

Recommendations for a minimal metadata set to aid harmonised discovery of learning resources

<u>Executive Summary:</u>	2
<u>Background:</u>	4
<u>Overview of Work:</u>	5
<u>Scope</u>	5
<u>Vision</u>	5
<u>Audience</u>	6
<u>Approach</u>	6
<u>Metadata Schema Mapping and Comparison</u>	7
<u>User stories</u>	9
<u>Community Feedback</u>	11
<u>Results & Recommendations</u>	11
<u>Minimal Metadata Set</u>	11
<u>Table 1. RDA Minimal Metadata for Learning Resources</u>	12
<u>Data Dictionary</u>	13
<u>Way forward</u>	14
<u>Implementation Testbed(s)</u>	14
<u>FAIRsFAIR- EOSC Future</u>	14
<u>SSH (Social Sciences and Humanities) Training Discovery Toolkit</u>	15
<u>Further implementations</u>	16
<u>Metadata Application Profile</u>	16
<u>Call for adoption stories</u>	17
<u>Future Work</u>	17
<u>Extended Documentation Focus Group</u>	17
<u>Related Work: Testing Core Characteristics of Learning Resource Collectors FG</u>	18
<u>Participate by joining the ETHRD-IG</u>	18
<u>List of Abbreviations</u>	18
<u>Acknowledgements</u>	19

Executive Summary:

As part of the [Education And Training On Handling Of Research Data IG](#) activities, the Minimal Metadata for Learning Resources¹ Focus Group recommends a [minimal set of metadata for learning resources](#). [See below for a view of the minimal metadata set.] By comparing and analyzing existing learning resource-related metadata schemes to find the overlaps, the group provides guidance on metadata elements that should be minimally required for purposes of learning resource discovery to those concerned with supporting or providing training resources.

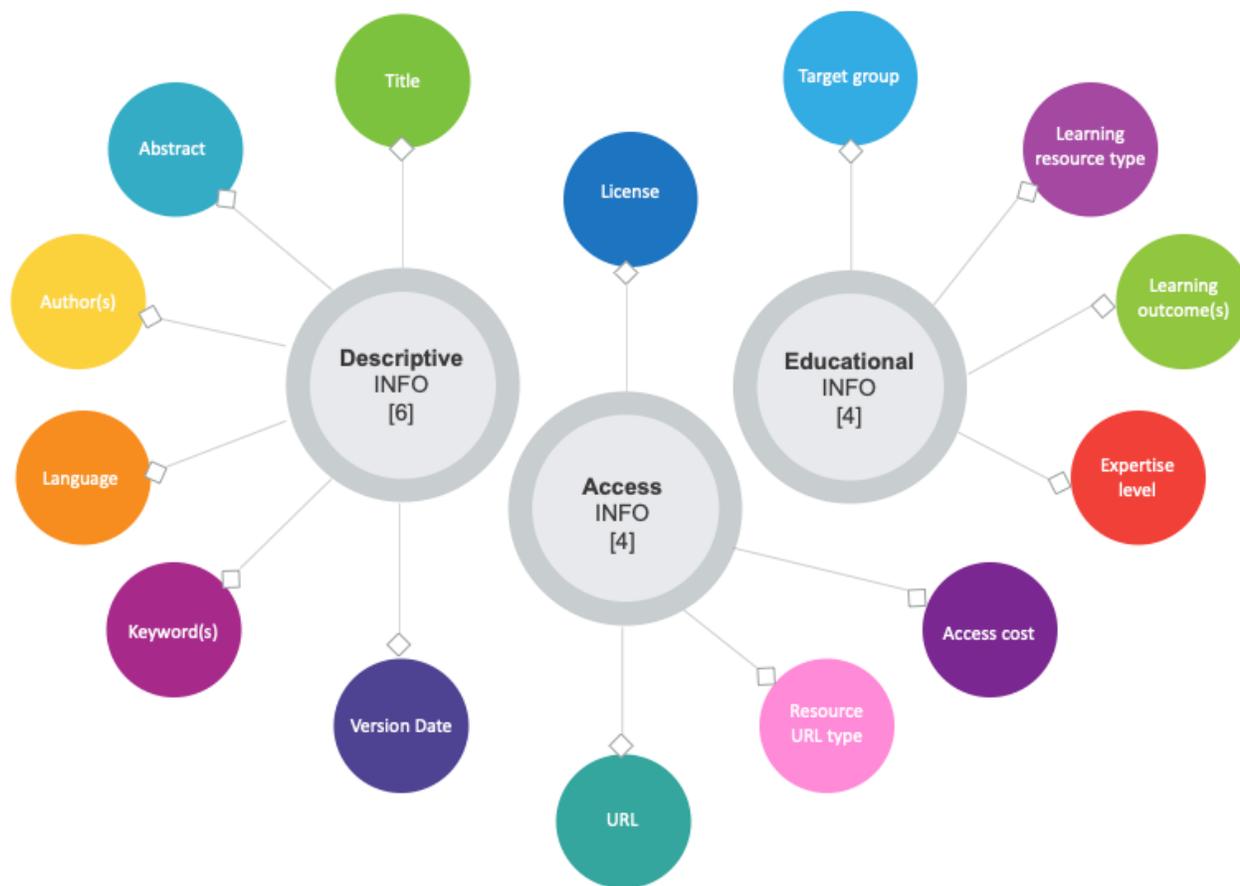


Figure 1: Minimum set of metadata for learning resources

¹ By learning resource we mean a persistent information resource that has one or more physical or digital representations, and that explicitly involves, specifies or entails a learning activity or learning experience. As an information resource it cannot be, for example, a person, object, and since it is persistent it cannot be an event (though it can be a record of an event). A learning activity or experience is one that has characteristics that may improve or measure a person's knowledge, skills or abilities. A learning resource may reference other supporting materials, creative works, tools etc that do not themselves meet this definition. <https://www.dublincore.org/groups/lrmi-task-group/>

The scope of this group's work was to focus upon discovery of learning resources, leaving discussions of reuse of those materials to a later group that planned to extend the minimal set of metadata and other forms of documentation.

The vision of the group was to provide assistance to learning resource searchers (learners of various types), resource creators (trainers, and resource aggregators (training service providers) by building upon previous work of the learning and training communities to facilitate resource discovery and reuse.

The results of the group's work to date include recommendations for a minimal set of metadata that can be used by learning and training stakeholders, a set of user stories used to both analyse and confirm the necessity of the metadata terms included in the minimal set, data dictionary that provides definitions and examples of how the minimal set can be used, an in-draft metadata application profile to facilitate adoption of the minimal metadata set, and an in-process testbed to further assess the utility and practicability of the minimal metadata set.

The results and recommendations of this group will be the foundation of further discussions by another ETHRD-IG focus group intending to extend the minimal metadata set as needed and add recommendations for further documentation for purposes of learning resource reuse.

Background:

The Research Data Alliance (RDA) Education and Research on the Handling of Research Data Interest Group (ETHRD-IG) was created with the focus upon “the exchange of information about existing developments and initiatives and promotion of training/education to manage research data throughout the data lifecycle.” In order to make this objective concrete, the ETHRD-IG identified three different areas in which the members work. These areas are to:

1. Enable the setting of quality standards for appropriate education and training programmes aimed at researchers and the professionals that support them, at all career stages;
2. Encourage the recognition of data skills amongst employees, employers, and professional bodies.
3. Prepare the ground for practical applications applying these standards in educational environments²

In early June 2020, interested ETHRD IG members began meeting to discuss an issue that had arisen as important to pursue in the previous RDA Plenary. That is, while many learning resources³ already existed that addressed many of the IG’s areas of concern (i.e., education and training on the handling of research data for researchers, commonly agreed upon standards for the data skills and competencies needed by both researchers and data support specialists), the group agreed that the discovery and re-use of those learning resources is more difficult and less widespread than desired. At the plenary breakout session, participants agreed that what made the most sense to do in order to reduce these barriers was for the ETHRD-IG to focus upon coming to an agreement on the range and depth of metadata that should be created for these learning resources along with recommendations on how the metadata should be created.

At the first meeting, the group decided that addressing this broad issue should be broken out into three different working focus groups: 1) identifying and recommending a set of minimal metadata elements, 2) extending the minimal set to include documentation (that included both metadata elements beyond the minimal set and other forms of documentation and guidance), and 3) addressing some of the issues faced by the organisations that support learning resource discovery and re-use including learning resource portals, catalogues, and registries.

² <https://www.rd-alliance.org/groups/education-and-training-handling-research-data.html>

³ While there was much discussion about what term to use for the educational and training materials to be considered as relevant for the focus groups’ work, the eventual decision was to use “learning resources” as that term. By learning resource we mean a persistent information resource that has one or more physical or digital representations, and that explicitly involves, specifies or entails a learning activity or learning experience. As an information resource it cannot be, for example, a person, object, and since it is persistent it cannot be an event (though it can be a record of an event). A learning activity or experience is one that has characteristics that may improve or measure a person's knowledge, skills or abilities. A learning resource may reference other supporting materials, creative works, tools etc that do not themselves meet this definition.
<https://www.dublincore.org/groups/lrmi-task-group/>

Early discussions of the group resulted in the merging of the Minimal MD set Focus Group (FG) and the Extended Documentation FG as the two topics are intrinsically related. The focus group working on issues related to the learning support organisations continued to work separately although the overall vision for the work of the metadata groups was that the implementation and adoption of the metadata / documentation focus groups' recommendations be taken up by the learning support organisations. Per these discussions, the work of the focus groups was divided as follows:

- the Minimal Metadata (MD) for Learning Resource group focus would be on identifying descriptive characteristics that should be applied to learning resources (LRs) for purposes of their **discovery** in training portals, catalogues & registries
- the Extended Documentation for Learning Resources group was to be focused upon metadata and other forms of documentation that would facilitate the **reuse** of those materials.
- the Learning Resource Aggregators group focus was to create a snapshot inventory of LR portals, catalogues & registries, and also identify a list/description of core characteristics of those portals, catalogues and services that would help their organisers, funders and maintaining organisations sustain them over time.

Overview of Work:

Scope

The scope of the Minimal Metadata group's work was to focus upon discovery of learning resources, leaving discussions of reuse of those materials to a later group that planned to extend the minimal set of metadata and other forms of documentation for purposes of facilitating reuse of learning resources. While seeming applicable, analysis and mapping of the full range of metadata terms considered to be important for implementing the FAIR principles (FINDABLE ACCESSIBLE INTEROPERABLE REUSABLE) for learning resources was not fully explored in this group's work, and so are discussed only briefly in the group's results and recommendations. Rather, more complete discussions on the applicability of both minimal, recommended and optional metadata terms and other documentation to the FAIR principles will be taken up by the ETHRD-IG's Extended Documentation Focus Group in the near future.

Vision

The vision for the minimal metadata Focus Group is to develop a core set that can be recommended as a means to:

- facilitate discovery across catalogues and other resources
- provide utility for different points of view by targeting
 - resource searchers e.g. learners,
 - resource creators e.g. trainers

- resource aggregators e.g. (training) service providers.
- improve interoperability between learning resource catalogues
- work towards establishing best practice in applying FAIR principles to learning resources
- provide a foundation for the work of the products from the other Focus Groups.

By establishing a core set we can help improve discoverability and, in the longer term, reduce duplication of effort by service providers and help identify gaps among existing and prospective learning resource creators and aggregators.

Audience

The set of recommended minimal metadata elements is intended to support key stakeholders including resource creators, individual trainers, and learning resource service providers who create or adapt learning resources, and also services that support their discovery. The group anticipates that the output of its work can thus be used by general and domain repositories which store and share learning resources, but also learning catalogues or registries, educational platforms and search portals that curate learning resources in order to improve the capability of both learners and trainers to find resources more effectively.

Approach

Part of the vision of the Focus Group was to draw on previous work that has helped facilitate the discovery of learning resources. In the first instance as a result, a desk exercise of metadata mapping and comparison across different metadata schemas was conducted. Alongside the metadata mapping, a set of user stories were developed against which the metadata were assessed. The desk-based activities were followed by a period of community consultation to gather feedback on both the minimal set and the relevance of the user stories.

Early in the process the group looked at what was missing and had specific discussions related to how the metadata elements important for discovery of learning resources might apply to the FAIR principles. It became clear that metadata which supported the FAIR principles provided an immediate benefit to training aggregators such as training portals, catalogues and registries. Indeed, the FAIR principles aim to improve findability for both humans and machines. Whenever structured metadata is the central topic, those not familiar with it find it difficult to realize what the immediate benefit for humans is. As more and more (meaningful) metadata are added to training materials (or any other research object, e.g., data), finding the metadata becomes easier for machines so that, for instance, search engines -- either generic or specialized ones, can present better results to humans. Most searchers are probably used to those quick summaries provided by general-purpose search engines, e.g., Google and Yahoo, whenever we search for a movie or recipe. These search engines benefit from metadata to present more accurate results to their users. In the same way, specialized search engines targeting researchers can also use metadata to improve their offerings. For these reasons, discussions about the FAIR principles and their relation to metadata for learning resources mostly focused

upon the findability advantages associated with a *minimal* metadata set. A more complete set of metadata and documentation will undoubtedly describe more fully the advantages of such structured documentation upon the accessibility, interoperability and reusability of learning resources.

The process used by the group provided the means for the collection of a “master list” of 84 metadata elements for learning resources to be whittled down to the current 14 elements.



Figure 2: Metadata analysis & recommendation process

The approach to the metadata mapping, user stories’ analysis, and community feedback are described more fully below.

Metadata Schema Mapping and Comparison

A number of metadata schemas for learning resources (LOM⁴, LRMI⁵, bioschemas⁶, etc.) already exist.⁷ These schemas can be fairly extensive or domain specific, and are widely implemented in learning resource collections and Open Education Resource (OER) repositories. They vary in the extent of the description of resources, and in technical implementation. Also, depending on the user view (training infrastructure, learner, trainer), they may not always be understandable or necessary to, or relevant for a searcher. By comparing several frequently used metadata schemas to identify common metadata elements as well as what was missing from the existing schemas, this focus group aimed to identify and recommend a set of metadata elements minimal enough to encourage their consistent use by resource creators or adaptors to describe the learning resources when making them publicly available. The additional benefit of consistent use of this minimal set would be to allow greater success in the learning resource’s discovery and ultimately, re-use.

⁴ We are using [MS22: Training materials for the ENVRI data centers are produced and available at the training portal](#) which is a recommended subset of LOM appropriate for training resources.

⁵ See https://www.dublincore.org/specifications/lrmi/lrmi_1/

⁶ See [Bioschemas/Training](#)

⁷ See [List of Abbreviations](#) section below.

In addition, the use of an agreed upon minimal metadata set should improve the chances that learning resources will be made more FAIR and, therefore, better meet the cross-collection and cross-community needs of learning resource users. The existence of a more commonly agreed upon set of metadata will help to meet a common need for aggregators of learning resource metadata to be able to interoperate with machine-actionable metadata from multiple collections, each of which may be using a different OER standard or schema to describe and expose their metadata.

The metadata mapping process often represents the most suitable method for guaranteeing interoperability across different systems/catalogues where the structural and semantic heterogeneity does not allow achievement of a common agreement with a given set of metadata fields. Further, a mapping can be used to develop a metadata application profile to help learning resource service providers make resources more widely discoverable and reused. Hence, in order to develop a metadata application profile (as is in process as part of this focus group's work), a crosswalk exercise has been performed.

In particular:

1. A set of already existing standardized schemes/profiles for learning resources was identified:
 - Learning Resource Metadata Initiative (LRMI);
 - EOSCpilot⁸
 - IEEE Learning Object Metadata (ENVRI-LOM used by the ENVRI-FAIR project and the LifeWatch ERIC Research Infrastructure);
 - Bioschemas Data Catalog profile;
 - Bioschemas Training profile;
 - Schema.org;
2. From the set of standardized schemas/profiles, a set of common characteristics describing the learning resources was identified and mapped across the different schemas. (For each element, a title and a description have been provided)
 - See snapshot of mapping to LRMI, EOSCpilot, and IEEE LOM at: https://drive.google.com/drive/folders/1gJSCaUAyL9c4n_skFksZrHQ91OyJ9pEd
 - See snapshot of spreadsheet with mapping to schema.org at: <https://drive.google.com/drive/folders/1A9o3lxjKhHRU0R5onyy8SIYDQc1WuR9f>
3. This initial mapping of the elements from the different schemas generated a list of 84 metadata element candidates. For this step, the FAIR guiding principles were analysed

⁸ Whyte, Angus, Leenarts, Ellen, De Vries, Jerry, Huigen, Frans, Kuehn, Eileen, Sipos, Gergely, Kalaitzi, Vasso, Dijk, Elly, Jones, Sarah, & Ashley, Kevin. (2019). D7.5: Strategy for Sustainable Development of Skills and Capabilities (1.1), p. 43: Zenodo. <https://doi.org/10.5281/zenodo.5095052>

in order to check if all needed and required metadata fields had been considered and included in the list of 84 metadata element candidates⁹.

4. The “master list” of metadata elements was then divided into “minimal required metadata” (mandatory fields that are necessary for the understanding/findability of the training materials) and “useful for extended documentation metadata” (recommended fields which inclusion is necessary to ensure the FAIRness of the training material).

In order to bring a wider point of view to bear from different perspectives (and stakeholders) on the analysis of which descriptors would be most relevant to use for finding appropriate learning resources, a “user story” approach was adopted.

User stories

Following the initial mapping across schemas in order to develop the minimal set of metadata elements, a series of user stories were created to represent different user types and wants.

The development of user stories enabled the Focus Group to assess the metadata through a range of different perspectives.

An initial list of user stories were created and then these were grouped together resulting in five user types: learners, trainers, service providers, developers and funders¹⁰. The 15 user stories are described below.

Learner¹¹ User Stories: A learner could be someone in the workplace doing self directed continuing professional development, or a student doing self directed learning, i.e. in all cases looking for their own resources, not using resources provided by or suggested by a teacher or instructor.

- (L1) Learner searching for training for specific skills;
- (L2) Learner looking for free courses on a specific topic;
- (L3) Learner searching for a specific type of learning resource, i.e. MOOC, self-study, webinar, Face to Face training opportunities;
- (L4) Learner wanting to find materials from a course

Trainer¹² User Stories: A trainer could be an individual or group of trainers. They may be located in an educational or workplace environment. A trainer can be anyone providing training even if it is not their formal role. We have assumed that each of these trainer “users” has a shorter term goal of discovery with a longer term goal of reusing the resources they discover.

- (TR1) Trainer looking for training resources for their own training;

⁹ Note: Previous work published by EOSCpilot was also important for its discussion of FAIR training materials. See: Kühn, Eileen, & Streit, Achim. (2017). D7.2: Interim report and catalogue of EOSC skills training and educational materials (1.0), p. 24: Zenodo. <https://doi.org/10.5281/zenodo.3395959> .

¹⁰ There is no specific user story for a researcher as it was felt the role of the researcher was absorbed into the user story either as a learner or potentially as a trainer.

¹¹ Learner - synonym: trainee, student

¹² Trainer - synonym: instructor

- (TR2) Trainer looking for existing modules to combine for a new course;
- (TR3) Trainers looking for resources in a given language);
- (TR4) A trainer wanting to share their materials

Service Provider User Stories: In this context a service provider may be a training provider who produces training material or a third party who aggregates and makes training materials available e.g a catalogue.

- (SP-1) Service provider wanting to promote their training resources;
- (SP-2) Service provider wanting to know where gaps are in subject coverage;
- (SP-3) Service provider wanting to provide an overview of training possibilities for their community

Developer User Stories: A developer could be working independently or be working with a service provider.

- (D1) Developer wanting to add to or create web interfaces;
- (D2) Developer wanting to import metadata into a different catalogue;
- (D3) Developer looking to reuse data about OER objects for different applications

Funder User Story: In the broadest sense a Funder is any organisation providing funding e.g a charity, a national research funder, an institution etc.

- (F1) Funder wanting to know which training resources have been created by a specific project

Working with the master list of metadata, each user story (e.g L1, T1 etc) reviewed all metadata elements for relevance to that use case with the emphasis on the F in FAIR of findability. Assessment was made as to whether a metadata element was required (yes/no), recommended or optional. When all the user stories had been completed, the group reviewed and harmonised the responses across the user stories to establish the minimal metadata set. Only elements that were considered to be necessary were added to the minimal set; elements that were optional or recommended will be considered in the context of an extended metadata set. This approach was endorsed by the consultation process and we were able to establish a set of 14 elements from the original 84 as illustrated in Figure 2 above.

The analyses by user stakeholders can be found at:

- Stakeholder User Story list: [Report_UserStory_StakeholdersList.pdf](#)
- Learner User Story Analysis: [Report_UserStory_Learner.pdf](#)
- Trainer User Story Analysis: [Report_UserStory_Trainer.pdf](#)
- Service Provider User Story Analysis: [Report_UserStory_ServiceProvider.pdf](#)
- Developer User Story Analysis: [Report_UserStory_Developer.pdf](#)
- Funder User Story Analysis: [Report_UserStory_Funder.pdf](#)

It should be noted that the minimal set was developed by consensus and there was not a unanimous agreement, as even 14 metadata elements are considered to be “too many” for a

minimal set by some stakeholders. Whilst the minimal metadata set is subject agnostic, a guiding principle for the set was that it would include specific elements of relevance for training materials and learning resources and was not just a generic schema that could be applied to any material or resource type.

Community Feedback

There has been close working within the ETHRD-IG of the different Focus Groups and representation from overlapping communities e.g Bioschemas.

Throughout the process of developing the minimal metadata set there has been engagement and consultation with the learning resources/training community. ETHRD-IG members have participated in workshops run by FAIRsFAIR and INFRAOESC 5b Training and Skills TaskForce¹³, Community of Practice for training coordinators¹⁴ as well as participation in the RDA Plenaries RDA P16 and RDA P17. During RDA P17, there were two workshops¹⁵ where the minimal set was presented and consulted on. Feedback from these events were incorporated into finalising the proposed minimal metadata set. Other feedback and analysis took the form of comparisons among recommendations made by other communities and experts. See the results of this comparison at: [Report RDAMinimalSetComparison P17](#).

Along with the consultations described above, members of the Focus Group have consulted within their own communities and networks. A number of initiatives are actively incorporating the minimal metadata set in their work as a result of these consultation activities. See the list of contributors in the [Acknowledgements](#) section below.

Results & Recommendations

Minimal Metadata Set

As previously mentioned, the set of recommended minimal metadata elements is intended to support key stakeholders who create or adapt learning resources, but also services that support their discovery. The minimal set can be used to improve the capability of both learners and trainers to find resources more effectively by general and domain repositories which store learning resources, but also learning catalogues or registries, educational platforms and search portals that curate aggregated learning resources .

¹³ Workshop on Harmonising Training Resource Metadata for EOSC Communities April 2021 (<https://doi.org/10.5281/zenodo.4769468>)

¹⁴ Workshop on the RDM training and support landscape September 2021, <https://www.opensciencefair.eu/2021/workshops/the-rdm-training-support-catalogue-landscape>

¹⁵ [Reducing barriers to and increasing global participation in data interchange through advancing discovery, access, and reusability of RDM training resources | RDA \(rd-alliance.org\)](#)

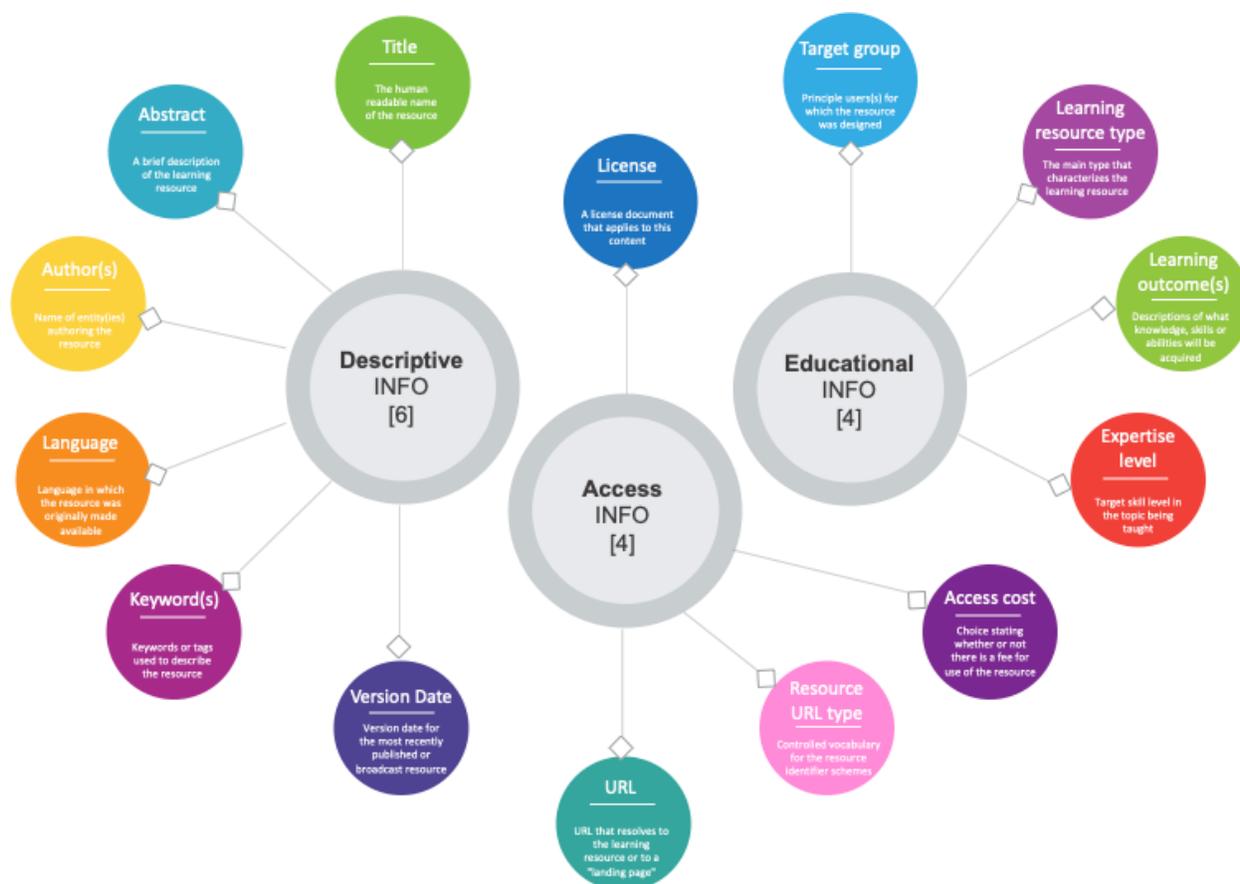


Table 1. RDA Minimal Metadata for Learning Resources

Element Name	Definition
Title	The human readable name of the resource.
Abstract / Description	A brief synopsis about or description of the learning resource.
Author(s)	Name of entity(ies) authoring the resource.
Primary Language	Language in which the resource was originally published or made available.
Keyword(s)	Keywords or tags used to describe the resource.
License	A license document that applies to this content, typically indicated by URL
Version Date	Version date for the most recently published or broadcast resource.

URL to Resource	URL that resolves to the learning resource or to a "landing page" for the resource that contains important contextual information including the direct resolvable link to the resource, if applicable.
Resource URL Type	Designation of the identifier scheme used for the resource URL, e.g., DOI, ARK, Handle.
Target Group (Audience)	Principal users(s) for which the resource was designed.
Learning Resource Type	The predominant type or kind that characterizes the learning resource.
Learning Outcome	Descriptions of what knowledge, skills or abilities a learner should acquire on completion of the resource.
Access Cost	Access cost: Choice stating whether or not there is a fee for use of the resource (CV = Y/N/Maybe with recommendation that further explanation of "Maybe" goes in the Description field for "It depends" or "It changes" explanations).
Expertise (Skill) Level	Target skill level in the topic being taught; example values include: beginner, intermediate, advanced.

Data Dictionary

A more complete explanation of the elements chosen to include in the minimal metadata set can be found in a data dictionary that includes the terms, definitions, type of value expected (i.e., text, date, etc.), usage notes, allowed values, examples, other constraints and recommendations for controlled vocabularies to use. See the full data dictionary at: [Report Snapshot DataDictionary.pdf](#)

In addition, there are examples of values to add for learning resources that are described for searchers with different perspectives and educational needs, i.e., formal education (e.g., academic or university training), professional development (e.g., to satisfy professional continuing education requirements), and informal education (e.g., learning a skill in Python, for example). See a snapshot of these examples for:

Formal Education at: [Examples for Formal Education](#)

Professional & Information Education at: [Examples for Professional and Informal Education](#)

Both the Data Dictionary and the Examples include links to controlled vocabularies for terms that warrant them. Note that the names of the controlled vocabularies on these documents have come from members of the ETHRD-IG and other education and training communities with experience in a variety of subject and skill domains related to research data generation and management. They are not necessarily relevant to all communities, however, and should be considered carefully by learning resource creators, trainers, and service providers to make sure

that they are appropriate for their target audiences, and for the infrastructure and support of the system using them. The U.K.'s JISC organisation has identified a number of considerations to use when choosing controlled vocabularies for your communities:

- Your users - are the terms used going to be meaningful to them?
- The nature and extent of your collection - if your collection is small, you're unlikely to need a highly detailed vocabulary
- The skills and available time of your cataloguing staff - some of these vocabularies will require experience or training to use properly
- Your community - it makes good sense to use vocabularies that similar collections are using
- Copyright issues - you may need to check whether permission or a license is required to use the vocabulary in the way you wish to

More information about controlled vocabularies can be found at this JISC Guide: [Controlled vocabulary](#). In addition, a recently published article in PLOS discusses how to make vocabularies FAIR at: [Ten simple rules for making a vocabulary FAIR \(https://doi.org/10.1371/journal.pcbi.1009041\)](#). Topics related to the appropriate use and extension of controlled vocabularies have been identified as important to further discussions on metadata and other forms of documentation by the Extended Documentation Focus Group.

Way forward

While a great deal of progress has been made in the time the focus group has been meeting, there is still more action needed and planned to test the minimal metadata set, to support and encourage its adoption, but also to move beyond a minimal metadata set to more complete documentation of learning resources for purposes of FAIR data implementation.

Current efforts to support the use and improvement of the minimal set include implementation testbeds, the development and testing of a metadata application profile using the Dublin Core - Tabular Application Profile template¹⁶, and a call for other adoption stories of the set.

Implementation Testbed(s)

FAIRsFAIR- EOSC Future

In addition to recommending a minimal metadata set, the focus group is interested in supporting efforts to assess the use and implications of the minimal metadata terms, definitions and type recommendations for different searcher target audiences, different users of the metadata

¹⁶ https://www.dublincore.org/blog/2020/dc_tabular_application_profiles/

including resource creators, trainers, and service providers, and different types of learning resource aggregators. As part of that interest, an effort is currently underway in the FAIRsFAIR project to create an implementation testbed in collaboration with various projects, mostly European, to inform the harmonisation of catalogues working in the EOSC context, and in particular the work beginning in the EOSC Future project to create a 'catalogue of catalogues'. Some of the issues to be tested include:

- Whether the minimal set can be used to exchange metadata between catalogues, to demonstrate the potential for improved discoverability across different catalogue needs and perspectives/roles; e.g.,
 - Does the minimal set offer an appropriate level of complexity?
 - Does the minimal set need more precision or flexibility to accommodate differences in constraints applied to the elements by the contributing catalogue providers?
 - Should certain controlled vocabularies be more strongly recommended?
- How may an application profile based on the minimal set facilitate interoperability across different platforms, e.g. through common validation methods and/or, harvesting/aggregation; what further guidance is needed on this to implement the set, in order to better facilitate accessibility, interoperability, and reusability of the set, and the resources themselves;
- Whether further explanations, guidance and/or recommendations are warranted for different users (e.g., on common or widely recognized practices for metadata input such as order entry for author names, inclusion of personal identifiers), especially trainers and service providers

The implementation testbed as currently planned will involve the following catalogue providers:

- [SSHOC](#)
- [ENVRI-FAIR](#)
- [ELIXIR/ TeSS](#)
- [EOSC Pillar](#)
- [NI4OS](#)
- [ESOC Synergy](#)
- [PaNoSC/ExPaNDS](#)
- [Data Management Training Clearinghouse](#)
- [EOSC Future project](#)

SSH (Social Sciences and Humanities) Training Discovery Toolkit

The SSHOC (Social Sciences and Humanities Open Cloud) project¹⁷ has developed the SSH Training Discovery Toolkit¹⁸ to improve the discoverability of good quality training resources in the SSH and beyond. The toolkit is meant as a tool for trainers and part of the train-the-trainer activities of the project. The structure of the metadata fields and the data model initially used in the Toolkit were evaluated in April 2021 using a draft of the RDA minimum metadata consisting of 10 minimum metadata. Based on this initial comparison a couple of issues were identified:

¹⁷ <https://sshopencloud.eu/training>

¹⁸ <https://training-toolkit.sshopencloud.eu>

- not all RDA minimum metadata were included
- not all RDA minimum metadata were included as mandatory fields
- information on some metadata like author could be found in multiple fields like author and responsible organisation

In the meantime, the draft of the RDA minimum metadata has evolved into a more elaborate list of metadata. In the upcoming months, the SSH training discovery toolkit will be updated to improve the metadata by adapting the data model to more closely align with the RDA minimum metadata, as well as implementing schema.org and including controlled vocabularies. The experiences of the team will be reported in the SSHOC project deliverable on the toolkit in March 2022.

Further implementations

1. The EOSC Future project customized the RDA minimal set of metadata for learning resources for the EOSC Training Catalogue and made it publicly available in the following wiki page at <https://wiki.eoscfuture.eu/x/zwOK>, to be updated according to the needs that could be identified during the development of the Training Catalogue.
2. The RDA minimal set of metadata is also used in the CoP FAIR Training Materials Focus Group.
3. Australian Research Data Commons checklist <https://doi.org/10.5281/zenodo.5003933> will be updated and they are looking at the minimal metadata set.
4. [OpenAIRE](#) learning platform - still under development, but planning to adopt the RDA Minimal MD set.
5. [CENTRUM VEDECKO-TECHNICKÝCH INFORMÁCIÍ SR](#) - a national training platform in Slovakia - still under development, but planning to adopt the RDA Minimal MD set.
6. [Catalogue of training materials on RDM in German-speaking countries](#) - still under development, but planning to adopt the RDA Minimal MD set.

Metadata Application Profile

We're hoping that the minimal metadata profile can help our target audiences create, validate, exchange, and select metadata, and also map between different sources of metadata. To further those efforts, we're hoping that the creation and use of a metadata application profile will further enable learning resource providers and aggregators to make more resources more widely discoverable, and reused. In collaboration with the Dublin Core Metadata Initiative's DC-TAP working group, we are in the process of developing a metadata application profile using its Dublin Core - Tabular Application Profile template¹⁹. The purpose of developing such a profile is to facilitate further implementations by providing, in machine-actionable form, rules that govern the creation and reuse of metadata instances. Its function would be to both explain the metadata but also to potentially constrain the metadata so that correct usage can be determined.

¹⁹ <https://github.com/dcmi/dctap/blob/main/TAPtemplate.csv>

Call for adoption stories

We encourage other learning resource communities to test the minimal set using the Data Dictionary and Example documents and let us know your feedback. If you do adopt the minimal set for use in describing the learning resources that you create, adapt or use, or within your service provider infrastructure, please consider sharing your adoption story with us.

Future Work

Once the minimal metadata set and associated guidance documents are finalized, and the implementation testbeds and the application profile, the work of the Minimal Metadata Focus group will be finished. ETHRD-IG members understand that the problem of learning resource reusability will not have been solved. As a result, future activities include taking discussions and efforts forward into an Extended Documentation Focus Group as well as testing the products of another related ETHRD-IG focus group on identifying core characteristics of learning resource aggregators and other service providers to help support and facilitate their long-term sustainability.

Extended Documentation Focus Group

The Extended Documentation Enabling Reuse of Learning Resources Focus Group will build on the work of the Minimal Metadata Focus Group to recommend extended documentation for learning resources which may include more, recommended and optional metadata to enable effective reuse of learning resources, to describe the context of learning resource delivery and to signal learning resource effectiveness. The group's focus will be on identifying and developing guidelines for documenting the information you need to have in order to effectively and confidently reuse learning resources. The intention for the further recommendations and guidance is to improve the ability of learners and instructors to reuse resources more easily and effectively, and to establish a jumping off point for indicating quality of materials within registries.

Examples of other metadata elements may include learning objectives, linking to educational frameworks such as skills and competency requirements, assessment of learning resource effectiveness, and contextual details for how the learning resources were used or presented. Outputs of this group may include checklist(s) for documentation that should accompany a learning resource, guidance on creating recommended documentation or controlled vocabularies, and recommendations for metadata scheme extensions that enable better reuse decisions. Efforts may include further exploring linked open data and machine-actionability and providing recommendations about these topics (aligned with FAIR implementation). The current version of the Scoping Document for Extended Documentation FG can be found:

<https://docs.google.com/document/d/1VTn6QJ2YmUIb235gVvpY2sApfzsgGUhx6RYo7oi7A0/edit?usp=sharing>

Related Work: Testing Core Characteristics of Learning Resource Collectors FG

The need for sustainability of learning resource service providers such as catalogues, portals and registries was a related, but important item of discussion at the original RDA Plenary where the issues related to the need for better discovery and reuse of learning resources arose. ETHRD-IG members recognized that it may sometimes be difficult to find such service providers when the funding for them, the ways that they are built, and the information they provide about themselves and their services are either not sustained, or not evident. As a result, another focus group was created that was intended to address issues related to the longer term survival and success of service providers. The Scoping Document for the Learning Resource “Collectors” is available at: [Outline/scoping document](#). Part of that group’s work is a document that identifies core characteristics of these kinds of organisations that the group suggests will help facilitate their sustainability over time. The document can be found at: [Core Characteristics of Learning Resource Collectors](#). These characteristics have also been vetted by key service providers and the learning resource communities, but are as yet untested. The document is is openly available as an output of ETHRD-IG. The IG would greatly appreciate knowing if the document has proven helpful to service providers and/or been used in conjunction with other sustainability analyses or efforts.

Participate by joining the ETHRD-IG

We encourage and welcome members of education and training communities to join our discussions and contribute to our outputs. Join the RDA Interest Group on Education and Training on Handling Research Data here: <https://www.rd-alliance.org/groups/education-and-training-handling-research-data.html>. You can post here to express interest in working with any of the focus groups mentioned, provide feedback, or relate adoption stories for any of the outputs discussed above. Alternatively, please contact one of the chairs listed on the ETHRD-IG page indicated above, and we would be happy to help you out with any questions regarding the operation of ETHRD-IG.

List of Abbreviations

EOSC - European Open Science Cloud
ETHRD-IG - Education and Handling of Research Data Interest Group
FAIR - Findable, Accessible, Interoperable, Reusable
FG - Focus Group
IG - Interest Group
LR - Learning Resource
LOM - Learning Object Metadata
LRMI - Learning Resource Metadata Initiative

Acknowledgements

This work has been funded through in-kind contribution of effort from the members of the RDA ETHRD-IG Minimal Metadata for Learning Resources Focus Group, including the following projects:

- FAIRsFAIR “Fostering FAIR Data Practices In Europe”, which has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558
- Data Management Training Clearinghouse which has received funding from Earth Science Information Partners, the U.S. Geological Survey’s Community for Data Integration, and the Institute of Museum & Library Services, National Leadership Grant # LG-70-18-0092-18

Focus Group members:

Nancy Hoebelheinrich - Knowledge Motifs LLC, Chair
Katarzyna Biernacka - Humboldt-Universität zu Berlin, Co-chair
Michelle Brazas - OICR (ISCB, GOBLET, Bioschemas Training, Bioinformatics.ca)
Leyla Jael Castro - ZB MED Information Centre for Life Sciences (Bioschemas Training)
Nicola Fiore - LifeWatch
Maggie Hellström - Lund university (Sweden) and ICOS research infrastructure
Iryna Kuchma - EIFL/OpenAIRE
Emma Lazzeri - CNR Italy
Ellen Leenarts - Data Archiving and Networked Services (DANS)
Paula Martinez Lavanchy - TU Delft
Elizabeth Newbold - Science and Technology Facilities Council (STFC)/ FAIRsFAIR project
Amy Nurnburger - Massachusetts Institute of Technology
Esther Plomp - TU Delft
Lucia Vaira - LifeWatch ERIC
Celia van Gelder - DTL Projects/ELIXIR-NL/Health-RI, ELIXIR Europe, GOBLET
Angus Whyte - Digital Curation Centre/FAIRsFAIR project

Other consultants:

Professor Marcela Alfaro-Córdoba - University of California, Santa Cruz