eleed e-learning & education

Issue se2023
2023

Towards a Standardized Machine-Readable Metadata Format for MOOC Platforms

Categorizing Fields of Study

Max Thomas
max.thomas@hpi.de

Thomas Staubitz
thomas.staubitz@hpi.de

Christoph Meinel
christoph.meinel@hpi.de

Hasso Plattner Institute
Prof.-Dr.-Helmert-Str. 2-3
Potsdam
Germany

urn:nbn:de:0009-5-57801

Zusammenfassung

Heutzutage sind viele Onlinekurse wie MOOCs (Massive Open Online Courses) verfügbar von verschiedenen Anbietern (z.B. edX, Coursera, openHPI, OpenWHO, iMOOX). Um Lernende zu unterstützen, wurden Aggregatoren wie Class Central oder MOOChub gegründet. Diese Aggregatoren enthalten Kataloge mit den Angeboten der Provider, was sie zu zentralen Eintrittspunkten für die Nutzenden macht. Diese Kataloge basieren auf Metadaten, die entsprechend formatiert sein müssen. Die Metadaten können dann auch zum Filtern von Kursen und von Empfehlungssystemen genutzt werden. Mit dem Aufkommen von mehr und mehr KI-basierten Empfehlungsservicen für Lerngelegenheiten und Lernpfadassistenten steigt der Bedarf an gut gepflegten und bedeutsamen Metadaten massiv an. In dieser Publikation stellen wir unsere Untersuchungen zu verschiedenen Systemen zur Kategorisierung des Studienbereichs, Themenbereichs oder Faches vor, die genutzt werden können, um existierende Metadatenformate zu verbessern. Eine Übersicht über Kategorisierungssysteme für Studienbereiche von verschiedenen Entitäten (z.B. international, national und private Organisationen) wird gegeben. Die Systeme werden in Hinblick auf ihren Nutzen in Metadatenformaten zur Beschreibung von Kursen verglichen. Die Ergebnisse werden genutzt, um unser Metadatenformat zu verfeinern und stellen einen weiteren Schritt hin zu einem standardisierten Metadatenformat für Kurse und automatisiert generierten Metadaten dar.

Stichwörter: e-learning; Metadaten; Standards; Studienbereich; MOOC; Massive Open Online Course

Abstract

Nowadays, there are many online courses like MOOCs (Massive Open Online Courses) available from different providers (e.g. edX, Coursera, openHPI, OpenWHO, iMOOX). To support learners, aggregators like Class Central or MOOChub were established. These aggregators hold catalogs with the offerings of the providers making them a central entry point for the users. Such catalogs are based on metadata, which needs to be formatted in a proper way. This metadata can then be used for filtering courses and recommendation engines also. With more and more emerging AI-based recommendation services for learning opportunities and learning path assistants, the need for well-maintained and meaningful metadata is growing massively. In this paper, we report on our research about different systems for categorizing the fields of study, topic, or subject, which can be used to enhance existing metadata formats. An overview of field of study categorization systems of different entities (e.g. international, national, and private organizations) is given. The systems are compared regarding their usefulness in metadata formats for the description of courses. The results are utilized to refine our own metadata format and represent a further step towards a standardized metadata format for courses and automatically generated metadata.

Keywords: e-learning; Metadata; Standards; Fields of Study; MOOC; Massive Open Online Course

1. Introduction

The first Massive Open Online Courses (MOOCs), in today’s predominant form of xMOOCs (more traditional teaching format than their predecessors, cMOOCs, which adhere to the connectivist learning theory (Kesim & Altinpulluk, 2015)), were offered by a group of professors at Stanford University in 2011 (Vardi, 2012). Already back in 2014, learners were confronted with a vast number of online courses in all imaginable subjects (Apaza et al., 2014). In general, this is a positive development as eager learners have a broad range of learning offers easily available. There is also competition among course providers that leads to a steady improvement of the offered material. However, at some point this offer became overwhelming and means of orientation need to be provided to the learners (Ibrahim, Yang, & Ndzi, 2017; Ibrahim et al., 2019).

These courses are provided by a wide range of different providers on different platforms. Before the Covid-19 pandemic, the number of providers slightly decreased, but lately, the market recovered again (Shah, Pickard, & Ma, 2022). The choice of platform is already a pre-filtering but might not be the filtering that is wished for. Consequently, several meta-platforms or aggregator systems appeared that combined the offers of multiple MOOC platforms. One of the earliest protagonists of this development was Class Central. In Europe, the MOOChub focusing on the German-speaking market or MOOCs4U combining the courses of the members of the European MOOC Consortium were established. There are similar efforts within the Open Educational Resources (OER) community such as the Open Educational Resources Search Index (OERSI), which connects OER repositories of universities and libraries (Klinger & Pohl, 2021). More recently, initiatives such as the digitale Vernetzungsinfrastruktur für die Bildung (German project connecting stakeholders from education and EdTech, DVB) in Germany or the Digital.Campus in the German federal state of Bavaria attempt to combine learning offers on an even broader spectrum than just MOOCs or OER.

As it becomes hard to navigate through this abundance of learning materials, it is exceedingly important to provide good filter and search options. Supporting the user to search for a course using filtering requires a sufficiently detailed set of relevant, well-defined metadata of the same quality throughout all courses, providers, and platforms (Stratakis et al., 2003). For plain filtering purposes, roughly sketched metadata sets can be sufficient or may even be better suited for human usage. For example, a level system that restricts itself to “beginner”, “advanced”, “expert” is more usable for a human user but does not have the required fine granularity for an automated recommendation service.

A well-defined and comprehensive metadata format is crucial for future AI-driven course recommendation services or systems that automatically create learning paths. Already a quick look into the literature—e.g. this non-exhausting list (Ibrahim, Yang, & Ndzi, 2017; Ibrahim et al., 2019; Gulzar, Leema, & Deepak, 2017; Guruge, Kadel, & Halder, 2021)—shows that this is becoming an increasingly important topic. Obviously, such services need access to cross-platform course catalogs for training and finding the best-suited offers. Ideally, the format of these catalogs is standardized in a way that all courses deliver the same set of metadata of the same quality. Unfortunately, many providers and aggregators are offering their catalogs in different formats. Therefore, a standard for course metadata is of fundamental interest to allow cross-platform functionality and (re-)search in this huge amount of data (Ruipérez-Valiente et al., 2022). Several initiatives are working to achieve this goal and, in our experience, the communication between them increases. The better standardized these formats are, the easier the development of search engines, learning path builders, and recommendation services.

2. Related Work

Various formats to deliver metadata exist already. In a very general approach, schema.org provides guidelines for the description of structured data on the web (Guha, Brickley, & Macbeth, 2016). Besides that the Dublin Core Metadata Initiative (DCMI) is dedicated to developing a metadata format for describing more universal resources and the ISO 15836 series of standards is based on the work of the DCMI (ISO 15836-1:2017; ISO 15836-2:2019).

These formats are also useful to describe online learning objects but might not always cover all the particular requirements. That is why more specific standards were developed on top of them, such as the international standard series ISO/IEC 19788. This standard is dealing specifically with metadata for learning resources (MLR) (ISO 19788-1:2011, 2011; ISO 19788-2:2011, 2011; ISO 19788-3:2011, 2011; ISO 19788-4:2014, 2014; ISO 19788-5:2012, 2012; ISO 19788-7:2019, 2019; ISO 19788-8:2015, 2015; ISO 19788-9:2015, 2015; ISO 19788-11:2017, 2017). Next to the ISO standard series, a variety of different formats, specifications, standards, and other guidelines on how to structure metadata of learning objects exists (Queirós & Leal, 2013; Bakhouyi et al., 2017). It is reasonable to focus on the most commonly used formats. The MLR, the Learning Object Model (LOM) (IEEE, 2020), the Sharable Content Object Reference Model (SCORM)—or better it’s successor xAPI (Experience API, Tin Can API) (Poltrack et al., 2012), which implements different standards and other formats (Parmar, 2012), and Dublin Core are widely known metadata formats.

Previous considerations about a common metadata format for MOOCs can be found here (Ebner et al., 2023).

3. Motivation

Especially providers of online learning opportunities are interested in a standardized metadata format describing MOOCs. Hence, our partners and we elaborated a metadata exchange format (MOOChub format, 2023), which developed into a de-facto metadata standard for MOOC platforms (openHPI; eGov-Campus; openSAP; iMOOX; openVHB; oncampus) and major aggregators in the German-speaking countries. It should be noted, that the attributes of this format are based on the "Course" schema of schema.org. For our use in the MOOChub API, we extended it and made it mandatory to list a course on MOOChub. Other platforms using our format are findig.sh, the Digital.Campus in Bavaria and MERLOT (a project within the GAIA-X digital ecosystem funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK)) (MERLOT, 2023). Moreover, just recently, the MOOChub format is officially supported by the DVB to integrate our educational offers there.

The underlying, freely accessible metadata format for OERSI, on the other hand, is the AMB (Allgemeines Metadatenprofil für Bildungsressourcen; en: General Metadata Profile for Educational Resources). AMB and the MOOChub format were developed independently but both are based on schema.org and they are compatible. Therefore, courses implementing the MOOChub format can be listed in OERSI.

At the time the MOOChub API format was developed, learning paths and recommendation services had not been in the focus. That is why the format currently lacks some information for training and usage of such AI-based tools. Discussions with other projects and partners in the context of MERLOT, MOOChub, and DVB revealed that most urgent required additional information are:

- course topics or fields of study,
- competencies,
- and difficulty levels.

The information on the fields of study is necessary for clustering courses and preselecting subsets of the data to improve the search performance. Information on the competencies conveyed by a course enables the matching of potential learners with suitable courses to close the competency gap. The same applies for the difficulty levels.

Considering the description of learning opportunities in different knowledge domains, it might appear that each domain requires a different, domain specific format. However, that would result in a high complexity and different formats within and across the platforms, which is exactly what we try to avoid. The main requirement is that we need to have a single but flexible format to serve all domains. From this further requirements for the metadata format can be derived:

- Every attribute value needs to be identified unambiguously.
- Every attribute value needs to be based on a framework that is accessible to the metadata consumer.
- Every attribute value needs to provide a meaningful description.

In the paper at hand, we present our investigation of a selection of existing categorization systems of “fields of study” (FoS) representing the subject or topic of a course. In our opinion, a proper categorization by FoS requires the usage of existing standards, allows for automated selection on the aggregator side, and can still be handled by humans on the platform side. Learners mostly select courses based on their learning objective (Ibrahim, Yang, & Ndzi, 2017), which is often defined by a FoS or a specific discipline (Ruipérez-Valiente et al., 2022). A representative survey on our platforms openSAP and openHPI (Figure 1) confirms this finding.

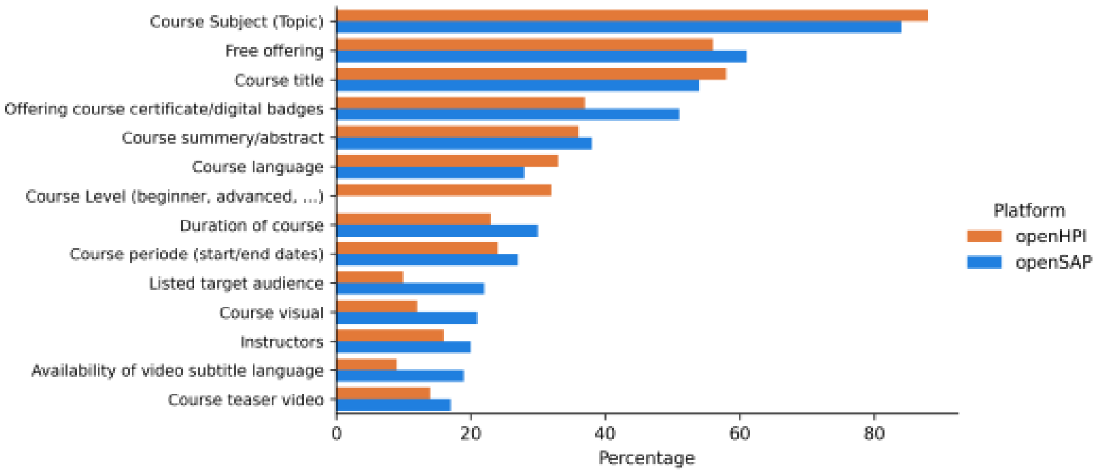

Figure 1: Survey on the platforms openSAP and openHPI why a user chose the course.

In the following, we summarized existing categorization systems of FoS with respect to their background, aims, and systematic/logic. We want to stress that every categorization of fields of study is merely arbitrary and there is no naturally given classification or separation of FoS as already stated in the literature (Krishnan, 2009).

4. Terminology

4.1 Standards, formats and systems

If we talk about format standards, we need to clarify what a standard is. The Cambridge Dictionary defines a standard as ”*a pattern or model that is generally accepted*”. This definition allows for a rather broad spectrum of formats that might be called standards. They are based on an agreement in a formal standardization process (formal standard) or simply by wide usage because it turned out to be the best way to deal with a certain problem or task (de-facto standard). Everyone can bring a solution into a formalized standardization process or providing a solution that many other stakeholders use. Within this document, we will use the term standard only for formal standards issued by officially recognized organizations (e.g. ISO, IEC, UNESCO). Besides these formal standards, we use the terms systems or formats even if these formats might be considered de-facto standards.

4.2 Hierarchical vs. Non-hierarchical

Hierarchical categorization systems can be represented by a tree structure. It shows a child-ancestor relationship of the different fields in which every field can have one or multiple sub-fields. As a typical example, we can take the ÖFOS (ÖFOS, 2012) categorization system where organic and inorganic chemistry are sub-fields of chemistry, which is a sub-field of the natural sciences. Sometimes the strictly tree-like structure is replaced by a graph-like structure with several ancestors for one field (cycles). Nevertheless, the hierarchy stays intact since the role of “child” and “ancestor” are clearly defined (directed graph).

Many FoS categorization systems are hierarchical because it allows easy navigation and fast finding of the FoS in question, if the user is familiar with the structure. A drawback of hierarchical systems is, that they are difficult to maintain. Modifications in the categorization system can have a strong impact on the structure.

Non-hierarchical categorization systems have no clearly defined child-ancestor-relationship. The FoS can be just gathered in a list. Finding a FoS can turn out quite challenging since there are no hints where to find it. In the worst case, it becomes necessary to iterate through the complete list. Some non-hierarchical categorization systems try to improve their searchability by using ordered lists. The order is then loosely correlated to the relationship of the FoS. A typical example of such a system is *HECoS* (HECoS, 2020).

Although the non-hierarchical categorization systems are much harder to be searched in general, they are also easier to maintain. Entries can just be added to the list or removed from it. There are no side effects like in the hierarchical system with its dependencies.

We consider mixed hierarchical/non-hierarchical systems to be hierarchical systems as the hierarchical component in such a hybrid system outweighs the non-hierarchical component. During our investigation, we also have not encountered any hybrid systems.

4.3 Levels

In this paper, we define the term “level” as the depth of the node holding the information of the FoS within a hierarchical system. Starting from the root (level 0) the level is identical to the depth. Some categorization systems use their own terms for the depth or level of a FoS e.g. broad, narrow, and detailed field. The reason for that is simply that the different hierarchical categorization systems have different tree heights. To that end, we will use the depth of the node as a ubiquitous definition and we will speak of level 1, 2 ... etc. fields or first, second, ... etc. level fields, respectively.

5. Methodology

To get an overview of already existing categorization systems used by different organizations, we first looked at other course platforms. This gave us a chance to analyze the FoS categorization systems of our partners. We checked the possibilities to filter the offered courses on their respective websites and investigated the categorization systems of two big international course platforms (Coursera, edX) and the members of the European MOOC Consortium (EMC): eduOpen, FutureLearn, Fun.MOOC, MiriadaX, OpenupEd, iMOOX and NAU.

In terms of official systems and standards, we examined the output of internationally recognized organizations in the area of education. The first investigated standard came from the United Nations Educational, Scientific and Cultural Organization (UNESCO). The "International Standard for Classification of Education" (*ISCED*) (ISCED, 2012) and the more FoS-specific extension "International Standard for Classification of Education - Fields" *(ISCED-F*) (ISCED-F, 2015) are the basis for European standards like the "European Qualification Framework" (*EQF*) and parts of the "European Skill, Competencies and Occupation" (*ESCO*) classification system. Since UNESCO is associated with the United Nations, it is a truly international standard with contributors all over the world and is designed to make different educational systems comparable. This makes it a good starting point for further investigations.

The next candidate with a large international impact was a standard from the Organization for Co-operation and Development (OECD). This organization has a strong focus on economic aspects but also publishes reports and other documents about educational systems. Furthermore, the expenses for research and development are investigated according to clearly defined guidelines published in the "Frascati Manual" (Frascati Manual, 2015). Similar to the *ISCED*, a supporting document for the Frascati Manual is published dedicated to the categorization of FoS (OECD Fields of Science and Technology, 2007).

After investigating these international standards, we focused on national standards and systems. Since educational programs differ between countries, national categorization systems are needed to fit the local special environment (Luijkx & de Heus, 2008; ÖFOS, 2012). For Germany, we elaborated an overview of the school subjects in the 16 federal states. We analyzed the curricula published by the respective Ministry of Education and Cultural Affairs on their official websites and looked into the system of fields of studies for institutes of higher education published by the Federal Statistical Office of Germany (destatis). In addition, the "Hochschulkompass" was investigated where state and state-recognized institutions of higher education of Germany to upload their course program catalogs. Together with the Austrian *ÖFOS* (Österreichische Systematik der Wissenschaftszweige; en: Austrian System of Academic Fields) published by Statistics Austria (Bundesanstalt Statistik Österreich), we established a comprehensive overview of FoS categorization systems in the German-speaking area.

We limited our investigation of categorization systems in other countries to higher education. We examined the national systems in

- France (INSEE – l'Institut national de la statistique et desétudes économiques, CNRS – French National Center for Scientific Research),

- Italy (Istat – Italian National Institute of Statistics),
- Spain (INE – National Statistics Institute),
- the United Kingdom (HESA – Higher Education Statistics Agency),
- the Netherlands (CBS – Central Bureau voor de Statistiek),
- and Portugal (Portuguese Government/OECD and FCT – Fundacao para a Ciencia e a Tecnologia).

We focused on these countries, as all of them are represented via a MOOC platform in the European MOOC Consortium (EMC). Additionally, the systems of the United States of America (NSF – National Science Foundation) and Canada (both NCES – National Center for Education Statistics) were investigated to widen the scope. Most of these categorization systems are provided by entities interested in statistical investigations of the education system of their respective country.

Finally, we searched the web for other organizations beyond the official governmental or administrative entities. As everyone can (and does) invent a categorization system of their own, we stopped this effort at a certain point as the results were too many.

Nevertheless, we included the *List of Academics Fields* (*LAF*) from Wikipedia, the *Microsoft Academics Graph* (*MAG; now openAlex*), the Dewey Decimal Classification (*DDC*) and the Library of Congress Classification (*LCC*). The latter two are important bibliographic categorization systems for FoS.

6. Analysis of other platforms

If we want to establish a standard for MOOCs, it makes sense to analyze how other MOOC platforms categorize their FoS (Table 1, Figure 2). These platforms offer between 70 and 12000 courses each (data from September 2022). The number of courses, however, does not correlate with the number of FoS and a platform with many courses might only provide a few FoS for categorization (Coursera: 11025 courses, 11 “subjects”), while some providers with fewer courses might have a much more detailed categorization (e.g. iMOOX: 270 courses, 35 FoS).

It is noteworthy, that Coursera does not only have a categorization by “subject” but also by “skill” (809 “skills”). A course on Coursera can be tagged with multiple subjects and skills to refine your search. The user can search only by “subject”, only by “skill”, or combine both leading to an inclusive disjunct result set (logical OR). We will only deal with the “subject” in this paper (skills and competencies frameworks will be the topic of a separate work).

This is different from other platforms like iMOOX with a hierarchical tree-like structure. iMOOX uses the national Austrian *ÖFOS* categorization system as the bedrock to define the category of their courses. However, they are only using the first two of the four levels of *ÖFOS*. Besides iMOOX, only FutureLearn uses a hierarchical categorization system.

Within the set of examples, only iMOOX is using a conjunct system that allows refining the user’s search by filtering by multiple FoS. That means the more filters a user applies the fewer results are given (the ones that match all filters; logical *AND*). In contrast, the filter systems of Coursera, edX, funMOOC, and NAU use inclusive disjunction (logical OR). Therefore, the number of courses offered to the user increases the more filters are applied. EduOpen, FutureLearn, and MiriadaX allow the users to choose only one field for filtering: a combination of FoS filters is not possible.

Provider	Fields[a]	Hier.[b]	Levels	Allow refinement by multiple tags
Coursera sub-jects	11	No	1	No (inclusive disjunction)
eduOpen	6	No	1	No (only one field)
edX	31	No	1	No (inclusive disjunction)
funMOOC	26	No	1	No (inclusive disjunction)
FutureLearn	124	Yes	2	No (only one field)
iMOOX	35	Yes	2	Yes (conjunction)
MiradaX	24	No	1	No (only one field)
Nau	10	No	1	No (inclusive disjunction)
[a]All data received September 2022, [b]Hierarchical				

Table 1: Overview of key data of MOOC platforms regarding FoS.

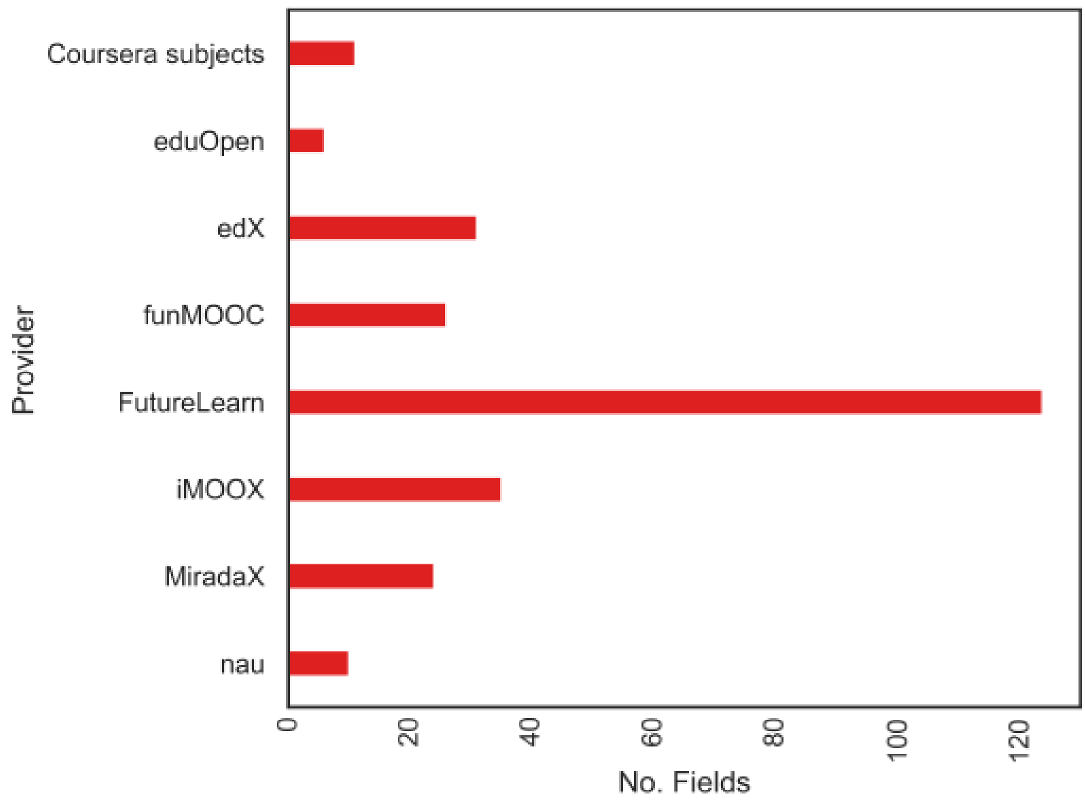

Figure 2: The number of FoS that can be assigned in the respective categorization system.

7. International systems and standards for categorizing FoS

After elaborating on an overview of FoS categorization systems of the different MOOC providers, we were interested in official categorization systems on this topic (see Table 2). First, we examined international standards supported by official administration and governmental entities.

A widely and internationally recognized classification system for FoS is the *ISCED-F*, which has been established in 2013 by the UNESCO Institute for Statistics (*ISCED-F*, 2015). *ISCED-F* and *ISCED* serve as the basis for e.g. *ESCO* and aim to make education systems comparable worldwide. The supplement *ISCED-F* is a document, especially for the categorization of FoS. The 148 different fields are under steady revision and allow cross-field classification.
Assigning multiple FoS to a learning opportunity should not be done according to the standard and every FoS has to be represented by a four-digit code (no reference to a first-level field possible). While the first two digits (including a leading zero if applicable) define one of the eleven first-level fields, the following two digits describe the second-level field. For example: “0788” - Inter-disciplinary programs and qualifications involving engineering, manufacturing, and construction (“07” - Engineering, manufacturing and construction).

The latest issue of the *OECD Field of Science and Technology* (*FOS*) *Classification* (OECD Fields of Science and Technology, 2007) defines 42 FoS divided into six first-level fields. The idea behind this classification system is to create a standard for statistical analysis of e.g. R&D expenditure. As in the *ISCED-F*, it is not allowed to assign more than one FoS; the rule is to choose the FoS by the majority of the content. The coding logic demands choosing a level 1 field – represented by a number between 1 and 6 – followed by a dot and the number representing the second level field. Example: “3.2” – Clinical medicine (“3” – Medical and Health Sciences). More recently, the OECD also published a quite similar classification to FOS in the Frascati Manual (latest issue from 2015) with the title “Fields of Research and Development” (*FORD*) (Frascati Manual, 2015). It was originally developed to support administration, industry, and scientific communities to categorize and organize their R&D resources. In theory, it can be used also to classify courses or educational programs. The coding logic and restrictions are the same as for the *FOS*.

Name	Fields	Hierarchical	Levels	Issuer	Issuing date
ISCED-F	148	yes	2	UNESCO	2013
FOS	42	yes	2	OECD	2007
FORD	42	yes	2	OECD	2015

Table 2: Overview of standards and categorization systems for FoS on the international level.

8. National systems and standards for categorizing FoS

Besides the international standards, there are also national approaches to define categorizations of FoS (Table 3, Figure 3). Interestingly, national categorization systems are more detailed if they are intended to describe educational systems. Systems taking a closer look at expenses for R&D purposes are often simpler.

8.1 Germany

The Federal Statistical Office of Germany (destatis) has published an overview of FoS in institutions of higher education (Fächersystematik for students, 2021). A four-digit code is used to assign an object to one of the 273 possible FoS. There is no coding logic after some revisions and a systematic search in the tree structure is not possible. Every FoS belongs to one of the 10 level 1 fields (“Systematik der Fächergruppe”) which are further divided into level 2 fields (“Systematik des Lehr- und Forschungsbereichs”). The third and last level of the system is the FoS (“Fachgebiet”). The destatis system was designed to serve as a guideline for further statistical investigations about the higher education system.

Example: “0810” Allg. und vergleichende Sprachwissenschaft (“01” – Geisteswissenschaften; “080” - Allgemeine und vergleichende Literatur- und Sprachwissenschaft).

We also tried to get an overview of the curricula in K-12 schools and found 342 school subjects (unique subject names as assigned by the respective ministry of education). The problem with the school subject is that one subject might exist in one federal state but does not have a counterpart in another one neither by name nor by content. In addition, subjects with the same content might be named differently or have different content under the same name. Such a categorization system has always to be connected to the respective federal state. For the school FoS there is no hierarchical system and a course might be tagged with and assigned to one or more FoS.

8.2 Austria

National categorization standards for FoS are often based on international standards and are modified for special national purposes. Austria, for example, has taken FOS from the OECD as a basis for their national system ÖFOS (ÖFOS, 2012).

This modified standard is significantly finer granular and more detailed than the original OECD system. It still has six level 1 fields (identical to FOS) but instead of 42 FoS organized in two levels, ÖFOS provides 1362 fields in four levels. Every FoS has to be described by a six-digit code. Shorter codes (referencing the higher levels) and giving a learning opportunity several assignments are not allowed. The first digit of the code describes the “Main group” (level 1) and the following two digits are the “Group” (level 2) within this main group. The fourth digit defines a “Sub-group” (level 3) and the last two digits define the “Field” (level 4). For example: “302055” Oncology (“3” - HUMAN MEDICINE, HEALTH SCIENCES; “302” – Clinical Medicine; “3020” – Clinical Medicine).

8.3 United Kingdom

The HESA (Higher Education Statistics Agency) of the UK provides two different FoS categorization systems. The “Higher Education Classification of Subjects” (*HECoS*, 2020) is a purely tag-based system with no hierarchy. Learning opportunities can be assigned to several of the 1092 possible FoS. There is no explicit coding logic but it is emphasized that the more or less randomly assigned codes can hint at loose relationships between the FoS. For example, close numbers can point to a cluster of similar or connected fields (example given on the *HECoS* homepage: *“the codes 100036 (optometry), 100037 (orthoptics), 100261 (ophthalmology) and 101511 (ophthalmic dispensing) are all related”*). *HECoS* is also the basis of the “Common Aggregation Hierarchy” (CAH) system (HESA, 2021). As the name already states, it is a hierarchical categorization system, which allows a more structured classification. In the latest version from 2021, there are 232 FoS under 27 level 1 categories (including a “non-applicable” branch). It was developed to improve *HECoS* and remove some inconsistencies. The coding logic follows a pattern, where all codes start with “CAH” defining the framework. It is directly followed by two digits (including a leading zero if applicable; or Z5 for “not applicable”) defining the first level field. The next number after a hyphen (“-”) specifies the second level field followed by another hyphen. Then, the last number is representing the third level (FoS). Example: “CAH19-03-04” Iberian studies (“CAH19” – language and area studies; “CAH19-03” – languages, linguistics and classics). Interestingly, some FoS are existing more than once under different broader fields and there are only 172 unique fields by name. We still consider it to have 232 FoS by unique paths/codes.

8.4 United States and Canada

The National Center for Education Statistics (NCES) as part of the United States Department of Education published the “Classification of Instructional Programs” (*CIP*) (CIP, 2020). The background of this classification is a taxonomic coding scheme for statistical reasons and an overview of the education system. With 2143 fields under 48 level 1 fields (“CIP family”) this system is quite detailed. Every learning opportunity can only have one FoS assigned depending on which category covers the majority of the content. The FoS is coded by six digits. The first two digits (including a leading zero, where necessary) define the first-level field. A dot and two digits for the level 2 field and two digits for the level 3 field (including a leading zero, where necessary) follow. Example: “54.0107” Canadian History (“54” – HISTORY; “54.01” - History). The CIP system is used in Canada, too. Placeholders in the original data stating: “Reserved for use by Statistics Canada.” were omitted for counting the FoS.

Besides CIP, the National Science Foundation (NSF) of the US has published a closely connected categorization system called “Survey of Earned Doctorates” (SED) (Fiegener, 2013). The version from 2013 has 303 fields under four level 1 fields and an hierarchical structure using another methodology for clustering then CIP. In SED the FoS are roughly divided into Science and Engineering and Non-science and Engineering. There is no real coding logic behind the numbers assigned to the FoS. Every field gets a three-digit code (including leading zeros if applicable) which points loosely towards a specific broader field.

8.5 France

The statistics bureau of France (INSEE) offers a categorization system designed to classify jobs (NAF, 2008). This system is called “Nomenclature d’activités française” (en: “French classification of activities”, NAF) and can be used also for categorizing courses. The original FoS categorization system was published in 2008 and reviewed in 2015. The 21 level 1 fields (“sections”) are coded by the letters A to U. The second level (“division”) is numbered consecutively by a two-digit code (including a leading zero, if applicable) but have no marking to which section they belong. A single digit separated from the division by a period represents the third level (“group”). The group is followed by a “class” (fourth level) which is also represented by a single digit. It follows directly the digit for the group without a separator. The fifth level (“sub-class”) is represented by a capital letter at the end of the coding. In the “sub-class” level 729 FoS can be assigned. Example: “80.20Z” Security systems service activities (80 - Security and investigation activities; “80.2” - Security systems service activities; “80.20” - Security systems service activities. This FoS belongs to Section N: “Administrative and support service activities” but this information cannot be derived directly from the code.).

The Centre National de la Recherche Scientifique (CNRS) from France uses a very simple categorization system. It is a non-hierarchical system with only 10 possible FoS. Just the name is used and there is no coding. CNRS uses this system for structuring its research fields.

8.6 Spain

Very similar to the French CNRS is the system used by the Spanish Instituto National de Estatistica (INE) to describe the doctorate studies by field (INE Fields of Doctoral Studies, 2009). Here, only six FoS can be assigned without any hierarchy. It also works without coding and only the name is used to define a FoS. From the same institute with a focus on R&D, there is another system with 14 FoS (Cristóbal, del Moral, & Olmos, 2013).

8.7 Italy

The Italian Instituto Nazionale di Statistica (Istat, 2022) works with a non-hierarchical system of 146 FoS. One could conclude a hierarchical order by coding. The six-digit code could be seen as a very loose clustering but is not really strict in assigning the FoS to a broader field. Like the others, this system is intended for use in educational statistics.

8.8 Netherlands

In contrast to the aforementioned systems, the Dutch “Standaard Onderwijsindeling” (en: “Standard for Classification of Education”, *SOI*) (Schaart, Moens, & Westerman, 2008) is related to *ISCED* but the FoS classification in this standard differs from *ISCED-F* to align with national requirements. SOI relies on a hierarchical structure with four levels: level 1 “sector groups”; level 2 “subsector groups”; level 3 “groups”; level 4 “minor groups”. In total, there are 454 “minor groups” to which a program can be assigned in the SOI version published in 2008. SOI is intended to support statistical investigations on the Dutch education system. Example: 4123 - Financieel beheer (4 - Juridisch, bestuurlijk, openbare orde en veiligheid; 41 - Juridisch, bestuurlijk; 412 - Openbaar bestuur; the detailed version of SOI is only available in Dutch).

8.9 Portugal

We found two different categorization systems from the Ministry of Science, Technology and Higher Education. The first document was created to shed light on research activities in Portugal and elaborate a strategy for further development. This “Evaluation on Research Units” comprises 24 FoS (FCT, 2005). It is not directly intended to be a hierarchically structured system but clustering the table of contents into six level 1 fields can be interpreted that way. The fields are numbered 1 to 24 and there is no coding logic. The second document from 2006 is a background report for an OECD assessment (OECD Background Report Portugal, 2006). In this document, the overall development of Portugal in terms of education is investigated. All enrolments at universities are analyzed leading to an overview of enrolled students by FoS. It is a simple hierarchical system with two levels with 23 FoS under 8 first-level categories.

Name	CC[a]	Fields	Hier.[b]	Lev.[c]	Issuer	Date[d]
Hochschulfächersystematik	DE	273	Yes	3	destatis	2021
School subjects[e]	DE	124	No	1	Ministries of Education	2022[f]
ÖFOS	AT	1362	Yes	4	Statistik Austria	2012
CAH	UK	232	Yes	3	HESA/Jisc	2021
HECoS	UK	1092	No	1	HESA	2020
CIP	US/CA	2143	Yes	3	NCES	2020
SED	US	303	Yes	5	NSF	2013
NAF	FR	729	Yes	5	INSEE	2015
CNRS	FR	10	No	1	CNRS	2022
INE Doctoral Studies	ES	6	No	1	INE	2022
INE expenditure on R&D	ES	14	No	1	INE	2013
Istat Course classification	IT	146	No	1	Istat	2007
SOI	NE	454	Yes	4	Statistics Netherlands	2008
Tertiary Education in Portugal	PT	23	Yes	2	MCTES	2006
Evaluation of Research Units	PT	24	Yes	2	MCTES	2005

[a] Country Code According to ISO 3166.
[b] Hierarchical
[c] Level
[d] Issue Date - The date the latest version was issued to the best of our knowledge. The first publication might be older.
[e] The German school system is very heterogeneous. All Data has to be seen as an overview of all 16 federal states of Germany.
[f] Year the data was received. The original data from the different Ministries of Education might be older.

Table 3: Overview of standards and categorization systems for FoS on the international level.

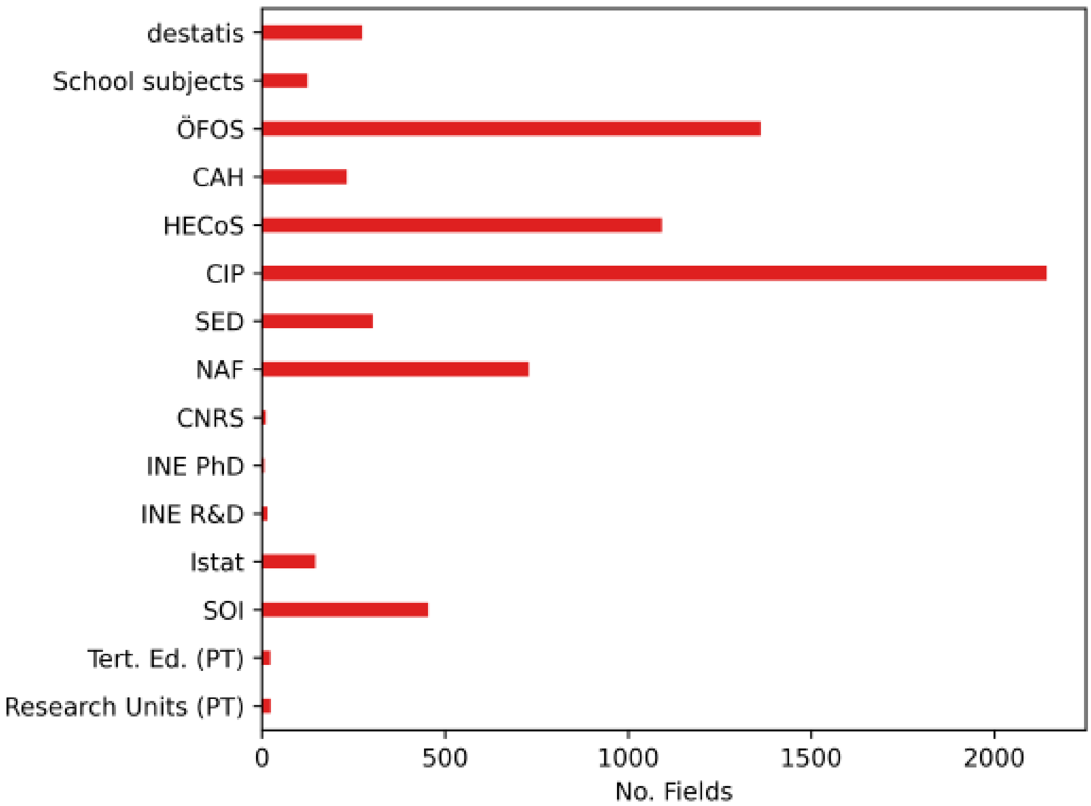

Figure 3: Comparison of potential numbers of FoS in different national categorization systems.

9. Other systems and standards for categorizing FoS

Not only official entities tried to find definitions for the FoS categorization but also private organizations. E.g., the Microsoft Cooperation has been working with the Microsoft Academics Graph (*MAG*) (Sinha et al., 2015) which was converted into *OpenAlex* in 2022 as a catalog of scholarly papers. The field ("concept") description from its respective Wikipedia article is taken as a basis. There are 65072 FoS making it the most detailed categorization system in this review. There are 19 level 1 fields (L0 concepts) and multi-tagging is allowed as well as referencing any level. A query on *OpenAlex* with a filter for concepts and a level greater than five led to an empty result list. Therefore, there are six levels in total (L0 to L5). There is no real coding logic and the resource is simply tagged by the name. The graph behind this classification is under steady revision (data from June 2022).

We also investigated the Wikipedia entry to academic fields (*LAF*) which could be used in a hierarchical manner. There are no rules for applying *LAF* as an FoS categorization system. It is simply a list of what is considered an academic field in the Wikipedia community. Nevertheless, the 2137 fields are clustered under the four main groups: Humanities and social science, Formal sciences, Natural sciences, Professions and applied sciences. As in *OpenAlex*, the FoS are just added as plain text to the learning opportunity.

Another approach is to list all existing university course programs in Germany. We found 9404 unique (by title) course programs according to the “Hochschulkompass”. Naturally, this leads to a non-hierarchical tagging-by-name system without coding. The purpose of the “Hochschulkompass” is to give guidance to people interested in studying.

Here, we also need to name two bibliographic systems: the Dewey Decimal Classification (*DDC*) and the Library of Congress Classification (*LCC*). Both are widely used in libraries around the world. While *DDC* provides 10 level 1 fields (numbers 0 to 9), the *LCC* uses letters in the first level (21 fields; “Main Class”). *DDC* is a logical 1000 FoS system representing the main classification with three numbers between 0 and 9: XYZ (X = first level, Y = second level, Z = third level; e.g. 613 - “Personal health and safety”). To specialize the FoS further, more numbers can be added after a “.” (e.g. 613.7046 - “Physical yoga”). Due to the high complexity, we will stick to the three main levels in this paper. In contrast to that, the *LCC* uses mostly letters for the second level leading to two-letter or three-letter codes (e.g. K - “Law”, main class and KF - “United States”, sub-class). After that, there are numbers coding the most detailed levels. Sometimes the numbering starts directly at the second level. Taking into account the strong variations in the *LCC*, it is unsuited for our purpose.

Name	Fields	Hier.[a]	Levels	Issuer	Date[b]
OpenAlex	65 072	yes	6	OurResearch	2022
List of Academic Fields	2 137	yes	6	Wikipedia	2022
Hochschulkompass	9 404	no	12	HRK[c]	2022
LCC	230[d]	yes	2[d]	Library of Congress	2023
DDC	1 000[d]	yes	3[d]	OCLC[e]	2023

[a] Hierarchical
[b] The date when the data was received
[c] HRK: Hochschulrektorenkonferenz (eng. Conference of Highschool Deans, Germany)
[d] main classification, adoption may lead to different versions of the system
[e] Online Computer Library Center

Table 4: Overview of standards and categorization systems of FoS by private organizations.

10. Pros and Cons of the different FoS categorizations for implementation

To determine which categorization systems are suitable for classifying FoS of learning opportunities like MOOCs, we need to clarify what are the relevant parameters.

One question was how detailed learning opportunities should be described. It is important that learners, teachers and administrators can understand the categorization system and find the FoS for their learning opportunities. The fields should not be too general but also not too detailed and complicated to use. Administrators and teachers need to know what to insert into the respective metadata field and learners need to find their desired learning opportunities by applying the filters correctly. So, we need to find a middle ground of granularity when looking at human limitations. For software-based recommendation services, this limitation is not applicable. Here, also very complex systems can be processed if the relevant data is machine-readable. To that end, we are looking for categorization systems providing data in a structured way and, preferably, with detailed information about the FoS.

A parameter describing the granularity is the number of FoS that can be assigned. The more FoS possible the more complex the assignment. Another indicator is the number of levels. With increasing numbers of FoS more levels should be provided to give structure to the growing amount of fields.

For MOOC providers it seems to be sufficient to define only a few, very broad fields (between 6, eduOpen and 121, FutureLearn). With this small number of FoS they use non-hierarchical structure or only two levels. The international systems are often designed in a more general fashion and have between 42 and 146 FoS organized in two levels. This broad structure takes into account the differences in the various education systems worldwide and allows for flexible adjustment to national needs. Therefore, it can be the basis for national systems, is still compatible with other education systems, and makes them comparable. The national FoS classification systems are often more detailed to depict the national and cultural specialties. Many of these systems are very structured with up to 5 levels.

The hierarchical national and international categorization systems are generally better suited for (AI) recommendation services. Often the data is structured and there are descriptions of the FoS available. In addition, developers of recommendation engines made it clear that they prefer to rely on categorization systems recognized by official entities (We had several meetings with partners from projects like DVB and MERLOT). Nevertheless, there are no hard statistics supporting this or that classification system for every system has its use case. In general, every herein investigated classification system can be supported and it is a political decision between the organizations who operate the learning platforms, the recommender engines, and the aggregators.

One should keep in mind the process of how the FoS categorization system was elaborated and revised since the world of FoS is under persistent development. The systems from official administration entities are often created and revised by a transparent process with public revision schedules and defined protocols. Many institutions offer to listen to stakeholders and gather feedback. For private organizations, it is often unclear to the public how a system is created. Changes in these systems might be published unexpectedly. This is the same problem with Wikipedia's list of academic fields. Everyone can alter the list at any time. Hence, we decided to use open public systems with transparent elaboration and revision processes primarily.

11. Which standards and categorization systems to implement

For the acceptance of our exchange format, it will be exceedingly essential to design it open for the needs of the platforms, aggregators, lectures, and users. The strategy is to implement FoS systems on-demand. Everyone is invited to suggest an extension of the current set of supported categorization systems but the presented considerations might be valuable guidelines. In general, every categorization system for FoS can be supported but we want to provide an initial selection of FoS classification systems. Otherwise, we fear everyone adds his or her own systems and the data gets incomprehensive for further use. This would be contradictory to our standardization intention.

One classification system we will include from the first day on is *ISCED* since it is an internationally approved standard and basis for other classification systems. It can be assumed that many parties interested in using our metadata exchange format can directly apply it. To support our idea of openness, we will implement *ÖFOS* as asked for by our partners and support it from the beginning.

Both systems meet the central requirements we elaborated (s. Motivation). They allow for unique identification of the FoS by providing shortcodes, which encode the complete path through the hierarchical system. The tree-like structure enables easy navigation through the system. With 146 or 1362 possible fields, respectively, both systems allow sufficiently detailed categorization. Additionally, the documentation of both systems is publicly available free of charge. However *ÖFOS* lacks a detailed descriptions of the FoS at the moment.

The decision to flexibly support a choice of systems directly influences the design of our metadata standard. For adding an FoS to the metadata the name and framework have to be given. Furthermore, a shortcode and a URI can be included for unique identification and a description can be written into the metadata. This way almost any FoS categorization systems are technically supported.

12. Conclusion

We have presented our findings for improving our current format according to the requirements. It also deals with the challenges of next-generation recommendation systems for learning paths.

After reviewing a non-exhausting selection of FoS categorization systems, *ISCED-F* was chosen because it is an international standard issued by the UNESCO and the fundament of many other systems. As another FoS system, we will support *ÖFOS* since it is a requirement by one of our partners and recognized by official Austrian entities. They follow a transparent and strictly rule-based revision process. Moreover, they allow for detailed selection of the FoS without getting lost in an extremely fine granular structure.

These standards can be easily implemented (e.g. JSON) to be used for the description of MOOCs and can be used for the design of individual learning paths. This is only possible with automated tools based on standardized FoS categorization systems. So, administrators of the platforms and courses can easily maintain the metadata. In the future, the selection of the FoS will be supported by AI and are a further step towards automatically generated metadata. This is another reason why standardized frameworks for FoS are needed.

Including the new field for FoS will be a useful addition to the already existing MOOChub API and will be of help to develop it further to a standard metadata format for course catalogs. To align our API format with the course schema with schema.org, we chose to add the information about FoS in the educationAligment field. This will not only allow for storing which FoS a course belongs to but also give some additional data like an URL to the standard reference.

Acknowledgements

Funding by the BMBF (German Federal Ministry of Education and Research) via the project HPI4NBP (16INB2029) is gratefully acknowledged.

References

Apaza, R. G.; Cervantes, E. V.; Quispe, L. C.; Luna, L. O.: Online Courses Recommendation based on LDA. In: SIMBig 2014 - 1st Symposium on Information Management and Big Data. CEUR-WS.org, Cusco, Peru, 2014, pp. 42–48. https://ceur-ws.org/Vol-1318/paper5.pdf (last check 2023-09-04)

Bakhouyi, A.; Dehbi, R.; Talea, M.; Hajoui; O.: Evolution of standardization and interoperability on E-learning systems: An overview. In: 2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET). IEEE 2017, Ohrid, Macedonia, pp. 1–8. https://doi.org/10.1109/ITHET.2017.8067789 (last check 2023-09-04)

CIP (2020). https://nces.ed.gov/ipeds/cipcode/resources.aspx?y=56 (last checked 2023-09-04)

Cristóbal, A., C.; del Moral, A. G.; Olmos, B. G.: Proyecto para la capitalizaci´on del gasto en I+D en los nuevos sistemas de cuentas nacionales: estimaci´on de su impacto sobre el PIB y compilaci´on de una cuenta sat´elite de I+D. Instituto National de Estatistica, Madrid, 2013. https://www.ine.es/ss/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3B+filename%3Ddoctrabajo2-2013.pdf&blobkey=urldata&blobtab (last checked 2023-09-04)

Ebner, M.; Koschutnig-Ebner, M.; Rampelt, F.; Serth, S.; Staubitz, T.; von Stetten, A.; Thomas, M.; Wittke, A.: Metastandard für den internationalen Austausch von MOOCs – der MOOChub als erster Prototyp. In: Zeitschrift für Hochschulentwicklung, 18(1), 2023, pp. 17–35. https://doi.org/10.3217/zfhe-18-01/02 (last checked 2023-09-04)

FCT: Evaluation of Research Units 2002-2004 Overall Report. Fundação para a Ciência e a Tecnologia (FCT) - Ministério da Ciência, Tecnologia e Ensino Superior, Lisboa, 2005.

Fiegener, M. K.: Science and Engineering Degrees: 1966–2010 (SED). National Science Foundation, Alexandria, VA, 2013.

Frascati Manual: Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development. OECD Publishing, Paris, 2015. https://doi.org/10.1787/9789264239012-en (last checked 2023-09-04)

Fächersystematik for students: Bildung und Kultur - Studierende an Hochschulen - Fächersystematik. Statistisches Bundesamt (Destatis), Berlin, 2021. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/_publikationen-innen-hochschulen-studierende-endg.html (last checked 2023-09-04)

Guha, R. V.; Brickley, D.; Macbeth, S.: Schema.org: Evolution of Structured Data on the Web. In: Commun. ACM, 59, 2016, pp. 44–51. https://doi.org/10.1145/2844544 (last checked 2023-09-04)

Gulzar, Z.; Leema, A.; Deepak, G.: PCRS: Personalized Course Recommender System Based on Hybrid Approach. In: 6th International Conference on smart computing and communications ICSCC. Elsevier 2017, Kurukshetra, India, pp. 518–524. https://www.sciencedirect.com/science/article/pii/S1877050917328314?via%3Dihub (last checked 2023-09-04)

Guruge, D. B.; Kadel, R.; Halder, S. J.: The State of the Art in Methodologies of Course Recommender Systems—A Review of Recent Research. Data, 6, 2021, pp. 1–30. https: //doi.org/10.3390/data6020018 (last checked 2023-09-04)

HECoS (2020). https://www.hesa.ac.uk/support/documentation/hecos (last checked 2023-09-04)

HESA (2021). https://www.hesa.ac.uk/support/documentation/hecos/cah (last checked 2023-09-04)

Ibrahim, M. E.; Yang, Y.; Ndzi, D.: Using ontology for personalised course recommendation applications. In: Computational Science and Its Applications – ICCSA 2017: 17th International Conference. Springer International Publishing AG 2017, Triest, Italy, pp. 426–438. https://doi.org/10.1007/978-3-319-62392-4_31 (last checked 2023-09-04)

Ibrahim, M. E.; Yang, Y.; Ndzi, D. L.; Yang, G.; Al-Maliki, M.: Ontology-Based Personalized Course Recommendation Framework. IEEEAccess, 7, 2019, pp. 5180–5199. https://doi.org/10.1109/ACCESS.2018.2889635 (last checked 2023-09-04)

IEEE (2020). IEEE Standard for Learning Object Metadata. IEEE 1484.12.1-2020, 10.1109/IEEESTD.2020.9262118. https://ieeexplore.ieee.org/document/9262118 (last checked 2023-09-04)

INE Fields of Doctoral Studies: Survey on Human Resources in Science and Technology. Year 2009, Characteristics of the Doctorate, Percentage of doctor-holders, by field of the doctoral studies and sex. Instituto National de Estatistica, Madrid, 2009.

ISCED: International Standard Classification of Education (ISCED) 2011. UNESCO Institute for Statistics, Montreal, 2012. https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf (last checked 2023-09-04)

ISCED-F: International Standard Classification of Education - Fields of education and training 2013 (ISCED-F 2013) Detailed field descriptions. UNESCO Institute for Statistics, Montreal, 2015. https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-fields-of-education-and-training-2013-detailed-field-descriptions-2015-en.pdf (last checked 2023-09-04)

ISO 15836-1:2017: Information and documentation — The Dublin Core metadata element set — Part 1: Core elements. International Organization for Standardization, Geneva, 2017. https://www.iso.org/standard/71339.html (last checked 2023-09-04)

ISO 15836-2:2019: Information and documentation — The Dublin Core metadata element set — Part 2: DCMI Properties and classes. International Organization for Standardization, Geneva, 2019. https://www.iso.org/standard/71341.html (last checked 2023-09-04)

ISO 19788-1:2011: Information technology — Learning, education and training — Metadata for learning resources — Part 1: Framework. Standard. International Organization for Standardization, Geneva, 2011. https://www.iso.org/standard/50772.html (last checked 2023-09-04)

ISO 19788-2:2011: Information technology — Learning, education and training — Metadata for learning resources — Part 2: Dublin Core elements. Standard. International Organization for Standardization, Geneva, 2011. https://www.iso.org/standard/46157.html (last checked 2023-09-04)

ISO 19788-3:2011: Information technology — Learning, education and training — Metadata for learning resources — Part 3: Basic application profile. Standard. International Organization for Standardization, Geneva, 2011. https://www.iso.org/standard/52774.html (last checked 2023-09-04)

ISO 19788-4:2014: Information technology — Learning, education and training — Metadata for learning resources — Part 4: Technical elements. Standard. International Organization for Standardization, Geneva, 2014. https://www.iso.org/standard/52775.html (last checked 2023-09-04)

ISO 19788-5:2012: Information technology — Learning, education and training — Metadata for learning resources — Part 5: Educational elements. Standard. International Organization for Standardization, Geneva, 2012. https://www.iso.org/standard/52776.html (last checked 2023-09-04)

ISO 19788-7:2019: Information technology — Learning, education and training — Metadata for learning resources — Part 7: Bindings. Standard. International Organization for Standardization, Geneva, 2019. https://www.iso.org/standard/60613.html (last checked 2023-09-04)

ISO 19788-8:2015: Information technology — Learning, education and training — Metadata for learning resources — Part 8: Data elements for MLR records. Standard. International Organization for Standardization, Geneva, 2015. https://www.iso.org/standard/62845.html (last checked 2023-09-04)

ISO 19788-9:2015: Information technology — Learning, education and training — Metadata for learning resources — Part 9: Data elements for persons. Standard. International Organization for Standardization, Geneva, 2015. https://www.iso.org/standard/62846.html (last checked 2023-09-04)

ISO 19788-11:2017: Information technology — Learning, education and training — Metadata for learning resources — Part 11: Migration from LOM to MLR. Standard. International Organization for Standardization, Geneva, 2017. https://www.iso.org/standard/65587.html (last checked 2023-09-04)

Kesim, M.; Altinpulluk, H.: A Theoretical Analysis of Moocs Types from a Perspective of Learning Theories. In: Proceedings of 5th World Conference on Learning, Teaching and Educational Leadership, 186, 2015, pp. 15–19. https://doi.org/10.1016/j.sbspro.2015.04.056 (last checked 2023-09-04)

Klinger, A.; Pohl, A.: OERSI – Die Suche nach OER für die Hochschullehre. 2021. https://open-educational-resources.de/oersi-die-suche-nach-oer-fuer-die-hochschullehre/ (last checked 2023-09-04)

Krishnan, A.: What are Academic Disciplines? Some observations on the Disciplinarity vs. Interdisciplinarity debate. 2009. https://eprints.ncrm.ac.uk/id/eprint/783/ (Last checked 2023-09-04)

Luijkx, R.; de Heus, M.: The educational system of the Netherlands. In: Schneider, S.L. (Ed.): The international standard classification of education (ISCED-97). An evaluation of content and criterion validity for 15 European countries. Mannheimer Zentrum für Europäische Sozialforschung, Mannheim, 2008, pp. 47–75. https://www.mzes.uni-mannheim.de/d7/de/publications/book/the-international-standard-classification-of-education-isced-97-an-evaluation-of-content-and-criterion-validity-for-15 (Last checked 2023-09-04) www.mzes.uni-mannheim.de/publications/misc/isced_97/luij08_the_educational_system_of_the_netherlands.pdf (Last checked 2023-09-04)

MERLOT (2023). https://merlot-education.eu/en/homepage/ (last checked 2023-09-04)

MOOChub format (2023). https://github.com/MOOChub/schema (last checked 2023-09-04)

NAF (2008). https://www.insee.fr/en/metadonnees/nafr2/section/A?champRecherche=true (last checked 2023-08-01)

OECD Background Report Portugal: Tertiary Education in Portugal Background Report prepared to support the international assessment of the Portuguese system of tertiary education. Ministério da Ciência, Tecnologia e Ensino Superior, Lisboa, 2006.

OECD Fields of Science and Technology: Revised Fields of Science and Technology (FOS) Classification in the Frascati Manual. OECD - Working Party of National Experts on Science and Technology Indicators. OECD Publishing, Paris, 2007. https://unstats.un.org/wiki/display/EC/Revised+Field+of+Science+and+Technology+%28FOS%29+classification+in+the+Frascati+Manual (last checked 2023-09-04)

ÖFOS (2012) https://www.data.gv.at/katalog/en/dataset/92750ae3-6460-3d51-92a7-b6a5dba70d3d (last checked 2023-09-04)

Parmar, A.: Review on Sharable Content Object Reference Model (SCORM). In: Second International Conference on Advanced Computing & Communication Technologies. IEEE Computer Society, 2012, pp. 409–411. https://doi.org/10.1109/ACCT.2012.95 (last checked 2023-09-04)

Poltrack, J.; Hruska, N.; Johnson, A.; Haag, J.: The Next Generation of SCORM: Innovation for the Global Force. In: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC). NTSA National Training and Simulation Association 2012, pp. 1–9.

Priem, J.; Piwowar, H.; Orr, R.: OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. In: 26th International Conference on Science, Technology and Innovation Indicators (STI 2022). European Network of Indicator Designers (ENID) 2022, Granada, Spain. https://arxiv.org/ftp/arxiv/papers/2205/2205.01833.pdf (last checked 2023-09-04)

Queirós, R.; Leal, J. P.: A Survey on eLearning Content Standardization. In: Lytras, M. D.; Ruan, D.; Tennyson, R. D.; De Pablos, P. O.; Peñalvo, F. J. G.; Rusu, L. (Eds.): Information Systems, E-learning, and Knowledge Management Research. Springer, Berlin, Heidelberg, 2013, pp. 433–438.

Ruipérez-Valiente, J. A.; Staubitz, T.; Jenner, M.; Halawa, S.; Zhang, J.; Despujol, I.; Maldonado-Mahauad, J.; Montoro, G.; Peffer, M.; Rohloff, T.; Lane, J.; Turro, C.; Li, X.; Pérez-Sanagustín, M.; Reich, J.: Large scale analytics of global and regional MOOC providers: Differences in learners' demographics, preferences, and perceptions. In: Computers & Education, 180, 2022, pp. 104426. https://doi.org/10.1016/j.compedu.2021.104426 (last checked 2023-09-04)

Schaart, R.; Moens, M. B.; Westerman, S.: The Dutch Standard Classification of Education (SOI). Statistics Netherlands, Den Haag, 2008.

Shah, D.; Pickard, L.; Ma R.: Massive List of MOOC Platforms Around The World in 2023. 2023. https://www.classcentral.com/report/mooc-platforms/ (last checked 2023-09-04)

Sinha, A.; Shen, Z.; Song, Y.; Ma, H.; Eide, D.; Hsu, B.; Kuansan, W.: An Overview of Microsoft Academic Service (MAS) and Applications. In: WWW '15 Companion: Proceedings of the 24th International Conference on World Wide Web. Association for Computing Machinery 2015. Florence, Italy, pp. 243–246. https://doi.org/10.1145/2740908.2742839 (last checked 2023-09-04)

Stratakis, M.; Christophides, V.; Keenoy, K.; Magkanaraki, A.: E-Learning Standards. 2003. https://www.dcs.bbk.ac.uk/selene/reports/Del21.pdf (last checked 2023-09-04)

Vardi, M. Y.: Will MOOCs Destroy Academia? Commun. ACM, 55, 11, 2012, pp. 5. https://doi.org/10.1145/2366316.2366317 (last checked 2023-09-04)