Co-funded by the European Union

ENCORE

Guidelines for assessing the Quality of Repositories of OER (ROER)

By UNIPI and Beam Me Up

Università di Pisa
Universidad de Salamanca
Chalmers University of Technology
Università degli Studi di Padova
Hellenic Open University
Fondazione Bruno Kessler
mylia
ValueDo
Knowledge Foundation @Reutlingen University
EADTU
Adecco Training
BeamMeUp
Reconnaître

Guidelines for assessing the Quality of Repositories of OER (ROER)

Project result number D3.2 by UNIPI and Beam Me Up

28 October 2022

Co-funded by the European Union

Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Executive Summary

Within WP3, Task 3.2 aims to realize a reference framework to measure the quality of the Open Educational Resources (OERs). Beam Me Up (BMU), task leader, has designed and realized the framework, using a downstream approach: this means collecting as many OERs as possible from the available sources, skimming them following the method described in the current document and finally obtaining an overall evaluation of the ROER obtained. Thus, the quality of a ROER is defined by the quality of OERs contained. Starting from literature review, BMU has determined two possible evaluation methods of OERs: a direct one, which determines the OER quality by analysing its content using qualitative guidelines, and an indirect one, which determines the quality of the OER analysing the quality of connected metadata, using a quantitative automatic method.

The following document contains all the steps followed to realize the automatic method. The method for the indirect evaluation could be used to automatically evaluate a large quantity of OERs; the manual method, for the direct evaluation, could be used to precisely evaluate a sample or a specific subset of OERs that requires much attention.

Table of contents

1. Introduction

1.1 Scope

The goal of task T3.2 is to assess the quality of Repositories of Open Educational Resources (ROER). For this purpose, BMU developed an algorithm able to quantitatively evaluate a ROER, using multiple measurement methods.

The quality of the ROER is mainly determined by the quality of each single Open Educational Resources (OER) collected in the repository. According to this idea, it is reasonable to consider a high quality REOR as "*the one rich of high quality OERs*". Therefore, to assess the quality of the ROERs, it is necessary to determine the quality of each single OER.

We identified two main approaches for evaluating an OER:

- **Direct Evaluation:** determining the OER quality by evaluating the content itself.
- **Indirect (proxy) Evaluation:** determining the OER quality by evaluating its metadata.

Using a **Direct Evaluation**, it is possible to identify a deterministic causal relationship between the result of the quality evaluation and the real quality of the OER; using an **Indirect Evaluation** it is only possible to assume a probabilistic relationship.

- Direct evaluation:

Deterministic link between quality evaluation results and OER's real quality.

If the result of the evaluation is high

then OER's quality should be high.

- Indirect (proxy) evaluation:

Probabilistic (P) link between quality evaluation result and OER's real quality.

Using a Bayesian formulation: ***P(OER's quality is high | result of the evaluation is high) > P(OER's quality is high)****.*

We decided to develop an algorithm able to perform an **Indirect (proxy) Evaluation**. The reasons of this choice are the following:

- An automatic evaluation of OER's content quality (**Direct Evaluation**), considering different kinds of file format (text, video…), leads to big coding challenges and it is considered out-of-scope for this project.
- At the meantime, the development of an automatic method was the only possibility to guarantee the evaluation of large quantities of E.R. in a systematic way.
- It's reasonable to consider the quality of the metadata itself a relevant and appropriate factor of a ROER's quality.

To conclude the explanation of the approach adopted in Task 3.2, we should remind that the aim of the WP is to collect good quality OER related to GDE skills, to develop a suitable repository for the ENCORE project. To reach this goal, it is possible to choose among different strategies. We can mainly distinguish between two: it is possible to evaluate the ROER quality before collecting the OERs, selecting the sources to collect only good quality OERs in the repository (**Upstream Assessment**), or it is possible to evaluate the ROER's quality only after collecting them (**Downstream Assessment**). What just described is shown in Figure 1.

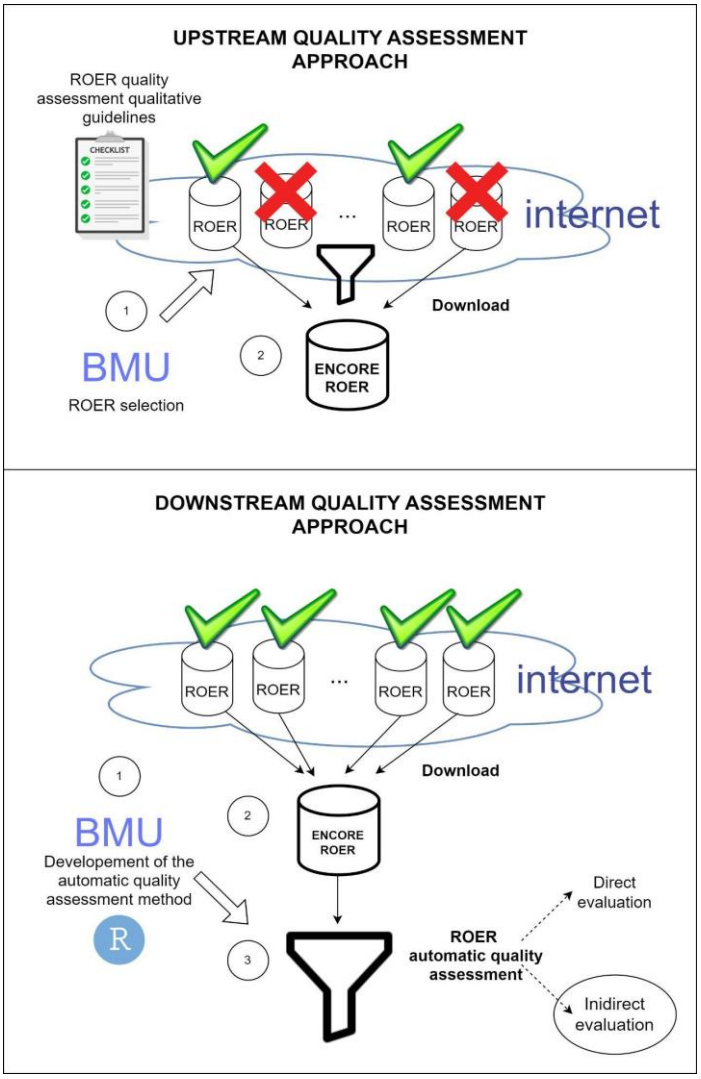


Figure 1 – The two possible strategies, the Upstream Quality Assessment, and the Downstream Quality Assessment (the one chosen)

We chose to carry out a **Downstream** ROER quality assessment, which means collecting as many OERs as possible from the available sources, and only after skimming them. This choice is due to the possibility given by this approach to maximise the advantage of analysing a large amount of data using an automatic method.

1.2 Debate around Educational Resources definitions

Before any further steps, it is important to describe what OERs are. IN the project we define them as: "*Educational materials of any kind (e.g., textbooks, worksheets, lesson plans, instructional videos, entire online courses, educational games) which can be freely used, adapted, and shared.*" This

definition is wide, so considering the goals of this task, we decided to deepen it, transforming it into an operational tool to determine what should be inside or outside of the ENCORE project scope.

Before defining what an OER (Open Educational Resource) is, we must define what an Educational Resource is. During time, many authors have proposed different definitions in the field of education and learning and different terms have been used to describe similar ideas such as Educational Materials, Educational Objects, Learning Objects (McGreal, 2004). Below, we summarize some of these definitions to come up with a common terminology for the Encore project.

Definitions

IEEE LTSC (2002) defines Learning Object (LO) as "*any entity, digital or non-digital which can be used, reused or referenced during technology supported learning*". This definition has an high level of granularity with respect to the goals of the present task.

Downes (2003) uses the example of tissue paper to argue that "*anything and everything can be used for learning and therefore must be considered to be a LO*". He added, also, that there is no reason to restrict a priori what counts as a LO.

Wiley (2000) settles on a definition of a Educational Resource as "*any digital resource that can be reused to support learning*" and Doorten, Giesbers, Janssen et. al. define LOs as "*any reusable resource, digital or non-digital that can be used to support learning activities.*" They also mention examples such as web pages, applications, textbooks, calculators, and microscopes.

Sosteric and Hesemeier (2004) emphasise the intent of the object more than the structure. According to them, an E.R. is "*a digital file (image, movie, etc.) intended to be used for pedagogical purposes, which includes, either internally or via association, suggestions on the appropriate context within which to utilize the object*." They claim that a newspaper article would not be an E.R. simply because it could be used for learning, it must be linked to "*pedagogical purposes*".

According to these authors, an information object becomes a E.R. when it is designed to be used by itself or in combination with other media objects to facilitate or promote learning.

1.3 Definitions for the ENCORE project

Keeping in mind what has been said above, we have three goals in developing our own definition of E.R.:

- Focus on e learning materials and contents that can be used through digital devices.
- Broad approach, not eliminating in advance any kind of formats (text, video, audio … etc).
- Focus on intrinsic aspects and not on pricing, legal or other side aspects.

Higgs et al (2003) provide a definition of Learning Object that contains many essential elements useful to our purposes. They define a set of six characteristics that a Learning Object must have to be considered as such: Instructional value, Independency, Shareability/Reusability, Interoperability, Discoverability and Context.

We consider this definition as the best for the practical needs of the ENCORE project, and therefore we decided to adapt it to adopt it.
First, a weakness of this definition is the focus on the dimensions of the E.R. given in the Shareable/Reusable characteristic (they say that "Learning Objects are small stand alone, reusable components that can be assembled..."). We think that it is impossible to define in advance a lower and upper bound on Educational Resource dimensions.
First, because it is very difficult to define which metric has to be used, i.e. time, number of educational goals, amount of information, etc. Secondly, because it's not essential to our purposes to give this kind of bound.

We think that is more useful not to give precise limits to the size of an E.R., but rather its size could be any designed by the creator in the act of developing a single, not immediately divisible, and stand-alone object.

Furthermore, we decide to introduce another element to the definition: the Educational Purpose. In fact, it is necessary to better explain and formalize this characteristic already mentioned in the Higgs' definition in the Instructional Value point, to indicate that the Instructional Value is the result of a precise and deliberate choice of the creator of the Educational Resource. Otherwise, the Instructional Value could be found in many objects in a rather arbitrary way. The utility of this additional feature will be clear in paragraph 1.5.

It is useful to overcome these weaknesses and to add some useful information taking up the Higgs's definition and developing a definition in-scope for the ENCORE project.

Definition of Educational Resources for the ENCORE project:

An Educational Resource is a means of transferring Instructional Value. Every other characteristic is meant to enable and maximise this purpose.

An Educational Resource should have the following characteristics:

- ***Instructional value*** *- An educational resource must have some intrinsic instructional value: it should result in a complete learning sequence, objective, skill or competency related to a certain field of knowledge.*

- ***Educational purposes*** *- An educational resource should be designed as a means of delivering Instructional Value, with a specific and clear educational intent from the creator.*

- ***Independent*** *- Educational Resources are discrete and coherent chunks of information, activities or assessment designed by a creator: they can contain a complete learning sequence and don't rely on other material to make sense.*

- ***Shareable/Reusable*** *- Educational Resources are stand alone, reusable components that can be assembled to provide resources in various learning environments, i.e., content developed in*

one context being transferable to another context. This characteristic is fundamental to leveraging any advantage in using educational resources.

- ***Interoperable*** *- Objects must be interoperable: content from multiple sources must work with different learning systems. To do this, they must be designed to conform to some standards.*

- ***Discoverable*** *- The potential learner should be able to find and figure out what an Educational Resource is. This often entails tagging it with appropriate descriptive metadata that will focus on linguistic semantics.*

- ***Context independent*** *- To maximize their reusability, educational resources are required to minimize the amount of information specific to a given context. However, this is often difficult and itis necessary to accept that Learning Objects could include context related information either within the object or by some external association to it.*

The introduction of the Educational Purpose characteristic, taking up the Sosteric and Hesemeier (Sosteric and Hesemeier, 2002) point of view, proves to be important in determining what should be considered an Educational Resource (see some exam ples in paragraph 1.5).

However, this characteristic puts outside the field of possibilities objects that can have great Instructional Value. Such objects, as said, must be considered in some way distinct from the proper Education Resources. To highlight this distinction, we define the Supplementary Educational Resources.

Definition of Supplementary Educational Resources for the ENCORE project:

A Supplementary Educational Resources is an Educational Resources, as previously defined, lacking Educational Purpose. The Supplementary Educational Resources are developed by their creators without any pedagogical intent.

Despite their similarities, this kind of Educational Resources should follow very different evaluation criteria from those of normal educational resources.

To define an Open Educational Resource, we adopt the previous definition with some additional features:

Definition of Open Educational Resource (OER) for the ENCORE project:

An Open Educational Resource is an Educational Resource (as previously defined) that can be freely used, adapted, and shared.

Lastly, we must define what a ROER is.

Definition of Repository of Open Educational Resources (ROER) for the ENCORE project:

A Repository of Open Educational Resources is a free access digital repository, where the Educational Resources can be efficiently searched and used by the user.

We identified on the web a list of digital repositories with free access. This list is provided in Appendix B.

1.4 Types of Educational Resources

In this paragraph, we describe the different types of Educational Resources. They can be different by their file format, e.g. videos, text, etc, but not only: each category can be differentiated into many other subcategories.

A general classification of media types is usually represented into a two-dimensional plane in which one axis represents the Time/space nature, and the other the origin of media.
An example of this classification is represented below in Figure 2:

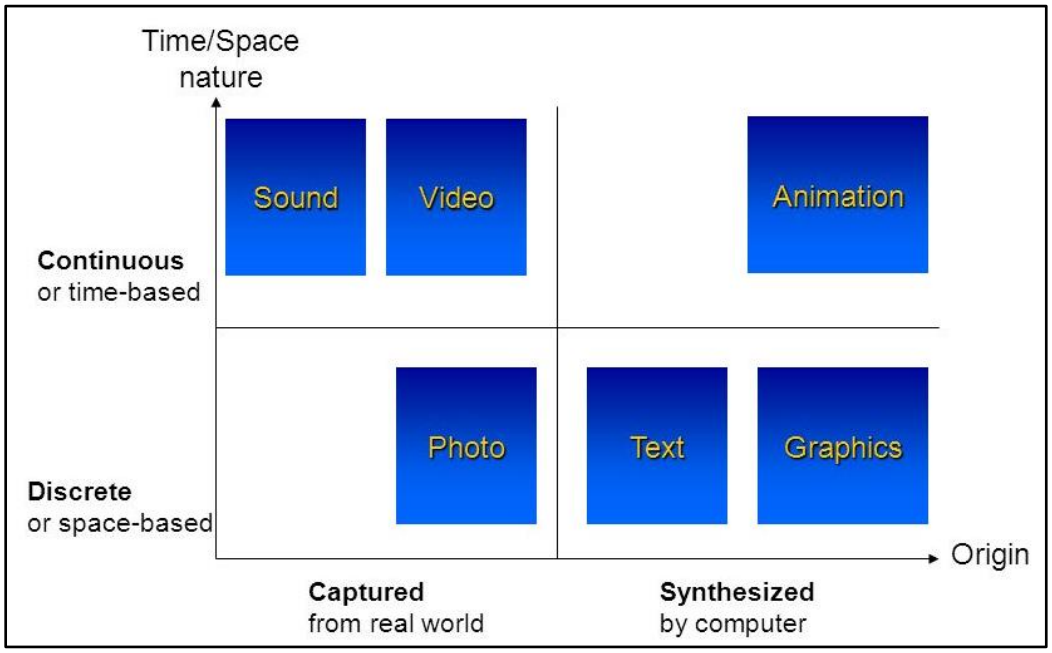


Figure 2 – Representation of media types in a two-dimensional plane in which one axis represents the Time/space nature, and the other the origin of media.

However, our task is more complex since there are Educational Resources expressed through format that do not fall within the classic multimedia content classification.

We propose here a classification useful for the ENCORE project based on six main categories:

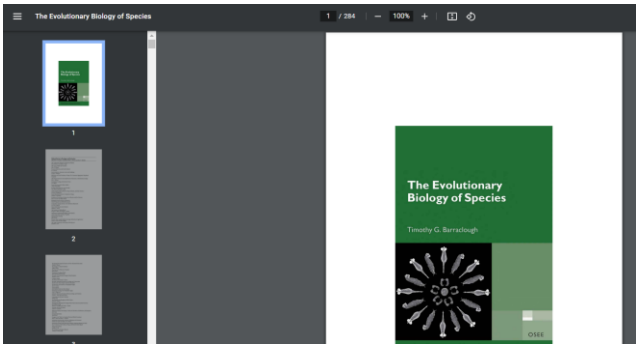
- **Text + Images:** prevalence of text, possible presence of images to a lesser extent in comparison of text.
- **Audiovideo:** videos of any kind, with or without audio.
- **Audio:** only audio.
- **Images + Text:** prevalence of images, single or multiple, with or without text.

- **Exe/Miscellaneous:** resources that present a mix of different kinds of media (like the ones previously stated), with the possibility of interactions from the user.
- **Other:** different formats not considered previously. Usually related to Supplementary Educational Resources.

Category	Examples	File format examples
Text + Images	Scientific papers	docx, pdf, pptx…
	Books	
	Slides	
Audiovideo	Short videos	MOV, AVI, MKV…
	Documentaries	
	Movies	
Audio	Podcasts	mp3, mp4, .wav…
	Songs	
Images + Text	Infographics	pdf, jpeg, png…
	Schemas	
	Diagrams	
Exe/Miscellaneous	Interactive e-learning courses	exe, HTML…
	Educational video games	
Other	Datasets	csv, xlsx, json, rds…
	Programming languages scripts	R, py…
	3D models	CAD, AMF, 3DXML…

Table 1 – Category of Educational Resource's type and related examples.

According to the definition of Educational Resources given in paragraph 1.4 some practical examples of different kinds of E.R. are listed in Table 2:

Name	Type and file format	Example image
The Evolutionary Biology of Species	Text + images; pdf format	

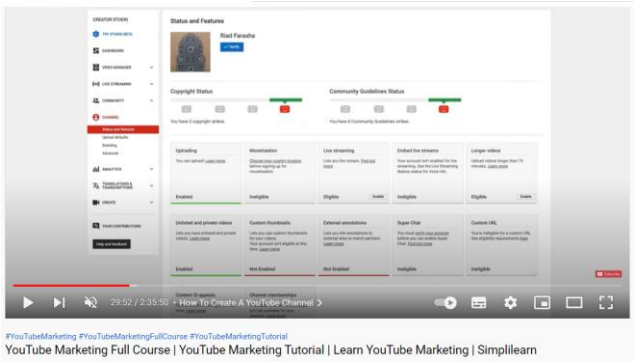
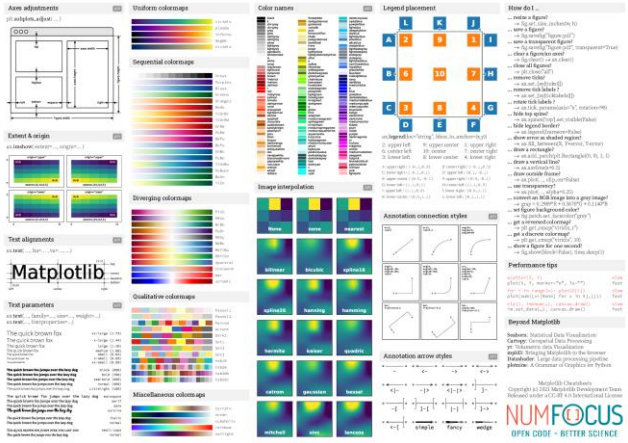
YouTube Marketing Full Course	Audiovideo; mp4	
Python Matplotlib cheat sheet	Images + Text; jpg	

Table 2 – Practical examples of Educational Resources of different types.

1.5 Challenging the Educational Resources' definition

Once we have defined what an OER is, we report a series of challenging examples of possible Educational Resources, that will help us to find the boundaries of our definition.

The examples are described arranged by their challenge level and each one is commented, explaining why it should be considered an E.R. or not.

1. A microscope

Figure 3 – A microscope.

Although this tool is of great use for many students around the world, this cannot be considered an E.R.. First of all, its physical nature compromises Shareably/Reusability property. Furthermore, its instructional value is not clearly definable: the definition provided in paragraph 1.3 tells us that an

Educational Resource “should result in a complete learning sequence, objective, skill or competency related to a certain field of knowledge.”

2. Live lectures

Figure 4 - Live Lectures.

They could not be considered as Educational Resources, since their instructional value is available only to people physically present, violating the Shareability/Reusability characteristics of paragraph 1.3.

3. Podcasts

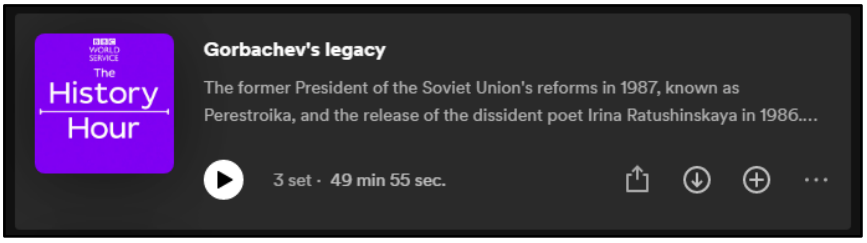


Figure 5 – Podcast example.

They are Educational Resources. It might be uncommon to consider just an audio to be something delivering an instructional value, but of course they do.

4. Infographics

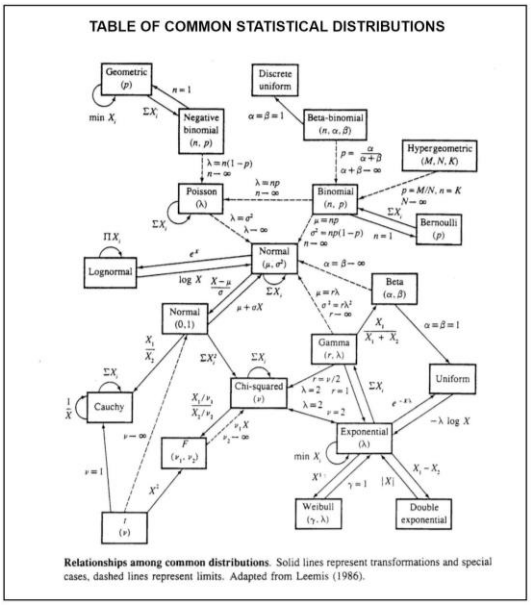


Figure 6 – Infographics.

They should be considered as Educational Resources. The weakness of this kind of resources is that they might be extremely synthetic, as they use just a short text/diagram for conveying meaning.

5. Project blueprints

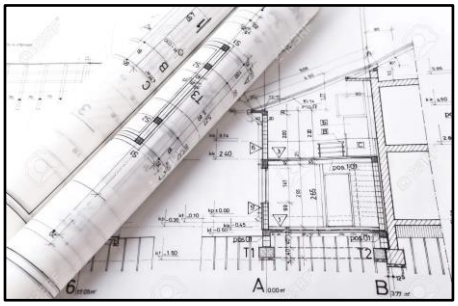


Figure 7 - Project blueprints.

They should not be considered Educational Resources because they lack in the Educational Purposes, in reference to the definition given paragraph 1.3.

6. Documentaries and educational games

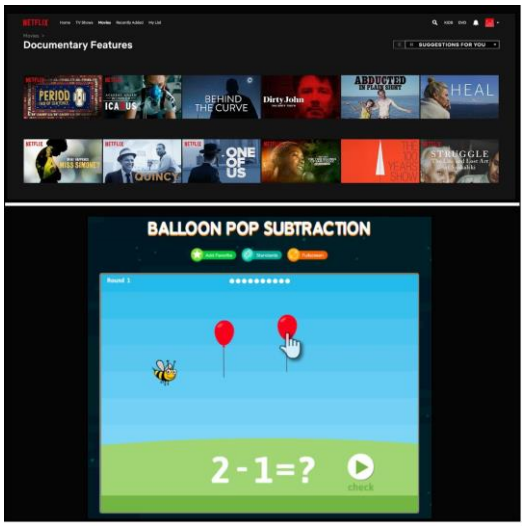


Figure 8 - Documentaries (above) and educational games (below).

They should be considered as Educational Resources if the educational purpose is greater than the entertainment one.

7. Movies and video games

Figure 9 - Movies (to the left) and video games (to the right).

Someone argues that it is possible to learn history, or other subjects, by watching movies or video games. While this is partially true, we must point out that there are far more effective methods to learn the same topics, if the learner is genuinely interested in them. Moreover, it often happens that the care towards the topics covered is not sufficient, precisely because the creators do not focus on this aspect.
These kinds of objects should not be considered as Educational Resources, as the entertainment purpose is far greater than the educational one.

8. E-learning full courses/E-masters

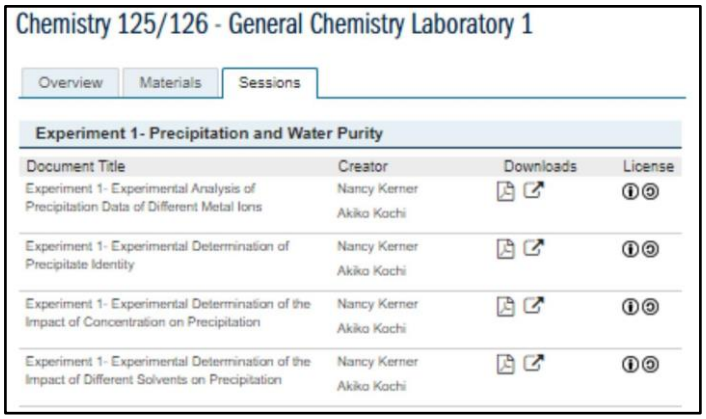


Figure 10 - E-learning full courses/E-masters.

They should not be considered as single Educational Resources, as they are collections of multiple Educational Resources, available as independent and self-contained objects.

9. Literature classics, e.g. War and Peace, L. Tolstoy

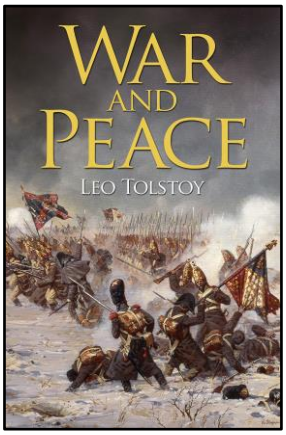
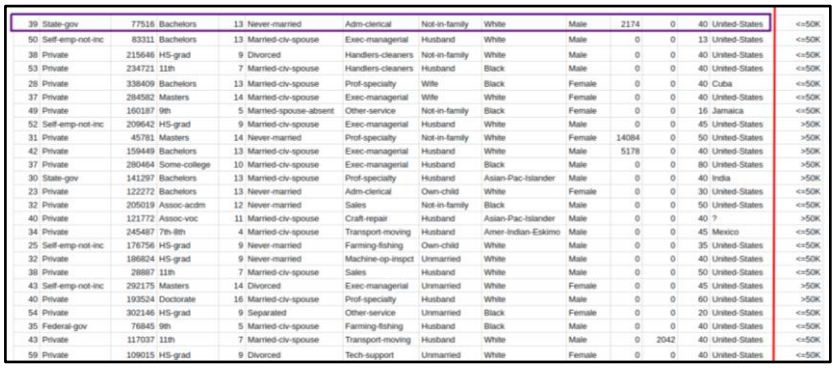


Figure 11 – War and peace by L. Tolstoy.

Even if its educational value is beyond any doubt, these kinds of books are not written with explicit educational intent. They should be considered as Supplementary Educational Resources.

10. Datasets



39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	<=50K
50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	<=50K
38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	<=50K
53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40	United-States	<=50K
28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0	0	40	Cuba	<=50K
37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	White	Female	0	0	40	United-States	<=50K
49	Private	160187	9th	5	Married-spouse-absent	Other-service	Not-in-family	Black	Female	0	0	16	Jamaica	<=50K
52	Self-emp-not-inc	209642	HS-grad	9	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	45	United-States	>50K
31	Private	45781	Masters	14	Never-married	Prof-specialty	Not-in-family	White	Female	14084	0	50	United-States	>50K
42	Private	159449	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	5178	0	40	United-States	>50K
37	Private	280464	Some-college	10	Married-civ-spouse	Exec-managerial	Husband	Black	Male	0	0	80	United-States	>50K
30	State-gov	141297	Bachelors	13	Married-civ-spouse	Prof-specialty	Husband	Asian-Pac-Islander	Male	0	0	40	India	>50K
23	Private	122272	Bachelors	13	Never-married	Adm-clerical	Own-child	White	Female	0	0	30	United-States	<=50K
32	Private	205019	Assoc-acdm	12	Never-married	Sales	Not-in-family	Black	Male	0	0	50	United-States	<=50K
40	Private	121772	Assoc-voc	11	Married-civ-spouse	Craft-repair	Husband	Asian-Pac-Islander	Male	0	0	40	?	>50K
34	Private	245487	7th-8th	4	Married-civ-spouse	Transport-moving	Husband	Amer-Indian-Eskimo	Male	0	0	45	Mexico	<=50K
25	Self-emp-not-inc	176756	HS-grad	9	Never-married	Farming-fishing	Own-child	White	Male	0	0	35	United-States	<=50K
32	Private	186824	HS-grad	9	Never-married	Machine-op-inspct	Unmarried	White	Male	0	0	40	United-States	<=50K
38	Private	28887	11th	7	Married-civ-spouse	Sales	Husband	White	Male	0	0	50	United-States	<=50K
43	Self-emp-not-inc	292175	Masters	14	Divorced	Exec-managerial	Unmarried	White	Female	0	0	45	United-States	>50K
40	Private	193524	Doctorate	16	Married-civ-spouse	Prof-specialty	Husband	White	Male	0	0	60	United-States	>50K
54	Private	302146	HS-grad	9	Separated	Other-service	Unmarried	Black	Female	0	0	20	United-States	<=50K
35	Federal-gov	76845	9th	5	Married-civ-spouse	Farming-fishing	Husband	Black	Male	0	0	40	United-States	<=50K
43	Private	117037	11th	7	Married-civ-spouse	Transport-moving	Husband	White	Male	0	2042	40	United-States	<=50K
59	Private	109015	HS-grad	9	Divorced	Tech-support	Unmarried	White	Female	0	0	40	United-States	<=50K

Figure 12 – Datasets.

They could be very useful for students of Statistical or data science subjects, but they lack of Educational Purposes. They should be considered as Supplementary Educational Resources.

11. Instruction manuals

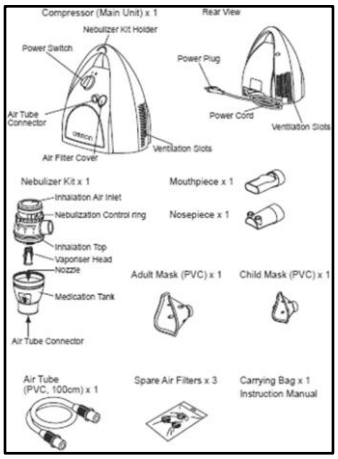


Figure 13 - Instruction manuals.

They are designed with Educational Purposes, they deliver Instructional Value, but they are greatly Context Specific, because they are useful only in relation to the object(s) they are related to and strictly to that specific object(s). For these reasons, they cannot be considered an Educational Resource.

To sum up, in Table 3 we collected the previous examples and stated when an E.R. definition characteristic is fulfilled (X) and when not (O).

	Instructional value	Educational purposes	Independent	Shareable/Reusable	Interoperable	Discoverable	Context independent
1.A microscope	O	O	O	O	O	O	O
2.Live lectures	X	O	O	O	O	O	O
3. Podcasts	X	X	X	X	X	X	X
4. Infographics	X	X	X	X	X	X	X
5. Project blueprints	X	O	X	X	X	X	X
6. Documentaries and educational	X	X	X	X	X	X	X
7. Movies and video games	X	O	X	X	X	X	X
8. E-learning full courses/E-masters	X	X	O	X	X	X	X
9. Literature classics	X	O	X	X	X	X	X
10 Datasets	X	O	X	X	X	X	X
11. Instruction manuals	X	X	X	X	X	X	O

Table 3 – Summary of challenging examples of possible E.R. Only the ones with a full set of X should be considered as E.R. following the definition give in paragraph 1.3.

1.6 Educational Resources' quality definition

Simplifying, we may define the Educational Resources' quality as the *capability to fully deliver the intended Instructional value to the target audience*. In other words, a good quality E.R. make the learner able to fully reach the explicit or implicit learning goals related to the E.R..

To maximise the capability of delivering Instructional Value, we can say that an E.R. should maximise first each of the characteristics reported in the paragraph 1.3 definition. To give a practical example, an Educational Resource that provides well-structured and well-written metadata is better (has better quality) than one that did not report it, because the first is able to maximise its Discoverability characteristic.

More generally, we could say that any aspect of an educational content that improves its ability to transfer Instructional Value affects the quality of the E.R. itself. For instance, if two books are about the same topic, but one does it in a more engaging way, it reduces the stress of the reader and allows a better transfer of the Instructional Value, therefore it can be considered of higher quality. This idea is represented in Figure 14.

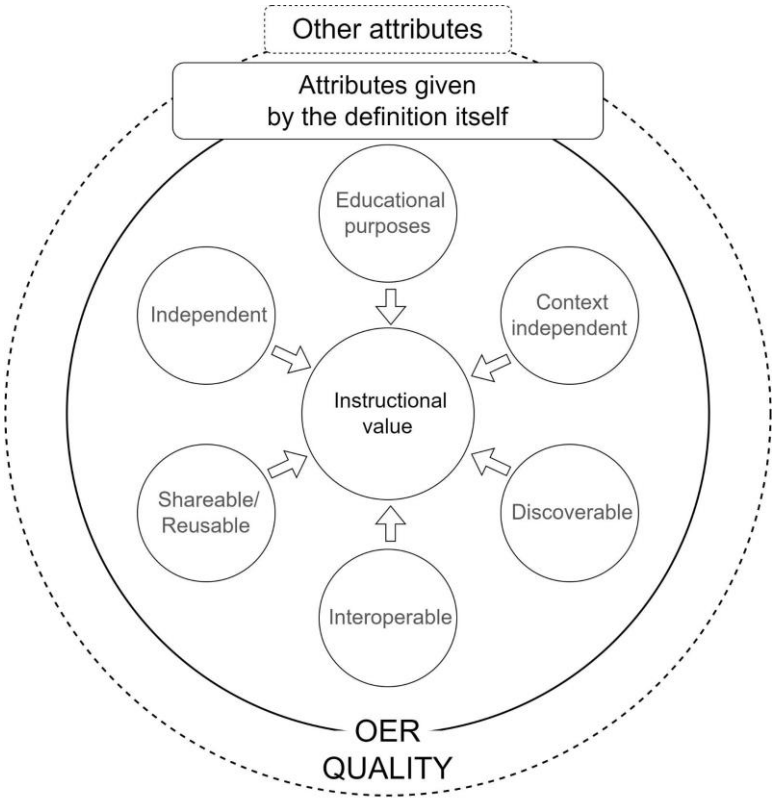


Figure 14 – Representation of E.R.'s attributes influencing its quality.

2. Literature review

2.1 Scope of the literature review

We analysed both quantitative and qualitative methods of OER's quality evaluation in literature, among scientific and non-scientific sources.

As mentioned in paragraph 1.1 we choose to focus on quantitative and automatic metadata-based methods. Therefore, quantitative methods analysed in literature are related mostly to an Indirect Evaluation, i.e. assessing the quality of E.R.'s metadata.

The Educational Objects analysed in literature are of any kind (books, video courses, podcasts, etc) and related to any subject.

Scientific literature is retrieved from different sources (e.g. Google Scholar, Web of Science, etc), while the main one is Scopus. The following table summarises the main queries used to perform the literature review on Scopus, divided into quantitative methods and qualitative methods queries.

Scope	Query	Retrieved papers
Qualitative methods of E.R.'s quality assessment	TITLE-ABS-KEY (("quality" OR "evaluation" OR "assessment") AND ("quantitative" OR "qualitative" OR "data driven") AND ("educational resources" OR "learning objects" OR "training content" OR "OER")) AND SUBJAREA (arts OR busi OR deci OR econ OR psyc OR soci) AND DOCTYPE ("ar")	185
	TITLE-ABS-KEY (("quality" OR "evaluation" OR "assessment") AND ("quantitative" OR "qualitative" OR "data driven") AND ("measure" OR "metric" OR "method") AND ("educational resources" OR "learning objects" OR "training content" OR "OER")) AND SUBJAREA (arts OR busi OR deci OR econ OR psyc OR soci) AND DOCTYPE ("ar")	95
Quantitative methods of E.R.'s quality assessment.	TITLE-ABS-KEY (("quality" OR "evaluation" OR "assessment") AND ("metadata") AND ("quantitative" OR "measure" OR "data driven" OR "metric" OR "method") AND ("educational resources" OR "learning objects" OR "learning" OR "training content" OR "OER")) AND SUBJAREA arts OR busi OR deci OR econ OR psyc OR soci) AND DOCTYPE ("ar")	128
	TITLE-ABS-KEY (("quality" OR "evaluation" OR "assessment") AND ("metadata") AND ("quantitative" OR "measure" OR "data driven" OR "metric" OR "method") AND ("educational resources" OR "learning objects" OR "learning" OR "training content" OR "OER")) AND DOCTYPE ("ar") SUBJAREA (comp)	313

Table 4 – Queries and number of papers retrieved in scientific literature.

Non-scientific literature is retrieved using Google searches. Queries are structured using terms related to education and we selected documents favouring those produced by the main for-profit and no-profit publishers, such as Udemy, Coursera, edX. Also. Most renowned academic publishers were included in the review, such as UC San Diego, Kennesaw State University, University of British Columbia (UBC), Dublin City University and others.

2.2 Scientific literature review

The main results of the scientific literature review are reported in the following table, summarising the most relevant papers identified, with their scope, method and results:

Qualitative methods:

Document	Scope	Method	Results
Elias, M., Oelen, A., Tavakoli, M., Kismihok, G., & Auer, S. (2020, September). Quality evaluation of open educational resources. In European Conference on Technology Enhanced Learning (pp. 410-415). Springer, Cham.	Evaluation performed by OER expert users on OpenCourseWare (OCW)	Evaluation by means of a list of 26 quality drivers, stratified on three dimensions: - Content Structure (CS). - Learning Content (LC). - Self-assessment (SA).	By OER users: - Content Structure is considered useful by 100% of users; - Learning Content is considered useful by 60% of users; - Self-assessment is considered useful by 80% of users.
Haughey, M., & Muirhead, B. (2005). Evaluating learning objects for schools. E-Journal of Instructional Science and Technology, 8(1), n1.	Evaluation performed by teachers in the K-12 education sector (virtual learning).	Evaluation by means of a list of 14 quality drivers, stratified on three dimensions: - Integrity. - Usability. - Learning. - Design. - Values.	The results obtained through the application of Learning Object Evaluation Instrument (LOEI) show that the most relevant criteria found for the evaluation of a learning content can be divided into three groups: - Accessibility criterion; - Student interface; - Pedagogical issues.
Peláez, A. R., Pullaguari, N. P., & Caro, E. T. (2011, April). Quality model proposal for educational material production in OCW sites. In 2011 IEEE Global Engineering Education Conference (EDUCON) (pp. 1074-1080). IEEE.	Quality model aimed to guarantee the quality of OpenCourseWare (OCW) resources.	Quality assurance model based on six main areas: - Development, distribution and licensing models. - Academic range. - Presentation to the user. - Evaluation and support material. - Technological and interoperability requirements. - Accessibility.	Evaluation has been performed on three careers, Economy, Computer Science and Pedagogy. 40% of courses are rated as Acceptable, 20% Desirable, These results could give a clue to the institution considering retributions and incentives to authors based on quality evaluations.

Table 5 – Literature review on qualitative methods of E.R.'s quality assessment.

Quantitative methods:

Paper	Scope	Method	Results
Ochoa, X., & Duval, E. (2009). Automatic evaluation of metadata quality in digital repositories. International journal on digital libraries, 10(2), 67-91.	Digital repository quality assessment.	They present a set of quality metrics for metadata, based on the Bruce & Hillman framework for metadata quality control, that is based on metadata: - accuracy; - completeness; - provenience.	Through statistical analysis, they found that several metrics correlate well with human evaluation and that the average of all the metrics are roughly as effective as people to flag low quality instances.
Tavakoli, M., Elias, M., Kismihok, G., & Auer, S. (2020, July). Quality prediction of open educational resources a metadata-based approach. In 2020 IEEE 20th international conference on advanced learning technologies (ICALT) (pp. 29-31). IEEE.	Automatic quality assessment on OERs, effectiveness of automatic compared to manual assessments.	They propose an OER metadata scoring model, and build a prediction model to anticipate the quality of OERs.	The accuracy of the quality assessment compared with a manual quality control is about 95%.
Aikoh, K., Isoda, Y., & Sugimoto, K. (2020, October). Data Profiling Method for Metadata Management. In 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA) (pp. 779-780). IEEE.	Metadata management, effectiveness of automatic compared to manual assessments.	Using feature extraction AI, authors extracted metadata from 34 documents. They then performed a manual judgement on the correctness of extracted results from documents.	In their experiment, they obtain the result that the correctness can be determined with 77% accuracy by the metadata profiling method they proposed.
Palavitsinis, N., Manouselis, N., & Sánchez-Alonso, S. (2013). Metadata quality issues in learning repositories.	Metadata quality assurance method applied on Learning Repositories.	They used a mixed approach, using expert judgments and quantitative methods. They used six metrics: - Completeness. - Accuracy. - Consistency. - Objectiveness. - Appropriateness. - Correctness.	They supported metadata definition process throughout each different phases of the repositories development and for each phase, they measured the resulting metadata quality.

Table 6 – Literature review on quantitative metadata-based methods of E.R.'s quality assessment.

2.3 Non-scientific literature review

The table below summarises the main findings of the non-scientific literature review:

Type	Document	Scope	Summary
Company	(Coursera) Drivers of Quality in Online Learning How to Increase Engagement, Satisfaction, Skill Development, and Career Impact Worldwide. Alan Hickey, Staff Data Scientist Alexandra Urban, Senior Teaching & Learning Specialist Eric Karsten, Data Scientist	Video courses, interactive video courses	Ex-post quality metrics, identifying four main quality dimensions: 1) Engagement. 2) Satisfaction. 3) Skill Development. 4) Career Outcomes.
	(Udemy) Course Quality Checklist To publish on Udemy, these are the minimum course requirements	Video courses, interactive video courses	Document divides quality into categories and defines a set of short criteria to assess the quality of video courses and interactive video courses. Focuses on technical quality criteria, course design and teacher's skills
Academic	Quality Online Course Checklist	Video courses, interactive video courses, live lectures, textual materials	Introduce quality criteria for instructional materials, learning activities and student support
	KSU Course Quality Checklist	Live lectures, textual contents	Focuses on pedagogy, structure, navigation, course objectives, module objectives course content, alignment, assignment, assessment communication, engagement, interaction, active learning
	Quality checklist: questions for designing and delivering online courses	Video courses, interactive video courses	Full online courses ontology of design elements and quality drivers

Table 7 – Non-scientific literature review on quantitative metadata-based methods of E.R.'s quality assessment

2.4 Key points of the Literature review

Referring to the qualitative methods of quality assessment, e.g. quality guidelines applied by teachers or auditors, most of the papers share some categories of investigation: the technological aspects, the pedagogical aspects, the one of accessibility.

The main effort from researchers has been devoted to identifying and list potential quality criteria, generating ontologies meant to be as complete as possible. However, the literature lacks a vision that includes all possible kinds (books, videocourses, etc) of E.R., recognizing the common aspects and at the same time the specificities of the different E.R. formats.

In fact, the E.R. quality drivers change according to the typology of E.R. and their relative features, e.g. the video resolution is an important quality criteria to evaluate video courses, but it is meaningless to evaluate a book.

Non-scientific literature is extremely accurate in defining evaluation criteria and metrics, providing actionable methods for the quality assessment, but they mainly focus on a specific category of E.R.: interactive video courses (which also mostly are short videos format). However, the non-scientific literature provides excellent ideas for the evaluation of this specific kind of E.R., with well-defined and practical checklists.

Relating to **automatic methods**, we mainly analysed approaches able to perform E.R.' indirect evaluation, assessing the quality of OERs' metadata. The idea of using metadata quality as a method of predicting the quality of training content is not new (Tavakoli & Elias, 2020). Most of the papers share some metrics used to evaluate metadata quality. The most relevant are:

- **Completeness**: a metadata instance should describe the resource as completely as possible. This means that most metadata fields should be filled in to make resources useful for any kind of service. While this definition is based on the static library instance view of metadata, it can be used to measure how much information is available about the resource.

- **Accuracy**: the information provided about the resource in the metadata instance should be as correct as possible; in fact, typographical errors, as well as factual errors, affect this quality dimension.

- **Logical Consistency** and **Coherence**: metadata should be consistent with standard definitions and concepts used in the domain. The information contained in the metadata should also have internal coherence, that means that all the fields describe the same resource and contain coherent information with Educational Resources' content.

- **Provenance** or **Authority**: The source of the metadata could be another factor to determine its quality. Knowledge about who created the content and the level of his expertise could provide insight into the quality of the instance.

Accuracy and logical consistency or coherence are often merged into a more general metric called the "Quality of free text" (Ureña-Cámara et al., 2019). The latter considers that there are neither lexical or spelling errors nor errors that would lead to an information misalignment compared to the content of the resource described by the metadata.

As far as *authority* is concerned, some potential issues emerge. The main one is the conversion of the authority parameter into a metric.

Another important consideration is that metrics are defined by taking metadata standards as a reference, therefore the choice of the standard can lead to different design choices in the definition of a set of metrics. This specific issue will be discussed in paragraph 3.1.

3. Automatic quality assessment criteria

3.1 Metadata standard

To perform an Indirect evaluation of OER quality (as described in paragraph 1.6), we have to choose a reliable standard that defines which metadata should be evaluated.

Dublin Core Metadata Element Set (DCMES)

Dublin Core Metadata Element Set (DCMES), also known as Dublin Core (DC) standard, is a set of fifteen "core" elements (properties) for describing resources. These fifteen-elements have been formally standardised as ISO 15836, ANSI/NISO Z39.85 and IETF RFC 5013. The resources described using the Dublin Core could be digital resources (video, images, web pages, etc.) as well as physical resources such as books or works of art.

The DC metadata set is the most widespread metadata format today, the result of a project started in 1995 [Dublin Core Metadata Workshop Series] with the aim of addressing the problem of finding information resources available on the net. Its characteristics - the result of a precise choice - are generality and simplicity: the standard has been designed as a minimum set of descriptive elements capable of representing any information resource - web page, audio, video etc. - that has a WEB address.

The Dublin Core Metadata Element Set (DCMES) is based on guideline principles that play an important role in the satisfaction of need related to educational resources.

The main principles can be declined in: *extensibility*, that stays for the creation of additional elements for the creation of specific metadata; *interoperability*, according to the national or local standards; *modularity*, that refers of the community to switch from a general schema for metadata to the use of a specific schema required to meet a group of educational needs.

IEEE Learning Object Metadata (LOM)

Another widely adopted metadata standard set for this purpose is IEEE Learning Object Metadata (LOM), a standard which has been designed especially for the description of educational resources. Although the adoption and influence of the LOM has been considerable, there are some issues to be highlighted.

First, the LOM conceptual data schema is not based on an abstract model shared with other metadata schema and does not align with the guideline principle of interoperability. In this way, it is impossible to import elements from other metadata schemas, such as Dublin Core one. Another important key point is the difference of approach between LOM and DCMES. LOM works on “instance”, that is a block of metadata, while Dublin Core wants to define individual terms and their rules of applications.

MAchine Readable Cataloguing - MARC

MARC stands for MAchine-Readable Cataloguing record; the formats are standards for the representation and communication of bibliographic and related information in machine-readable form.

"Machine-readable" means that one particular type of machine, a computer, can read and interpret the data in the cataloguing record. "Cataloguing record" means a bibliographic record, or the information traditionally shown on a catalogue card. The record includes (not necessarily in this order): 1) a description of the item, 2) main entry and added entries, 3) subject headings, and 4) the classification or call number.

Metadata Object Description Schema (MODS). Metadata Object Description Schema (MODS) is a schema for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. This schema was created to bridge the gap between the simplicity of DC and the complexity of MAchine Readable Cataloguing (MARC). This could create the impression that MODS was developed as a competitor to the Dublin Core because DC started as a very simple scheme, but with its element refinement qualifiers it is becoming complex; meanwhile, MODS is simplifying MARC, which is known to be very complex, but always refers to regulations that are not general but specific to the field of application.

In the end, we choose the **Dublin Core (DC) standard:** the main reasons why DC is the most suitable standard for evaluating E.R. are that:

- It is applicable to all types of information resources.
- It is simple to understand.
- It is of general use for the "discovery" of information resources (unlike the catalographic descriptions that generally follow specific regulations of the sector of application: libraries, museums, video libraries ...).

The following table contains the metadata defined by the Dublin Core and their descriptions:

Dublin Core Element	Description
Title	A name given to the resource.
Subject	The topic of the resource.
Description	An account of the resource.
Creator	An entity primarily responsible for making the resource.
Publisher	An entity responsible for making the resource available.
Contributor	An entity responsible for making contributions to the resource.
Date	A point or period of time associated with an event in the lifecycle of the resource.
Type	The nature or genre of the resource.
Format	The file format, physical medium, or dimensions of the resource.
Identifier	An unambiguous reference to the resource within a given context.
Source	A related resource from which the described resource is derived.
Language	A language of the resource.
Relation	A related resource.

Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
Rights	Information about rights held in and over the resource.

Table 8 – Dublin Core metadata standard.

3.2 Overview of chosen metadata quality criteria

Once a metadata standard is chosen, we must define which quality drivers chose to evaluate OER's metadata quality.
We chose the following **quality drivers**:

A. Metadata completeness (**Completeness**).
B. Publishers' authority (**Authority**).
C. Description's quality of free text (**Quality of free text**).

Quality drivers meaning, the related hypothesis and quantitative metrics chosen are summarised in the following table:

Quality driver	**Meaning and hypothesis**	**Metrics**
A.	Meaning: a metadata instance should describe the resource as completely as possible. HP1: A good quality OER likely has well cared (complete) metadata. HP2: A good quality ROER should have well cared metadata	Completeness
B.	Meaning: Knowledge about who created the content and the level of his expertise could provide insight into the quality of the instance. HP1: An OER from a renowned publisher is a good quality one HP2: A good quality ROER has as much as possible OERs from renowned publishers.	Authority: publisher recognition from a list of reliable publishers
C.	Meaning: is the ease with which a reader can understand a written text. HP1: A good quality OER has a well articulated description. HP2: A good quality ROER has as much as possible OERs with well articulated descriptions.	Readability

Table 9 – Quality driver meaning, hypothesis and related metrics.

Quality driver's metric are explained in paragraph 4.1 (A), paragraph 4.2 (B), paragraph 4.3 (C).

3.3 Putting the metrics together

A full understanding of the overall quality assessment method could be obtained in Appendix A. In this paragraph we want to explain the logic with which the different quality drivers and metrics are put together.

Each of these quality drivers and metrics has a different function in the automatic quality assessment. The following schema resumes the overall method (Figure 15).

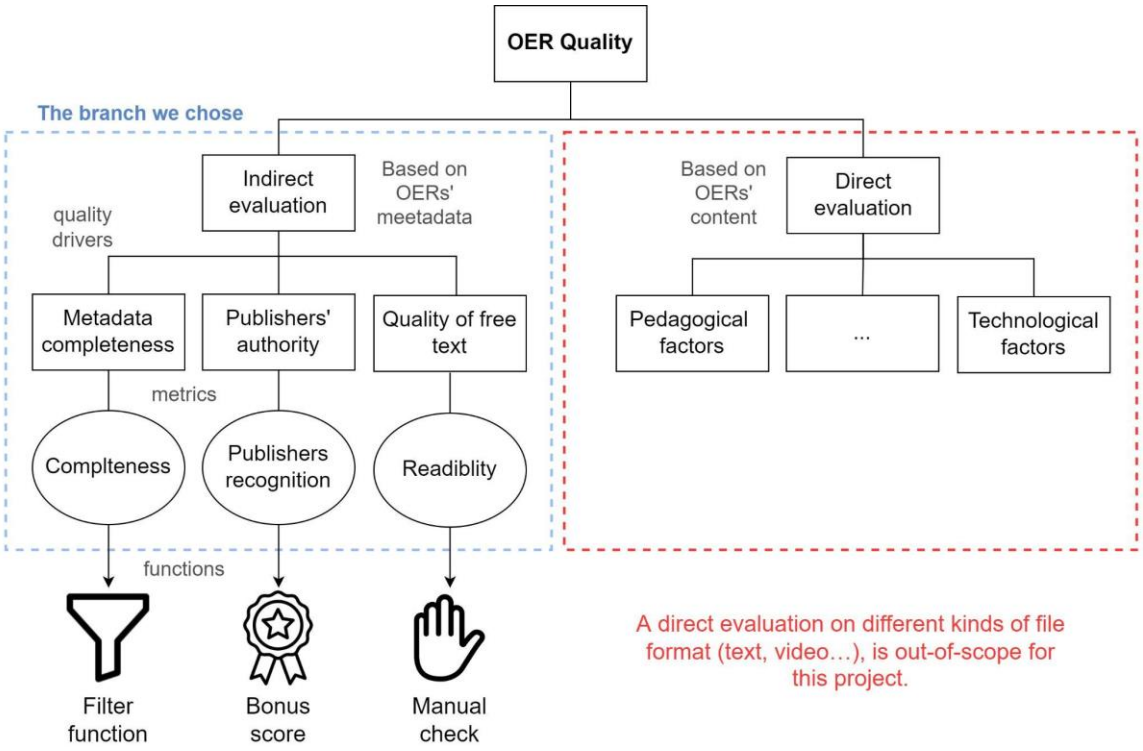


Figure 15 – A resume of the strategy chosen to measure OER's quality.

- The "**Completeness**" metric is used as a filter metric. OERs that do not have a minimum number of correctly filled metadata should be removed or put aside.
- The "**Authority**" metric is used to give a bonus score to some OERs. OER's with recognized publishers are considered to have a greater probability of having a good quality.
- The "**Readability**" metric is used just to manually check some OER with extreme values.

4. Quality metrics

4.1 Metadata completeness

4.1.1 Metadata applicability and importance

Considering the scope and the objectives of the ENCORE project, not each metadata defined by the Dublin Score Standard has been considered and they have not the same importance to achieve our goals.

There are two DC metadata fields (as seen before in Table 8) considered as not relevant for the automatic quality assessment:

- **Language**: the ENCORE project it's only about English language OERs.
- **Rights**: the ENCORE project it's only about open educational resources.

Therefore, these two metadata will not be covered further in this document and in the algorithm developed.

Furthermore, we must group metadata by their importance in the context of the ENCORE project and the quality assessment. We split the metadata into two categories:

- **Mandatory:** OER lacking metadata from this group should be removed from the ROER.
- **Optional**: OER presenting this kind of information have to be considered better in quality than the others.

Title, Subject, Creator, Date and Format metadata fields are considered the *mandatory* information with which an Educational Resource can be made Discoverable (in relation to the definition given in the paragraph 1.3). This is an essential condition to effectively developing the potential of an Educational Resource. As a consequence, lack of these metadata is considered as a strong signal of low quality OER. Furthermore, we also consider the **Description** metadata field as *mandatory,* not only because of its importance in making an E.R. discoverable and comprehensible to the user, but also because in the Encore Project Task 3.4 this metadata field will be useful to find relevant information.

On the other hand, we considered the **Publisher** metadata field as *optional*, because an OER could be developed by an independent creator that does not have any Publisher. All other DC metadata fields are considered *optional*, so as a further indication of the good quality of the OER. If these kinds of metadata are missing, it's not necessary to remove the OER.

The following table summarises what just explained:

Dublin Core Element	Group	Explanation
Title, Subject, Creator, Date, Format, Description	Mandatory	This set of metadata constitutes the basis for the identification of an OER.
Publisher, Type, Contributor, Identifier, Source, Relation, Coverage	Optional	This set of metadata facilitates potential students or educators searching for OERs.

Table 10 – Groups of DC metadata fields defined for the development of the Completeness metric.

4.1.2 Scoring method

The scoring method of OER using the Completeness metric is described in the following schema (Figure 16):

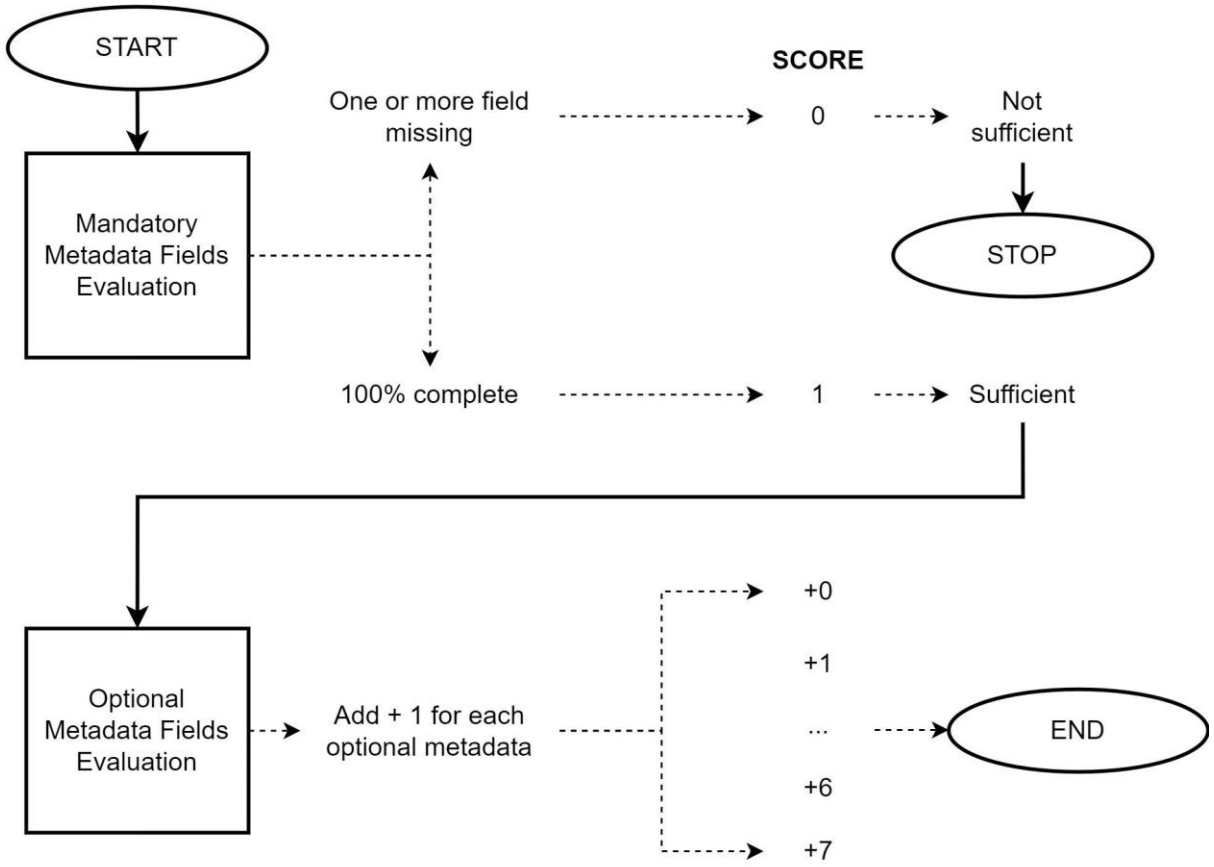


Figure 16 – Scoring method of OER using the Completeness metric. OER lacking mandatory metadata should be filtered from the ROER.

We chose to make no distinction in the scoring of each Optional metadata because we have no concrete evidence to determine if one metadata is more important than another.

The nine different possible outcomes of the scoring method are the one shown in Table 11.

Completeness outcomes	Description
0	Not sufficient. OER should be filtered.
1	Only Mandatory metadata field are filled. OER should be maintained.
2, 3, 4, 5, 6, 7	We have other information in addition to Mandatory ones, but something is missing
8	Full metadata. OER likely of a good quality.

Table 11 – Possible outcomes of the metadata Completeness scoring method.

4.1.3 Metadata completeness algorithm

About some metadata, Dublin Core suggests possible (but not unique) compliance rules, e.g. they suggest to use the DCMI Type Vocabulary to fill the “Type” field, meanwhile other metadata doesn't receive any suggestions.

In literature, there are no specific standards to define the correct filling of D.C. metadata used to describe Educational Resources, so we adopt a few straightforward criteria, summarised in the following bullet points:

- The metadata is not a missing value (NA values);
- The metadata contains at least two alphanumeric characters (alphabetic characters A through Z, numeric characters 0 through 9);
- The metadata contains less than 10.000 alphanumeric characters.

The following flowchart shows the steps of the algorithm with the scoring method (Figure 17):

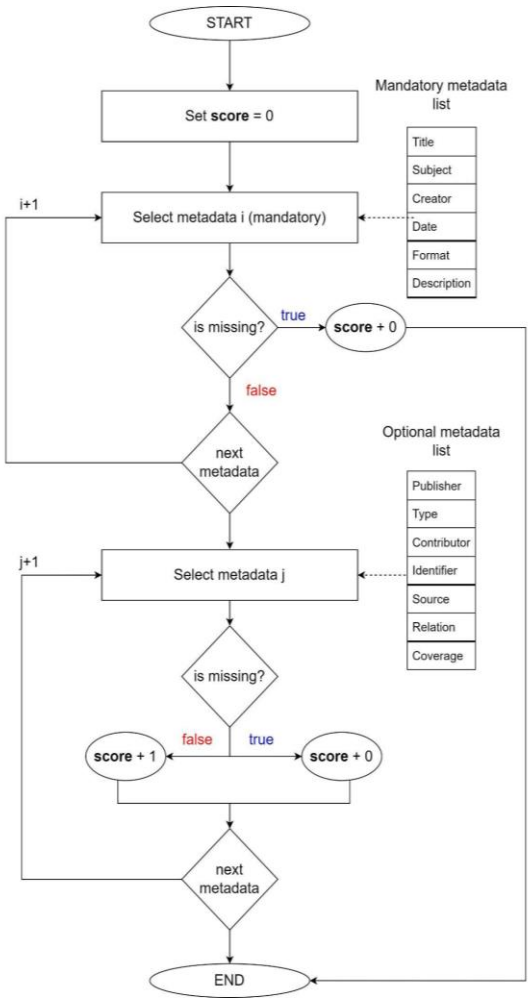


Figure 17 – Metadata Completeness measurement method.

4.2 Publishers' authority

4.2.1 Publishers' authority

To recognize the publisher of an OER stored in a repository, we must rely on an external corpus of data. Unfortunately, no complete data sources containing a full list of E.R.'s publishers are available on the market or over the internet (considering any kind of E.R.: video courses, books, podcasts ecc.). Thus, we have designed a method to create the dataset.
Considering the large amount of data ad data sources available, we decided to collect publisher information following these criteria:

What and why: we collected most important academic publishers of educational contents (mainly books), assuming that an academic publisher has to be considered a reliable publisher.

Where: data included in the list should be collected from reliable data sources that include publishers from all over the world.

When: active publishers in the last 40 years.

How much: at least the most famous publishers should be included in the list.

4.2.2 Academic publishers and universities list

The "Academic publishers and universities list" development pipeline is shown in the flowchart below (Figure 18):

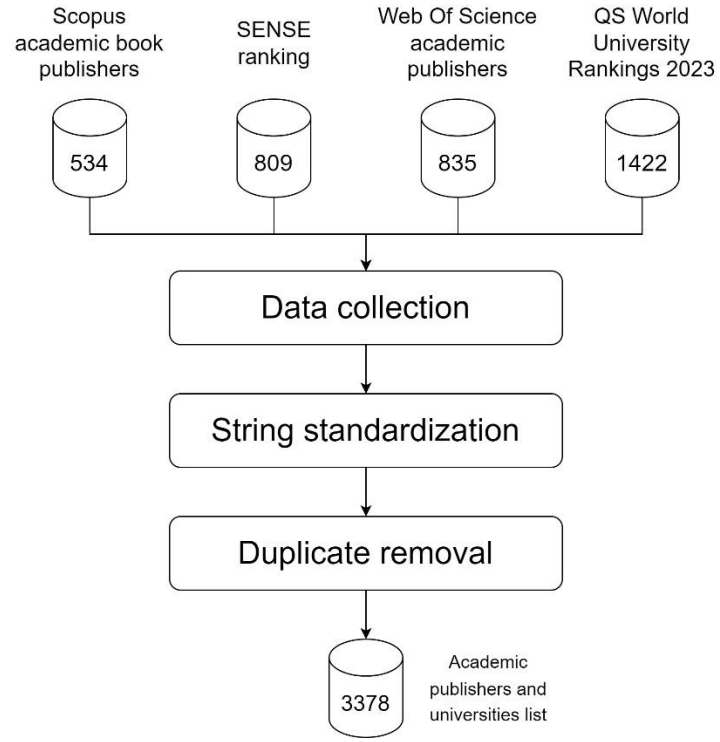


Figure 18 – Academic publishers and universities list development pipeline.

In the following table, we describe the previously shown data sources and the amount of data collected.

Data source	Description	Number of publishers
Research School for Socio-Economic and Natural Sciences of the Environment	Ranked list of academic publishers compiled by the data source	809
Clarivate Web of Science	Flat list of academic publishers compiled by the data source	835
Scopus	Flat list of academic book publishers taken from a sample of 30000 highest citation books published from 1980 to 2022 on Scopus.	535

Table 12 – Data sources used to develop the Academic and universities publisher list.

We must consider that in the lists there are lots of duplications of publishers' names (duplication rate equal to 7%.). The final overall number of unique publishers' names is 1985.

Analysing the list, we noticed that often an academic publisher (e.g. Oxford University Press) is substituted by the name of the university itself (e.g. Oxford University). To include in the recognition task this kind of inaccuracies, we add to the academic publisher list a list of worldwide most important universities, taken from the QS World University Rankings 2023, including 1422 universities names.

The full list, named "*Academic publishers and universities list*", is made of 3378 items and an example is presented in the table below:

Name	From
American Chemical Society	Scopus, SENSE
Channel View Publications	Scopus, SENSE, Web Of Science
Massachusetts Institute Of Technology	QS World University Rankings 2023
…	…

Table 13 – Examples of academic publishers included in the Academic publishers and universities list.

Then, we consider as recognized academic publishers the one included in the list described above. If a publisher is not included in this list, it does not mean that it is a low-quality publisher. The purpose of this list is basically to include all the most internationally renowned.

4.2.3 Publishers recognition rules

Once the academic publishers list has been completed, as explained in the previous paragraph, we must define the criteria with which the items in the list are recognized in a ROER.
First, given an OER, it is important to understand where to find the different metadata.

Dublin Core Standard provides precisely the "Publisher" metadata, described as “An entity responsible for making the resource available”, as shown in Table 8.

To recognize the “Publisher” of a given OER in the ROER, if it’s not a missing value, it can be found by searching the string, in the Academic publishers and universities list just described. To perform this task, there are two ways:

- Search the exact string match (strict method).
- Admit slight variation in string matching (flexible method).

We chose the second method. In fact, variations in the data entry could appear frequently, due to lexicographical errors, word inversions or usage of pseudonyms and should be bypassed.

Some examples of strings variations due to incorrect data entry:

Grammar errors: Univesrity of Pisa instead of University of Pisa.

Word inversions: Pisa University instead of University of Pisa.

Pseudonyms: UniPi instead of University of Pisa.

To avoid this kind of problems and perform a flexible string matching approach, we used the following methods:

Issue	Method	Description	Threshold
Grammar errors	Levenshtein distance	The minimum number of single-character edits (insertions, deletions or substitutions) required to change one word into the other	< 2. Only the variation in one character is allowed.
Word inversions	Cosine similarity	Measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them	= 0. Identify the same words in different orders.
Pseudonyms	Pseudonyms substitution	List of 170 most common universities pseudonyms matched with full names	\

Table 14 – Issues and method used to solve them in the Publishers’ recognition algorithm.

To measure the Levenshtein distance and Cosine similarity, we used the well-known *“stringdist”* R programming language package, using the *stringdist()* function.

4.3 Quality of free text

4.3.1 Evaluating metadata quality of free text

As explained in paragraph 3.3, we want to evaluate the quality of free text only to add useful information related to the ROER's content and to allow a manual check on the database. Therefore, the application of this metric should not be considered as a fundamental part in the automatic evaluation of the OERs stored in the ROER.

The free text we chose to evaluate is the one contained in the DC's "Description", as shown in Table 8. In fact, it is reasonable to evaluate the quality of free text when:

- data entry is not defined by a standard format or in a limited range of options;
- the amount of text expected is enough to allow the evaluation ("enough"depends on the chosen metric).

We chose to evaluate the Quality of Free Text metadata quality driver using a Readability metrics, as already done many times in literature (see chapter 2).

There are a lot of readability metrics known in literature, the following table shows some of them:

Readability metric	Formula inputs	Scale/benchmark
Flesch_Kincaid:	words per sentence, syllables per word	Scores related to school grades
Flesch reading ease:	words per sentence, syllables per word	90-100 : Very Easy 80-89 : Easy 70-79 : Fairly Easy 60-69 : Standard 50-59 : Fairly Difficult 30-49 : Difficult 0-29 : Very Confusing
Gunning_Fog:	long/complex words as part of the total	Score above 12 is considered hard
Coleman_Liau:	letters per 100 words, sentences per 100 words	Used by U.S. Office of Education to calibrate the readability of all textbooks for the public school system
Dale_Chall:	presence of "hard" words	ADJUSTED SCORE GRADE LEVEL 4.9 and Below Grade 4 and Below 5.0 to 5.9 Grades 5 - 6 6.0 to 6.9 Grades 7 - 8 7.0 to 7.9 Grades 9 - 10 8.0 to 8.9 Grades 11 - 12 9.0 to 9.9 Grades 13 - 15 (College) 10 and Above Grades 16 and Above (College Graduate)
Automated Readability Index (ARI):	letters per word, words per sentence	The Automated Readability Index outputs a number that approximates the age needed to understand the text
Linsear_Write:	percentage of long words (3 syllables or more)	results in a "grade level" measure
Spache:	sentence length, unfamiliar words (3 grade +)	For grades 3 and under. For higher, use Dale_chall

Table 15 – A summary of readability metrics known in literature.

We chose to use the *Flesch reading ease metric*, because of its simplicity and its wider use in literature than other metrics.

It is hard to define if a given OER, regardless of its file format, subject, etc, should have a low readability (complex and sophisticated text) or a high readability (simple and plain text), but we think it would be useful to check extreme values in both that directions.

4.3.2 Flesch reading ease metric

Flesch reading ease metric provides an absolute scale with which it is possible to evaluate the reading ease of a given text.
The formula to calculate the Flesch reading ease is the described by the following equation:

$$206.835 - 1.015(total\ words/total\ sentences) - 84.6(total\ syllables/total\ words)$$

The ranges of values to understand the metric results are the following:

Score	Description
100.00–90.00	Very easy to read. Easily understood by an average 11-year-old student.
90.0–80.0	Easy to read. Conversational English for consumers.
80.0–70.0	Fairly easy to read.
70.0–60.0	Plain English. Easily understood by 13- to 15-year-old students.
60.0–50.0	Fairly difficult to read.
50.0–30.0	Difficult to read.
30.0–10.0	Very difficult to read. Best understood by university graduates.
10.0–0.0	Extremely difficult to read. Best understood by university graduates.

Table 16 – Flesch reading ease metric scores description.

To calculate Flesch reading ease has been used the *quanteda* R programming language package, with its function *textstat_readability().*

Some examples of readability score in Udemy courses' descriptions:

High reading ease (near 100): short sentences, easy words.

The course helps you to answer these questions walking through different areas: Costs, Profit, Break-even, Brand Value, Competition, Steps Processes, Sales, Price Review

- Do I see a drastic change in sales or costs?

- What are my trends month over month, year over year?

- What is my estimation on my business`s future profitability?

- Shall I invest more?

- What are my biggest expenses? Can I somehow reduce those?

[...]

Low reading ease (near 0): long sentences, difficult words.

My course explains in "plain English" all the requirements of ISO 9001:2015 discussing the intent of each sub-clause of the standard, offering solutions that can be used for the implementation of a quality management system along with examples and, of course, providing guidance on how to audit a quality management system be it internal or external audit. After going through all the lessons in the course you will have a perfect understanding of ALL the requirements of ISO 9001:2015.

[...]

5. Conclusions

The quality assessment of Educational Resources is a complex topic, open to numerous possible developments, both as regards the direct evaluation of the content itself, and the indirect evaluation through E.R.'s metadata. In fact, the very definition of E.R. and their quality is currently under discussion and can be seen through different perspectives (McGreal, 2004).

The quality assessment of Repositories of Educational Resources, as seen in Chapter 1, is closely linked to the one of E.R. In fact, although the design and the strictly technical functionalities of the database affect the user experience, it is mainly determined by the quality of the E.R. that can be found inside and by the metadata with which they are described.

The method developed in this document refers to an Indirect quality evaluation approach, based on E.R. metadata. The developed method has the advantage of being extremely efficient and of guaranteeing the possibility of analysing large quantities of E.R., because it could be completely performed by an algorithm. It provides an overall assessment of the quality of the repositories and it's able to filter unwanted contents. The result, however, can be improved in several respects:

- the metadata Completeness metric control rules can be improved checking for more subtle inconsistencies in metadata fields, e.g. a generic text in place of the E.R.'s Title;
- The "*Academic publishers and universities list*" can be improved including non-academic publishers and, moreover, each item can be qualified by a publisher reliability ranking;
- The Quality of Free Text metric can be evaluated through more complex systems of textual analysis, capable of detecting grammatical errors, syntactic inconsistencies and other more complex textual features.

In addition, more reliable systems could be developed in the future for the cross-checking of the consistency between the different metadata fields and comparison between the metadata and the training content itself.

Appendix.A - Automatic quality assessment procedure

1. Steps of the automatic quality assessment

The current Appendix explains the methodology followed for the automatic ROER quality assessment and the reasons that guided our design choices.

The steps to follow are five:

1. **Code execution**
2. **Residuals analysis**
3. **Readability scores check**
4. **ROER quality metrics resume**
5. **Reporting**

The result of this evaluation should be documented in a specific report as an output, planned in step 5.

1.1 Step 1: Code execution

We prepared the code to execute using the R programming language.
The code needs the ROER as input in a tabular file format. The execution will be performed running the R script prepared. The output will be the ROER with some added variables, as shown in the following flowchart (Figure 19):

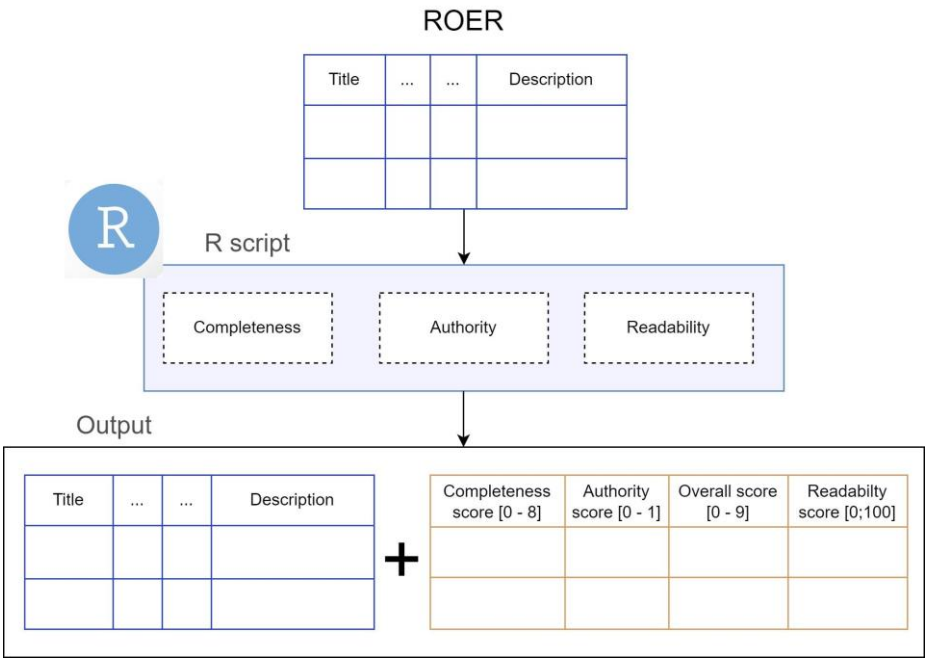


Figure 19 – An overall schema of the input and output of the algorithm.

The output variables shown above, are:

- **"Completeness score":** a discrete numeric variable ranges from 0 to 8, measuring Completeness metric, as explained in paragraph 4.1.

- **"Authority score":** a boolean True/False variable showing if the publisher has been recognized as an academic publisher, as explained in paragraph 4.2.

- **"Overall score":'** a discrete numeric variable ranges from 0 to 9, obtained as the sum of Completeness and Authority scores (only when Completeness is > 0, otherwise the Overall score is set to 0).

- **"Readability score":** a numeric variable ranges from 0 to 100, as explained in paragraph 4.3.

1.2 Step 2: Residuals analysis

As explained in chapter 3, the Metadata Completeness metrics is intended as a filter. Thus, it is suggested to filter the OERs that got a 0 score in Step 1.

On the other hand, it's important to verify that the overall quality of these 0 score OERs is low, to validate the quality assessment method. A method to do so is to take a sample of these kinds of OERs and manually check their quality (we suggest a stratified proportional sampling method on the OERs Type) .

It's also important to synthesise the result of this evaluation in a document called "ROER quality assessment report". If the overall quality of this group of OERs is established to be on average low, is it possible to eliminate them all from the ROER and to validate the method.

It is also possible not to remove these OERs, but rather putting them asid, e.g.. if on average they are valuable OERs but they have just a metadata problem, it will also be possible to fix it later.

1.3 Step 3: Readability scores check

As explained in paragraph 4.3, the Readability metric can be useful to identify outliers in the ROER, e.g. OERs out of our target, aimed at children and so described with a childish language.

To perform this manual check, we suggest first of all to arrange the ROER by the Readability score in an ascending order and check the first OERs appearing, evaluating the possibility of eliminating unwanted content and then doing the same arranging the ROER in a descending order.

The results of this evaluation and the result of any removals should be reported in the "ROER quality assessment report".

1.4 Step 4: ROER quality metrics resume

To resume the results of the ROER evaluation, giving an overall quality assessment, is useful to plot the histogram of the "Overall score" variable added to the ROER, to verify the percentage of OERs falling in each bin.

We could expect a distribution similar to the ones shown in the following figure (Figure 20):

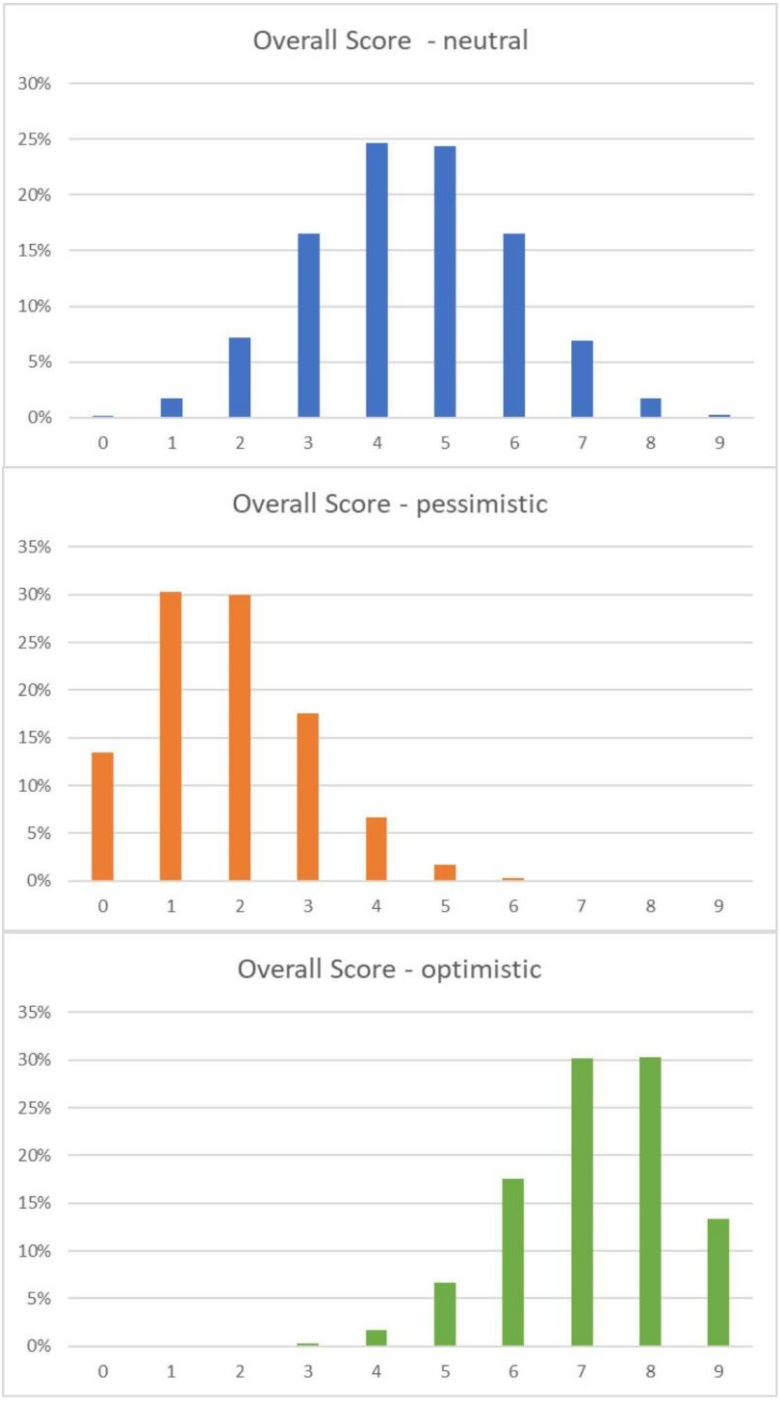


Figure 20 – Overall score examples.

- Generally, extreme values (e.g. score 0 and 9) should be less widespread, giving a bell-like shape to the histogram (is it possible to approximate the expected distribution as a Binomial distribution).
- The higher the average value, the better is the quality of the ROER (pessimistic scenario = left skewed, optimistic scenario = right skewed) .

Strong variations from this description should be considered as index of low ROER quality or wrong appliances of the automatic assessment.

Other metrics to take into account are:

- % recognised publishers / unrecognised or missing publishers;
- % removed OERs / total amount of OERs;
- distribution of missing metadata (below in Figure 21).

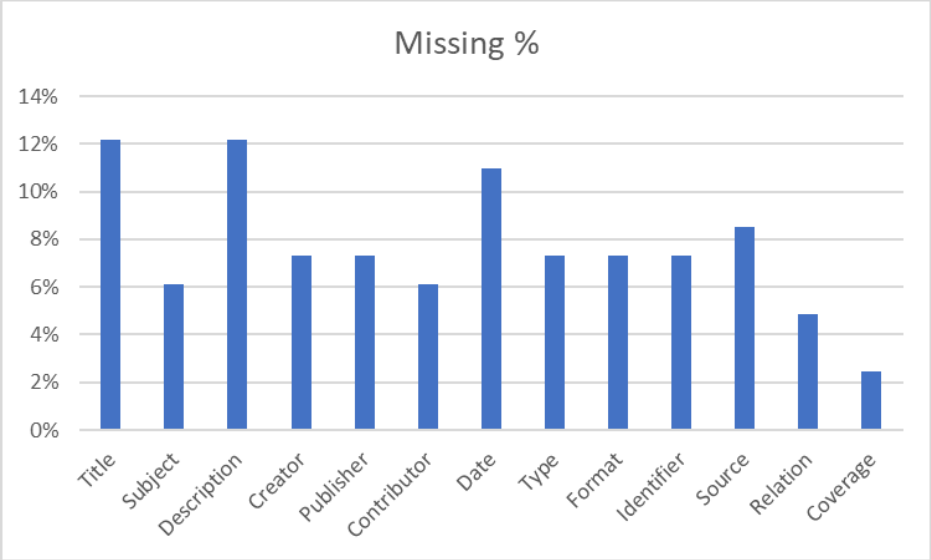


Figure 21 – Examples of distribution of missing metadata among DC metadata fields.

1.5 Step 5: Reporting

The result of this Automatic quality assessment has been tested on a preliminary set of OERs. In the project, after the collection of all the Green, Digital and Entrepreneurial OER that will be considered in ENCORE, it will be tested on a larger set of items. As an extra work, to increase the quality and scientific soundness of the project, we aim to write a document having the following contents:

1. **ROER quality statistics resume**, summarising the results of the automatic evaluation as described in Step 4.
2. **Residuals analysis,** summarising the results of in Step 2 and 3.
3. **Conclusions,** in which is described the output of the automatic assessment, summarising how much OERs have been eliminated, how much OERs have been fixed, and the final number of rows and variables of the ROER.

Appendix.B - List of ROERs

Name	Link
AMSER	https://amser.org/index.php?P=BrowseResources
HippoCampus	https://hippocampus.org/HippoCampus/?user=hippocampus
MERLOT	http://www.merlot.org/merlot/index.htm
NSDL	https://nsdl.oercommons.org/
OER Commons	https://www.oercommons.org
Teaching Commons	https://teachingcommons.us/
The Mason OER Metafinder (MOM)	https://oer.deepwebaccess.com/
OpenLearn	https://www.open.edu/openlearn/
OERu	https://oeru.org/
Saylor Academy	https://www.saylor.org/
OASIS	https://oasis.geneseo.edu/index.php
OERTX	https://oertx.highered.texas.gov/
Open Michigan	https://open.umich.edu/
OpenStax	https://openstax.org/
Open Textbook Library	https://open.umn.edu/opentextbooks/

Bibliography

Aikoh, K., Isoda, Y., & Sugimoto, K. (2020, October). Data Profiling Method for Metadata Management. In 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA) (pp. 779-780). IEEE.

Course Quality Checklist, Udemy (2016).

Digital Learning Course Quality Checklist, Kennesaw State University (2022).

Downes, S. (2001). Learning objects: Resources for distance education worldwide. International Review of Research in Open and Distance Learning (July). Retrieved September 12, 2003 from http://irrodl.org/content/v2.1/downes.html

E-Learning Course Quality Checklist, Art Providers (2020).

Elias, M., Oelen, A0., Tavakoli, M., Kismihok, G., & Auer, S. (2020, September). Quality evaluation of open educational resources. In European Conference on Technology Enhanced Learning (pp. 410-415). Springer, Cham.

Haughey, M., & Muirhead, B. (2005). Evaluating learning objects for schools. E-Journal of Instructional Science and Technology, 8(1), n1.

Hickey, A., Urban, A., & Karsten, E. (2020). Drivers of Quality in Online Learning.

Higgs, P., Meredith, S., & Hand, T. (2003). Technology for sharing: Researching learning objects and digital rights management. Flexible Learning Leader Report 2002.

Király, P., & Büchler, M. (2018, December). Measuring completeness as metadata quality metric in Europeana. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 2711-2720). IEEE.

Kiramang, K. (2008). A COMPARATIVE ANALYSIS ON DUBLIN CORE (DC) AND METADATA OBJECT DESCRIPTION SCHEMA (MODS).

Kurilovas, E., Bireniene, V., & Serikoviene, S. (2011). Methodology for evaluating quality and reusability of learning objects. Electronic Journal of e-Learning, 9(1), pp39-51.

McGreal, R. (2004). Learning objects: A practical definition. *International Journal of Instructional Technology and Distance Learning (IJITDL)*, *9*(1).

Moise, G., Vladoiu, M., & Constantinescu, Z. (2014). MASECO: a multi-agent system for evaluation and classification of OERs and OCW based on quality criteria. In E-Learning Paradigms and Applications (pp. 185-227). Springer, Berlin, Heidelberg.

Ochoa, X., & Duval, E. (2009). Automatic evaluation of metadata quality in digital repositories. International journal on digital libraries, 10(2), 67-91.

Online/Blended Learning Course Quality Checklist - Commonwealth of Learning Quality Checklist & Quality Matters

Palavitsinis, N., Manouselis, N., & Sánchez-Alonso, S. (2013). Metadata quality issues in learning repositories.

Peláez, A. R., Pullaguari, N. P., & Caro, E. T. (2011, April). Quality model proposal for educational material production in OCW sites. In 2011 IEEE Global Engineering Education Conference (EDUCON) (pp. 1074-1080). IEEE.

Questions for Designing and Delivering Online Courses, Dublin City University (2020)

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. Educational Psychology Review, 32(1), 77-112.

Sosteric, M. & Hesemeier, S. (2004). A first step toward a theory of learning objects. In R. McGreal (Ed.), Online education using learning objects. London: Routledge/Falmer.

Tavakoli, M., Elias, M., Kismihok, G., & Auer, S. (2020, July). Quality prediction of open educational resources a metadata-based approach. In 2020 IEEE 20th international conference on advanced learning technologies (ICALT) (pp. 29-31). IEEE.

Ureña-Cámara, M. A., Nogueras-Iso, J., Lacasta, J., & Ariza-López, F. J. (2019). A method for checking the quality of geographic metadata based on ISO 19157. International Journal of Geographical Information Science, 33(1), 1-27.

Weibel, Stuart, L., & Traugott, K. (2000). The Dublin core metadata initiative. D-lib magazine 6.12.

Wiley, D. (2000). Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy. In D.A. Wiley (Ed.), The instructional use of learning objects Online version (Chapter 1). Retrieved September 12, 2003 from http://reusability.org/read/chapters/wiley.doc

Sitography

Centre For Innovation In Teaching And Learning (Ed.). (n.d.). Learning Objects (LOs) versus Open Educational Resources (OERs). https://blog.citl.mun.ca/instructionalresources/learning-object-vs-open-educational-resource-oer/#:~:text=Learning%20Objects%20are%20a%20type,%2C%20re%2Dpurposed%20or%20modified.

DCMI: Creating Metadata. (n.d.). Dublin Core. https://www.dublincore.org/resources/userguide/creating_metadata/#Type

Introduction to Multimedia. (n.d.). https://slideplayer.com/slide/7100961/

Library of Congress. (n.d.). Understanding MARC Bibliographic: Parts 1 to 6. Library of Congress. https://www.loc.gov/marc/umb/um01to06.html#part4

McGreal, R. (n.d.). Learning Objects: A Practical Definition. https://www.itdl.org/Journal/Sep_04/article02.html

Naydenova, G. (2021, April 5). Readability metrics — easy to understand and even easier to apply. https://galiusha.medium.com/readability-metrics-easy-to-understand-and-even-easier-to-apply-f00028fe4953

Networking 2002 Forum What is a Learning Object? (2002, August 21). Flexible Learning | Personal Loan. Retrieved October 17, 2022, from http://www.flexiblelearning.net.au/

QS World University Rankings 2023: Top Global Universities. (2022, June 8). Top Universities. https://www.topuniversities.com/university-rankings/world-university-rankings/2023

Rankings of academic publishers. (n.d.). Wikipedia. https://en.wikipedia.org/wiki/Rankings_of_academic_publishers

UC San Diego. (n.d.). Quality Online Course Checklist. Digital Learning. https://digitallearning.ucsd.edu/instructors/online-course/qm-checklist.html#4.-Instructional-Materials-&-Le

Co-funded by the European Union

Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s)only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them

.