Learning Resources Metadata: Opportunities and Challenges in the German "Innovationswettbewerb INVITE" Program

Sheikh Faisal Rashid[1,*], Pascal Hürten[2], Insa Reichow[1] and Lutz Goertz[3]

[1]Educational Technology Lab, German Research Center for Artificial Intelligence (DFKI), Alt-Moabit 91c, 10559 Berlin, Germany

[2]Institut for Interactive Systems (ISy), Technische Hochschule Lübeck, Mönkhofer Weg 239, 23562 Lübeck, Germany

[3]mmb Institut GmbH, Folkwangstraße 1, 45128 Essen, Germany

Abstract

The increasing use of digital learning resources in vocational education and training (VET) has created a need for effective management and discovery of these resources. Learning resource metadata, which provide descriptive information about digital resources, have the potential to improve the discovery, reuse, and sharing of learning resources. However, the use of metadata in VET poses a number of challenges. This paper presents the opportunities and challenges of using metadata for learning resources in the context of the German funding program "Innovationswettbewerb INVITE". It discusses selecting metadata and metadata standards for learning opportunities, learners, and digital credentials within the INVITE projects and how, in contrast, international experts select metadata and metadata standards. One of the challenges frequently mentioned is the effort needed to assign large amounts of metadata. This paper thus explores the use of Large Language Models (LLMs) as a tool to address this challenge.

Keywords

Educational Metadata, Metadata Standards, Vocational Education and Training (VET), innovationswettbewerb INVITE, Large Language Model (LLMs)

1. Introduction

The vocational education and training (VET) sector is undergoing significant change, driven by the need to prepare learners for an increasingly demanding and rapidly changing workforce. The progress in the use of digital learning infrastructures and artificial intelligence (AI) has created new opportunities for educational providers to develop individualized learning paths, but also new challenges in terms of managing and delivering quality learning resources to

DCMI-2024 International Conference on Dublin Core and Metadata Applications
*Corresponding author.
faisal.rashid@dfki.de (S. F. Rashid); pascal.huerten@th-luebeck.de (P. Hürten); insa.reichow@dfki.de (I. Reichow); goertz@mmb-institut.de (L. Goertz)
https://www.dfki.de/web/ueber-uns/mitarbeiter/person/shra03 (S. F. Rashid); https://www.th-luebeck.de/isy/team/unser-team/pascal-huerten-1/ (P. Hürten); https://www.dfki.de/web/ueber-uns/mitarbeiter/person/inre01 (I. Reichow); https://www.mmb-institut.de/kontakt/ (L. Goertz)
0000-0003-1551-8884 (S. F. Rashid); 0000-0000-0000-0000 (P. Hürten); 0000-0002-7799-8409 (I. Reichow); 0000-0001-9453-2021 (L. Goertz)

https://doi.org/10.23106/dcmi.952436347

learners catering to their preferences and competencies. The effective management and discovery of these resources remains a significant challenge. Learning resource metadata, which provides descriptive information about digital resources, can potentially improve the discovery, reusability, and sharing of learning resources both within and across digital learning platforms.

In this work, we present the opportunities and challenges in selecting and adopting suitable metadata and metadata standards in different phases of the digital learning process in the context of the "Innovationswettbewerb INVITE". INVITE funds 35 projects aimed at networking and developing innovative platforms for vocational education and continuing training and the common use of standards. It is funded by the German Federal Ministry of Education and Research (BMBF) with a total of 88 million euros between 2021 and 2025 [1].

1.1. Research questions and methods

Metadata provides crucial information about the content, structure, and context of objects [2]. In digital learning, it helps describe learning resources, learners, and digital credentials [3]. However, managing this metadata is often time-consuming and costly. Despite many available standards, no single established solution exists, especially in domains without top-down regulations [4, 5]. The Innovationswettbewerb INVITE aims to create interoperable educational infrastructures and provides an opportunity to study how metadata are selected by research projects in a bottom-up fashion. The following research questions are addressed as part of this paper:

- Which metadata and standards are chosen for learning opportunities, learners, and digital credentials within INVITE projects?
- How do international experts select metadata and standards?
- What challenges arise in deciding on and assigning metadata standards?

A qualitative two-step procedure involving workshops and expert interviews was employed. Three virtual workshops with ten INVITE projects were conducted to discuss their metadata use for learning resources, learners, and digital credentials. These workshops included preparatory guiding questions, brief inputs from projects, and open discussions. The insights gathered were documented for further analysis. Additionally, guided interviews with five international experts from the US and UK were conducted to understand their approaches to metadata standards [6]. Experts were searched by looking at various educational service platforms, technology firms, and educational standard development organizations. From 20 experts contacted with a request for an interview study, five interviews were finalized. These interviews were analyzed for common themes and contrasting views.

In the process, it became apparent that one of the INVITE projects already uses AI and Large Language Models (LLMs) to automate the extraction of metadata from unstructured course descriptions: The INVITE project WISY@KI focuses on the identification of relevant metadata and annotation of learning outcomes as defined by the ESCO standard [7]. It is observed that LLMs can significantly help to minimize the manual effort involved in assigning metadata to learning resources. The procedure is explained as part of this paper to address one of the challenges mentioned by various experts.

https://doi.org/10.23106/dcmi.952436347

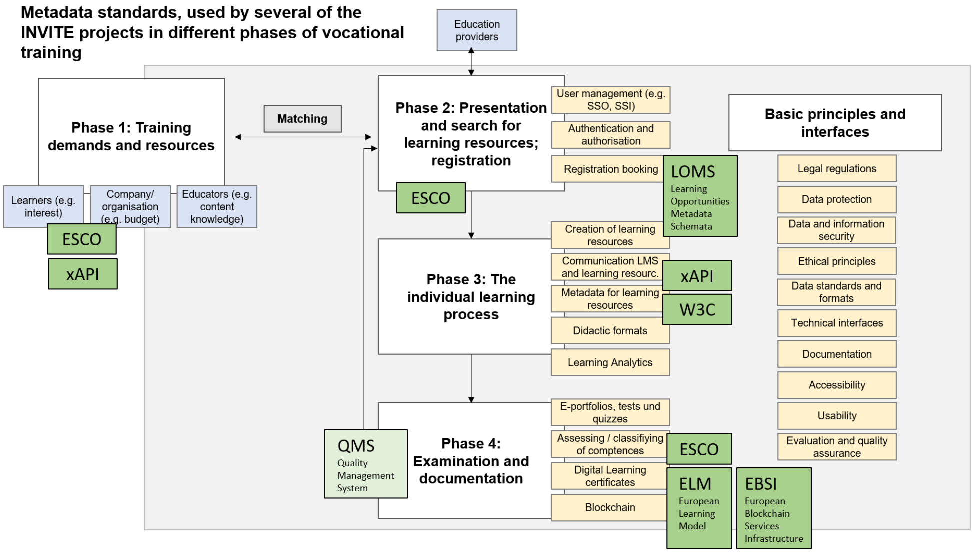

Figure 1: Integration of metadata standards used by the INVITE projects in a phase model of the learning process.

2. Results from INVITE Metadata Workshops

The background for the workshops is a model of the learning process (see Figure 1), including four phases (white boxes) [4]. The learning process typically begins with identifying needs and resources for training (Phase 1) to find a matching course. After deciding on a course, registration is necessary, which usually requires user data in terms of user credentials, personal and educational information (Phase 2). The individual learning process is captured in Phase 3 and may include learning analytics to measure, present, and predict learners' progress. The process concludes with examinations and documentation of the learned material, culminating in the issuance of certificates that showcase the competencies or skills achieved (Phase 4).

The three central workshop topics metadata for (1) learning opportunities, (2) learners, and (3) digital credentials were integrated into the model [8]. According to the results of the workshops, the most frequently used metadata standards were added to the process model (green boxes). It is already apparent that there is no "one-size-fits-all" standard for all learning situations. Different metadata standards have emerged and proven themselves in different contexts. Only ESCO's competency standards appear in different phases. In the following, the standards considered within INVITE are further explained.

ESCO (European Skills, Competences, Qualifications, and Occupations): ESCO is a European classification system that links skills, competencies, qualifications, and occupations. It is intended to improve the transparency and comparability of qualifications in Europe. ESCO turned out to be the overarching standard in the INVITE projects and is used to describe competencies that learners possess, that are taught in learning opportunities, and that are ultimately stated in certificates of completion. One challenge frequently discussed in the workshops, is the adequate mapping of the diversity of professional competencies and qualifications in different

https://doi.org/10.23106/dcmi.952436347

countries and sectors. Definitions of skills and competencies can vary by context, leading to ambiguities, inconsistencies, and gaps. To address this challenge, INVITE projects take the approach of classifying job offers and complementing ESCO in a context-specific way.

xAPI (Experience API) is a technical standard for capturing and tracking learners' experiences and activities, regardless of the learning platform or application used [9]. xAPI is used to communicate between learning content and the learning management system (LMS) and contains information about learners, activities performed, and learning content to provide detailed data about the learners' learning behavior and progress. The information stored in a Learning Record Store (LRS) can be used by the LMS for a comprehensive analysis and adaptation of the learning environment [4]. In INVITE, xAPI is used to realize e.g. learning analytics and data tracking.

The **Learning Opportunities Metadata Schemata (LOMS)** standard describes the metadata of learning opportunities [10]. In INVITE, the LOMS standard of the European Learning Model is used in combination with ESCO for the representation of skills and competencies, but also for adaptive learning within courses. The standard should not be confused with the IEEE Learning Objects Metadata (LOM) standard [11]. This standard has been deemed no longer relevant by the INVITE projects.

W3C standards form the basis for the exchange of metadata in INVITE. The W3C standards are technical specifications developed by the World Wide Web Consortium to ensure the interoperability, security, and accessibility of the World Wide Web. Examples include **Verifiable Credentials** (data model for digital education credentials), **WCAG** (web accessibility), and **WAI-ARIA** (accessibility for dynamic web content) [11].

The **European Learning Model (ELM)** aims to establish a semantic vocabulary for learning in Europe. By standardizing technical terms, it enables seamless data exchange across borders for different scenarios, including the recognition of educational documents. The new open standards-based ELM v.3 is an interoperable data model that ensures compatibility with ELMO and the EBSI Diploma Use Case and is linked to existing frameworks and classifications (e.g., EQF, ESCO, ISCED-f), in particular the W3C Verifiable Credential data model. It is also available in all 31 languages of Europass [12]. INVITE projects use ELM, among other things, for Verifiable Credentials.

The **European Blockchain Services Infrastructure (EBSI)** aims to provide a trusted and secure infrastructure for the use of blockchain technology in the EU. Currently, several use cases are being pursued: ESSIF – a Self-Sovereign Identity model in Europe, digital education credentials or traceable documents [4]. Some INVITE projects participate in EBSI in the early adopters program and thereby contribute in a non-academic context.

A central challenge mentioned by several projects remains the complete and uniform filling of metadata fields. **Quality management systems (QMS)** can play a decisive role. For example, a QMS could ensure that metadata for learning content and certificates are filled in a standardized and correct manner. In INVITE QMS is used for learning outcomes. In the workshops, it was also discussed that AI could be used to validate the quality of the machine-filled metadata fields.

https://doi.org/10.23106/dcmi.952436347

3. Results from Expert Interviews

This section discusses how experts are using metadata in their educational applications. The experts from the USA are referred to as E-USA1 to E-USA4, and the expert from the UK is referred to as E-UK. The findings are presented in three subsections: 3.1 provides practical usage of metadata by experts, 3.2 describes challenges in using metadata and how experts seek to cope with these issues, and 3.3 highlights the metadata standards discussed by the experts.

3.1. Practical Usage of Metadata - Why experts use metadata and how

Experts mentioned the following uses of metadata in the educational context: (1) searching the educational content or offers, (2) recommending educational content or offers, and (3) describing competencies.

(1) One of the crucial functions of metadata is providing descriptive information that helps users find relevant resources more efficiently. It also helps in the reuse of learning resources. To find things more easily, some experts mentioned the importance of **context**.

Opinion E-USA1: *"In the whole learning ecosystem, the idea is not only to be aware of the available learning opportunities but also knowing the "context" describing the learner's individual credentials and history to realize the actual need of the learner while performing the search (instead of what the learner wants). In this way, users can search for learning opportunities through a system that can better understand what is available in terms of learning resources as well as the context from which that user is originating."*

Opinion E-USA2: *"In the corporate sector, things are constantly changing that increases the complexity. We need to assign many labels to the same content because either the context has changed over time, or the user is looking at things with a different context in his mind."*

The context can be added to metadata with the help of ontologies by providing a structured framework for representing and organizing knowledge about a specific domain.

(2) Another important use of metadata is **content recommendations** to provide relevant suggestions to learners based on their interests or learning goals [13]. Metadata associated with user profiles, course attributes, learner feedback and ratings, and contextual factors, can be used to generate personalized and relevant suggestions to learners, helping them discover courses that match their interests, level of expertise, and learning objectives. One expert envisioned an advanced, implicit recommendation of educational content at the point of need:

Opinion E-USA2: *"The mature phase of recommendation would be giving learners exactly what they might need, at the point they need it, or even (based on the learning experience) recommending stuff without being explicitly required. The goal is to deliver something to someone before they even know they need it. This can be achieved with the help of "learning context" and "learning experience" metadata (IoT, AR/VR, user behavioral data, etc.) along with environmental triggers."*

(3) Metadata are also required to describe **competencies**. Competencies are often structured and contextualized within competency frameworks. These are specific knowledge, skills, or abilities that a learner should acquire during a course. Associating competency metadata with learning resources enables learners and organizations to identify the most relevant resources for developing and accessing specific competencies (competency-based search). E-USA3 mentioned the use of competencies from the "Open Competency Framework Collaboration (OCFC)", which

https://doi.org/10.23106/dcmi.952436347

is a US-focused organization, but the standards defined can be used internationally.

3.2. Coping with problems in using metadata

One major problem mentioned by experts is **how to get complete and accurate information for the metadata fields**. Mostly, metadata fields are filled in manually by individuals using their knowledge of the learning resource. However, the experts emphasized that a certain level of training and knowledge is necessary. Additionally, ensuring consistency in the data can be a challenge. Motivating individuals to complete all relevant fields can also be a difficult task. For example, although one expert believed that a metadata framework like SCORM is robust and easy to adopt, the individuals assigned to fill in the fields sometimes only make minimal effort, resulting in less useful outcomes. To ensure that the provided content is of high data quality or has been tagged accurately, experts suggest training individuals to understand the significance and content of a taxonomy or metadata field. Alternatively, subject-matter experts or educators should tag their own resources to avoid inaccurate entries. An alternative approach would be to write metadata prior to developing the learning resource, including all necessary information.

Another issue faced by the experts in their application scenarios is the use of non-standard metadata that hinders **interoperability** of the system to another, as people may use different schemas for recording metadata. Ideally, common metadata fields should prevail, ensuring seamless transformation. However, executing such transformations can pose significant challenges. A solution lies in using standardized metadata schemas as well as standards to fill the metadata fields.

The **User Interface** is another issue. What user interface is good for filling the metadata fields efficiently? There are various options, e.g. free text, fixed list items to Excel worksheets. If a lot of free text is to be filled, people do not fill it out most of the time. A fixed list is also problematic, it is not dynamic and cannot be changed and if you change you need to re-tag everything.

3.3. Metadata Standards – Which metadata standards are used by the experts?

To date, several learning metadata standards exist that can support educational institutions, content creators, and learners. For example, the IEEE Standard for Learning Object Metadata (LOM) and Dublin Core metadata standards have been in practice for the last few decades [14]. E-USA1 discussed the use of Dublin Core, LOM, and other standards like Learning Resource Metadata Initiative (LRMI) [15], Credential Transparency Description Language (CTDL) [16], and Schema.org [17] in the development of IEEE P2881-Learning Metadata Terms (LMT) [18]. LMT aims to bridge different communities and standards to serve as a shared vocabulary for describing digital objects related to learning. E-USA3 additionally mentioned the use of SCORM [19] for content management and the use of xAPI [20] for the representation of learning activities. The major weakness of xAPI is that it offers a standard for communicating activities but does not provide a schema for describing these activities. E-UK revealed the use of LRMI [15] and CTDL standards. The LRMI allows us to describe the educationally significant characteristics and relationships of a resource, e.g. to describe the relationship of resources to competencies (e.g. "teaches a particular competency") or the type of resources (is it an instructional video or a

https://doi.org/10.23106/dcmi.952436347

textbook). It also encompasses details like the typical learning time and target audience. LRMI can be used in recommender systems, allowing for more personalized and effective resource recommendations based on specific user preferences and needs.

4. Leveraging Large Language Models for Enhanced Interoperability and Skill Standardization in the WISY@KI Project

Diverse and unstandardized documentation in educational systems makes it difficult to assess learning outcomes and recognize learning achievements. To address this, the WISY@KI project leverages Large Language Models (LLMs) to bridge the gap between diverse learning offerings and learner needs. In this section, we explore the potential of LLMs in standardizing skills and competencies, with a focus on the ESCO standard, and discuss the benefits in the context of the WISY@KI project.

4.1. Methodology

Extraction of structured data: Our approach[21] addresses the challenge of extracting relevant metadata from diverse, unstructured, and semi-structured documents, including PDFs, HTML, XML, and JSON files. For instance, we can extract metadata such as learning outcomes, prerequisites, educational level, and workload from module descriptions. By leveraging LLMs and data model definitions, we can structure these documents and extract valuable information. This enables the transformation of unorganized data into standardized, machine-readable formats, facilitating its use in various downstream applications and creates a good foundation for further standardization efforts.
Data Standardization: The extracted metadata is standardized to ensure comparability across systems by aligning the extracted data with established taxonomies such as the ESCO classification. Our data standardization approach consists of three key steps. First, we use clustering to group the extracted learning outcomes by similar topics, rather than treating them as a whole. This step is crucial in ensuring that every aspect of the described learning is covered in the subsequent step. Next, we utilize a vector database that stores competency descriptors from the target taxonomy, each represented as a sentence embedding. This allows us to efficiently retrieve the most relevant competencies for each aspect of learning. To generate these sentence embeddings, we employ a fine-tuned BI-Encoder Model[22], which enables semantic search and facilitates the matching of learning outcomes with relevant competencies. In the final step, we use a specialized Cross-Encoder Model[23] to refine the candidate competencies, following the retrieve-and-rerank approach described by Reimers (2024) on the SBERT documentation[24]. Additionally, we employ other non-AI related algorithms, like domain-specific filtering, to further improve accuracy. This ensures high precision in matching learning outcomes with standardized competencies, enhancing the accuracy and reliability of the skill standardization process while improving efficiency and scalability.

https://doi.org/10.23106/dcmi.952436347

4.2. Addressing Challenges and Limitations

The development of LLM-based skill extraction systems is a complex task, and several challenges and limitations must be addressed to ensure their efficiency, accuracy, and scalability. We address three key areas: model fine-tuning, human dependency, and continuous model adaptation.

Model Fine-Tuning To optimize operational efficiency and improve accuracy, we transitioned from directly using large language models (LLMs) to leveraging specialized models wherever possible. Inspired by Guo et al. [25], we utilized LLMs as "Trainer models" to generate synthetic datasets. These synthetic datasets were then employed for fine-tuning smaller, specialized BI- and Cross-Encoder models. To ensure real-world applicability, we complemented this approach with human-evaluated data. The effectiveness of fine-tuning using synthetic data prompted us to shift focus towards generating human-validated data for further training. Consequently, we developed new, fine-tuned models that substantially improved performance in retrieving standardized skills for German course descriptions. Table 1 compares the performance of specialized models and their non-finetuned counterparts in retrieving the most relevant skills for course descriptions. The base model, *intfloat/multilingual-e5-base*, was fine-tuned to create *isy-thl/multilingual-e5-base-course-skill-tuned*. We observed that fine-tuning significantly improved domain-specific performance. Additionally, further enhancing the model by reranking the candidate results from the fine-tuned BI-Encoder with a fine-tuned Cross-Encoder model (*isy-thl/bge-reranker-base-course-skill-tuned*) resulted in even better performance, although with increased computation time.

Table 1
Performance comparison of base and fine-tuned models in retrieving relevant skills for course descriptions.

Model	recall@3	recall@5	recall@10	mrr@10	avg_time
intfloat/multilingual-e5-base	0.073	0.114	0.170	0.322	0.343
isy-thl/multilingual-e5-base-course-skill-tuned	0.231	0.305	0.445	0.699	0.343
isy-thl/bge-reranker-base-course-skill-tuned	0.253	0.363	0.521	0.742	10.028

Human Dependency and Continuous Model Adaptation While human verification is crucial for ensuring the relevance and accuracy of metadata, it introduces potential bottlenecks. To address this, we integrated our approach as an API into existing editorial workflows, incorporating human-in-the-loop feedback as training data for future models. Additionally, we decoupled taxonomies from models, enabling quick adaptation to changes in educational standards without the need to retrain models. This approach allows for updating taxonomies and language models separately, ensuring scalability and flexibility.

5. Discussion

The example of the ten projects in the German funding program INVITE demonstrates that different organizations and companies can agree on common metadata for educational purposes.

https://doi.org/10.23106/dcmi.952436347

They implement standards tailored to their needs in a grassroots approach, enabling them to exchange data and learning content and expand their offerings. This is a significant step forward and could serve as a model for others. Experts we interviewed confirmed the importance of metadata and the need for appropriate standards. In the past, many standards were neglected and fell into oblivion due to a lack of awareness. However, recent years have seen positive developments: metadata standards are now well-known, and various working groups and standardizing committees are increasingly coordinating their efforts. Clear, accessible, and well-maintained standards are essential for widespread adoption.

Leveraging large language models (LLMs) for educational metadata management offers transformative advantages, improving interoperability, accuracy, and user-centricity in educational systems. By extracting metadata from various documents, a seamless learning ecosystem can be envisioned where learners can make informed decisions about their education. The proposed methodology aims to extend beyond course module descriptions to encompass the entire learning journey, creating dynamic skill profiles that reflect individuals' skills and achievements. Additionally, adopting these models has significantly reduced the overwhelming workload in the education sector associated with manually digitizing, modernizing, and structuring existing databases of module descriptions.

Acknowledgments

This work is supported by the German Federal Ministry of Education and Research (BMBF) through the program "Innovationswettbewerb INVITE". We would also like to express our sincere thanks to the experts who took part in the interview study and to all those who contributed to the preparation and comments of this paper, especially our colleagues from the INVITE-Meta project.

References

[1] BIBB, Innovationswettbewerb INVITE, 2024. URL: https://www.bibb.de/de/120851.php.

[2] A. J. Gilliland, Introduction to metadata, Getty Research Institute, Los Angeles, CA, 2008, pp. 1–19.

[3] M. Oellers, S. Rörtgen, Kompendium: Didaktische Metadaten, 2024. doi:10.25656/01:29235;10.5281/zenodo.10828758.

[4] I. Reichow, M. Hochbauer, L. Goertz, Standards und Empfehlungen zur Umsetzung digitaler Weiterbildungsplattformen in der beruflichen Bildung: ein Dossier im Rahmen des INVITE-Wettbewerbs, Bundesinstitut für Berufsbildung, 2021. URL: https://res.bibb.de/vet-repository_779586.

[5] S. Rörtgen, R. Brenner, H. Zimmermann, M. Hupfer, A. Zobel, U. Lucke, Metadata Standards in National Education Infrastructure: Development of Evaluation Criteria and Their Exemplary Application, in: 21. Fachtagung Bildungstechnologien (DELFI), Gesellschaft für Informatik e.V., 2023.

[6] S. Rashid, L. Goertz, I. Reichow, Metadata for Learning Processes – Results of an

https://doi.org/10.23106/dcmi.952436347

International Interview Study, Preprints, 2024. doi:https://doi.org/10.20944/preprints202402.0889.v1.
[7] European Commission, What is esco?, 2024. URL: https://esco.ec.europa.eu/en/about-esco/what-esco.
[8] L. Goertz, S. F. Rashid, E. Vogel-Adham, A. Vogt, A. Wilhelm-Weidner, Metadatenstandards im Innovationswettbewerb INVITE, 2023. doi:https://doi.org/10.25656/01:27697.
[9] Rustici Software, xapi overview, 2024. URL: https://experienceapi.com/overview/.
[10] European Union, Publishing of Qualification and Learning Opportunity Data Documentation, 2024. URL: https://europa.eu/europass/system/files/2020-07/Documentation_publishing%20of%20Q%20and%20LO%20Data_v2.0.pdf.
[11] IEEE, 1484.12.1-2020 - IEEE Standard for Learning Object Metadata, 2024. URL: https://ieeexplore.ieee.org/document/9262118.
[12] Europass European Union, Upcoming launch of the european learning model v3, 2024. URL: https://europass.europa.eu/en/news/upcoming-launch-european-learning-model-v3.
[13] I. Reichow, K. Buntins, B. Paaßen, H. Abu-Rasheed, C. Weber, M. Dornhöfer, Recommendersysteme in der beruflichen Weiterbildung. Grundlagen, Herausforderungen und Handlungsempfehlungen. Ein Dossier im Rahmen des INVITE-Wettbewerbs, 2022. doi:https://doi.org/10.25656/01:24517.
[14] P. Barker, L. Campbell, Metadata for Learning Materials: An Overview of Existing Standards and Current Developments, Technology, Instruction, Cognition, and Learning 7 (2010) 225–243.
[15] Dublin Core, LRMI, 2024. URL: https://www.dublincore.org/specifications/lrmi/.
[16] Credential Engine, Credential Transparency Description Language (CTDL), 2024. URL: https://credentialengine.org/credential-transparency/ctdl/.
[17] Schema.org, Schema.org, 2024. URL: https://schema.org/.
[18] IEEE SA OPEN, Learning Metadata Terms, 2024. URL: https://opensource.ieee.org/lmt/lmt.
[19] ADL, Sharable Content Object Reference Model (SCORM®), 2024. URL: https://adlnet.gov/past-projects/scorm/.
[20] ADL, Experience API (xAPI) Standard, 2024. URL: https://adlnet.gov/projects/xapi/.
[21] P. Huerten, WISY@KI-API: Competence-Analyser, https://github.com/ild-thl/wisyki-api, 2023. Accessed: 2024-04-26.
[22] P. Huerten, Course Skill Retrieval Fine-tuned BI-Encoder Model, https://huggingface.co/isy-thl/multilingual-e5-base-course-skill-tuned, 2024. Accessed: 2024-07-17.
[23] P. Huerten, Course Skill Reranking Fine-tuned Cross-Encoder Model, https://huggingface.co/isy-thl/bge-reranker-base-course-skill-tuned, 2024. Accessed: 2024-07-17.
[24] N. Reimers, SBERT retrieve rerank, https://www.sbert.net/examples/applications/retrieve_rerank/README.html, 2024. Accessed: 2024-04-26.
[25] X. Guo, Y. Chen, Generative AI for Synthetic Data Generation: Methods, Challenges and the Future, arXiv cs.LG (2024). URL: https://arxiv.org/abs/2403.04190v1, cC BY 4.0 license.

https://doi.org/10.23106/dcmi.952436347