. Artificial Intelligence in Education: Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky (Vol. 1831, pp. 797–803). Springer. https://doi.org/10.1007/978-3-031-36336-8_121

[17] Bosche, B., & Strauch, A. (2023). Qualitative Verfahren der Kompetenzanerkennung zu Validierungszwecken - theoretische Einordnung und Praxisbeispiele. In M. Schmid (Ed.), Handbuch Validierung non-formal und informell erworbener Kompetenzen. Disziplinäre, theoretische und konzeptionelle Zugänge (pp. 287–306). wbv Publikation.

[18] Anderson, L. W., & Krathwohl, D. R. (Hrsg.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.