THE ROLE OF METADATA IN THE DESIGN OF EDUCATIONAL ACTIVITIES

Paul Libbrecht, Ulrich Kortenkamp

Didactics of Mathematics, Martin-Luther-University of Halle-Wittenberg

There exist several repositories of learning resources, some of them specific to mathematics, each with its own specificities. In the process of understanding the usages of the metadata in repositories, preparing for the Open Discovery Space federation [1], we attempt to describe the role of the metadata for the potential users of the learning resources: what purposes it serves, and when it is useful or not. We do this based on the log-books of actual maths teachers of the i2geo platform whose experience and professional practice of teaching sets particular utility of the metadata: how they find the resources, assess the resources' qualities, and evaluate the cost of possibly needed adaptations: all depends on the practice.

INTRODUCTION: METADATA IN EACH CONTEXT

Metadata is understood to be the data about data – a fairly generic concept, probably as general as the notion of a resource. Our basic concern in this paper is learning resources: electronic information sets, which can be used by an educator or learner to support their teaching or learning processes. More precisely, we are interested in the information that is encoded in repositories besides the resources itself: the metadata.

Multiple repositories of learning resources exist, including LeMill (lemill.net), i2geo (i2geo.net), Curriki (curriki.org), Merlot (merlot.org), Agrega (agrega.educacion.es), each with its specificity: for example, Curriki allows to search by educational levels for sections of ages, while Merlot does not allow such a search, and Agrega supports search by the exact educational levels (of Spain). Each of these repositories was created with a target population and a target set of learning resources in mind; based on these, the toolset and the metadata structure [2] was chosen. However, the impact of these choices of metadata schema and tool sets on the day-to-day practice of teachers is little explored. In this paper we aim at qualifying the role of the metadata in the usage practice of (math) teachers helped by log-books of actual teachers.

The differences between the learning resources repositories above make it so that it is easier to search for resources for a given educational level using Agrega or Curriki than using Merlot or generic web search engines (such as Google): in the two latter cases, only words can express the query and they do so quite ambiguously when one considers simple examples such as the word quatrième representing two different ages depending on the country one is living in.

This paper aims at inspecting the explorative processes of the selection of learning resources. The usage of learning resources by teachers, as well as the social and professional development that follows of this usage, have been studied in such works as the book of Gueudet, Pepin, and Trouche (2012), but *they assume that the resources are already found* and study their impact on the teaching. In this paper, we study the selection process (searching, choosing, evaluating) preceding the use that seems to be little explored. The utility of the information that is displayed to the educators so as to decide on the re-use the relevance of a learning resource should be measured with a look to the professional life of teachers: we contend that re-using learning resources can help the teachers introduce innovative learning practices and support the usage of software in classroom. This paper does not discuss this hypothesis, but the question: Which information should a sharing environment display for each resource so as to trigger re-use? Differently said: What exchange vocabulary can be used between producers and recipients of learning resources?

Thus, this paper aims particularly at the theme 1.1 of the CERME working group: *design and use of technologies and resources: quality issues*.

Definitions

The *learning resources* we are interested in are any form of a digital artefact that is ready to be used by educators or learners to support the teaching and learning process. Learning resources can be found in published works such as textbooks or their supplementary materials, they can be found in the portfolio of the experience of each teacher, or they can be found in sharing platforms. The concept of “learning resource,” which is equivalent to that of a “learning object” (Wiley, 2000), is used by all current *Open Educational Resources* repositories. These repositories create a context, which allows the learning resources to be found and, later on, to be re-used.

For this paper, we shall call *metadata* of a learning resource in a given environment any information recorded directly about a learning resource that is not included inside the learning resource itself. Thus, metadata includes a description or a caption, annotations indicating the target educational level or an instructional type or a snapshot of the learning resource. Partially standardized formats exist to encode metadata (LOM, DC-ED, LRMI) and may help the exchange between different platforms (container websites). Metadata records are generally split in sections such as: general, authorship, rights, pedagogical metadata, and technical metadata. Beyond the metadata, one often calls *paradata* the set of data about a learning resource that has been recorded following a particular view or usage of the resource. This includes ranking statements, records of how many students have succeeded, or comments on the resource. While paradata is not exactly metadata (and it often lies in separate places than metadata), it may often serve the same role.

Although our investigations have a potential of application beyond teachers of mathematics in Europe, they are our focus: The tools, the resources, and the annotations vocabulary in our study are designed for them and by them.

Outline The paper first informs about the objectives of this research: the design of the Open Discovery Space platform. It then describes the principles the i2geo log-books approach which constitutes our experimental basis. Then it presents the roles of the metadata, computer-wise and didactic-wise. These roles are then instantiated in an interpretation of the log-books. General remarks form the conclusion.

OPEN DISCOVERY SPACE

The research described here is intended as a basis of the platform design process of the Open Discovery Space portal. This portal will be the result of the EU project of the same name, a broad project gathering 52 institutions across Europe and about 20 learning resources repositories in many subjects, including mathematics.

Open Discovery Space will federate multiple existing learning resources repositories already on the web. Among others, i2geo platform (Kortenkamp et al. 2009), Cosmos portal, open-science-repository, organic edu-net or edu-tube-plus. Overall, these repositories provide several hundreds of thousands of learning resources that shall be made available through the project. (Megalou et al., 2012).

This federation will be enriched by a social network, by students' delivery tools which should empower teachers to re-use learning resources including features as far as the analytics services that allow to know if and how learners have used the resources, and by optional extension-servers which support a deep integration into the school infrastructures. Open Discovery Space is a EU project running from April 2012 to March 2015; as of this writing, its design is being articulated.

The platform design process includes the elaboration of a vision of the usage of the portal in the design of educational activities for secondary school. The vision is to be complemented by scenario (or lesson plan) templates, which will support teachers in the application of alternative didactical approaches.

The project gathers technology enhanced learning specialists in the field of science (notably biology and physics), language learning, and mathematics. It aims at serving the complete range of stakeholders involved in secondary school life.

In order to describe concrete scenarios of usages of the Open Discovery Space portal, user stories have been written [3] and the design process of teaching activities using the expected platform is being sketched: this is where the learning resources sharing platform are expected to enter the learning, and thus where metadata becomes important.

To understand this process, reports of the experience with other platforms are gathered. In this paper, the log-books of usages of the i2geo platform are discussed.

THE I2GEO LOG-BOOKS OF THE RESOURCING PROCESS

During the Inter2geo project (which ran between 2007 and 2010, see Kortenkamp, 2009), a team of active teachers attached to the INRP in Lyon (France) decided to gather to discuss and attempt the usage of the i2geo platform and to report about it. This effort was lead by Jana Trgalová and Sophie Soury Lavergne. The objective of this report was to help to guide the elaboration of the platform so as to make it useful for the work of teachers. Log-books were filled tracing the discoveries made and the expectations that were felt. They can be read, in French, at http://i2geo.net/Coll_Group_IREM-INRP-AcademiedeLyon/LogBooks. These log-books are all dated and sometimes represent the i2geo platform in a very preliminary state. Many of the issues have been dealt with in the meantime, be it on the platform level or on the level of resources. Moreover, some of the log-books mention resources which have been changed in the meantime (internally or externally). These log-books, however, should be read with the perspective of informing how these platform usages have an impact relevant to the teaching activities of these teachers.

We shall review several of the log-books below. They all involved a simple *resourcing process*: start with a need for a future teaching occasion, formulate search queries, skim through each of the probable results, identify the useful ones, try each of them, file a quality evaluation, attempt in class, file another quality evaluation.

THE ROLES OF METADATA

In this section, we describe the roles that we propose the metadata can have in the software activities of a learning resources sharing platform. We differentiate the technological functions and the didactical functions.

Technological functions of metadata

We are interested in the following functions that a computer program can perform with metadata within the activities around learning:

DISPLAY: When a person browses a learning resource within a collection of resources, metadata is presented. Parts of the metadata can then be read or seen by humans; this can help to recognise a resource. Good examples of rendered metadata include the title and description, the media-type (typically as an icon), and the educational level. The display can also include paradata.

SEARCH or FILTER: Using several retrieval methods, it is possible to find the resources that match particular metadata values. This includes browsing a taxonomy and clicking the links or entering a text and showing its matches. The information of the metadata is the basic search ingredient. Search engines generally apply multiple levels of matching between queries and matching resources so that the results list appears to be sorted by *relevance*: for example a word found in the title is more important than in the description or learning resource, a didactical function (e.g. reference, handout, demonstration, …) matching a query in the metadata of a resource that only has one such function is more important than such a function in a resource that dozens of such functions.

RECOMMEND: Based on recommendation algorithms, automatic searches can lead to suggestions of learning resources for users. Recommendation is similar to searching, but the search criteria are given by the software and search is usually not initiated by the user, but by the platform itself.

INPUT: A user that contributes a learning resource, and one who updates it, has the possibility to input or modify most of the metadata.

Didactical Functions of Metadata

Such basic functions as above help an educator to perform a number of actions that are useful for his or her teaching preparation and implementation process.

SELECT: Within a *resourcing* process, teachers routinely seek learning resources that could support their teaching. This generally involves cycles of search, preview, trial, and refinements until something applicable for their objectives and conditions is found.

Selection involves an elaborate dialectic activity between the usage of search tools, the observation (and thus evaluation) of the displayed metadata, the available (or missing) resources, the attempts of usage, and the refinements of the search.

PUBLISH: When users feel that a learning resource would be valuable to contribute so that others can take advantage of it, a basic record of metadata is populated with information that the user considers to be useful. Doing this he or she has an idea how to present the source so that it will be displayed adequately and that expected search queries will show it.

ADAPT: Finding the right resource is most commonly an imperfect quest which needs a complementary adaptation process. For example, one needs to adapt wording, the technical conditions of use (e.g. make PDF out of Word, find the exact link, package into a different format, cut irrelevant pieces…). The cost of adaptation is generally compared to the benefit of re-use as discussed in Libbrecht (2011).

ORGANIZE: Course planning and resource publishing often require the resources to be grouped and labelled. This activity allows a collection of content to be presented along a structure that is practical to get an overview (for example a thematic grouping, or a lesson plan).

DEPLOY: When it is time to get to the classroom, a publishing process happens: a print, the creation of a resource in the learning management system, an indication on the blackboard, an assignment... These processes can end in the classroom (for in-class activities) or later. This is generally the time when the learning resource is ready for the students' use (e.g. when an interactive exercise is properly linked so that most learners will be able to just click and start it).

RATE: During the usage of the learning resource, and during the selection process, a constant critical eye is exercised. The output of this critical eye is a judgement of the quality that is published, typically, on the sharing platforms. Various forms of rating exist, from simple star-based judgements to elaborate multi-dimensional questionnaires such as that of i2geo (Trgalová et al., 2011). This creates paradata.

SUGGEST: A more general form of than rating, suggesting is commonly done in social networks of teachers and learners (for example via Facebook or Twitter). It involves transmitting the information about a learning resource from one person to another (or several). The suggestion should invite the recipient to explore the learning resource by formulating characteristics that are relevant for him/her. This can be done via email, for example, exchanging a URL and indicating or summarising the particularly interesting metadata facets.

THE ROLE OF METADATA REPORTED IN LOG-BOOKS

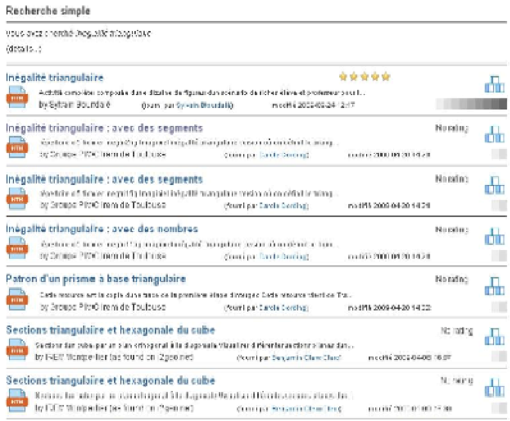
The roles we have described above appear in the INRP-log-books mentioned above. They show which metadata property and function has when led to a decision. We will summarize the analysis for several of these log-books below.

Triangular Inequality: Perfect but...

Log-book: JdB-inegalite-triangulaire.pdf Our teacher searches interactive geometry resources about the triangular inequality (*inégalité triangulaire*).

No concept is registered for this, so he searches for these words and finds a page-full of search results.

Because some resources have these words inside the title, they are listed earlier. The teacher reads the metadata excerpts of the search results: a bit of the description, the ranking … (see the Figure for an example)



Based on the title and the descriptions, the annotated level and topics, and the didactical functions, he can select a resource that seems appropriate: inégalité triangulaire. This teacher has decided on this resource because of its title which matches exactly the expectations.

In this case, it is a linked page, which contains students' and teachers' sheets as well as 9 interactive exercises.

Our teacher can test individual parts of the resource, making sure it is ready to try for the students. Thanks to the teacher sheet, he can plan that this activity will take two course-hours in lab and can book the rooms accordingly. He skips a part so as to save time. The second sheet is the starting sheet.

During the course, he realises that some of the computers are missing a classical requirement (Java, Flash). Moreover, that day had a very low network bandwidth. Both of these technical issues lead to a loss of time of 15 minutes (of 50).

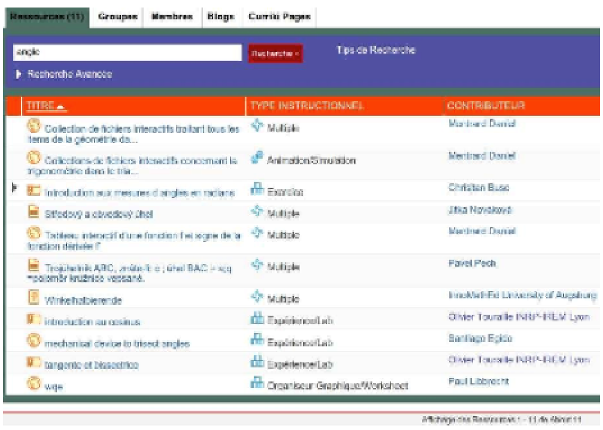
The teacher notes that this resource is perfect, and he rates it highly, but he notes that he would wish his usage to be a bit different and that, since he cannot adjust the resources, he only can tell the students to follow things differently.

Metadata fields used: all that is displayed in the search results and in the resources' info (title, description, levels, instructional function).

Actions: done: select, deploy, rate; wished: adapt, organise.

One of the most important criteria this teacher sets forth and has successfully encountered is the completeness of the didactical details (teacher's sheet, students' sheets, time estimates…). To our knowledge, no metadata structure encodes this completeness.

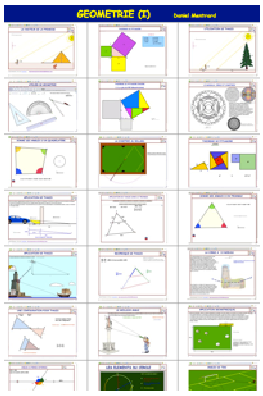
Corresponding Angles: Too Coarse Resources



Log-book: [JdB-angle.pdf]() Our teacher searches for *angle* using the search tool available then: a plain text search tool similar to that found on Curriki.

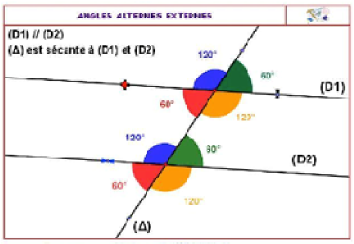
She wishes to find resources that allow students to infer the relationship between parallelism and corresponding angles.

As is usual with plain text search, the result includes multiple unrelated results, only two seem to be closely related. She then tries to search for the occurrences of the plural word *angles*, which gives unrelated results.



For these two, she looks deeper, opens the resource view, and opens the linked URLs. These URLs are large collections of resources, such that inspecting each of them takes much time.

One of the resources she finds in this big collection matches exactly in theme but the pedagogical approach bothers her: She wants the students to discover the relationship between the angle equality and the parallelism themselves but that interactive construction fixes the parallelism at start.



At the end, she creates a new resource [Angles correspondants]() which corresponds to her intents. She expects to use it as a demonstration tool in classroom.

Metadata used: title, description, authors. Missing the use of a more precise topic.

Actions: done: select, deploy, publish, rate; wished: adapt.

Remarks: This log-book is a classical story of *poor metadata*: the topic our teacher expected is a relatively precise topic, a topic that needs two words hence is difficult to search for with precision. As a result, our teacher uses more general terms and has to sort through a pile of irrelevant resources which she can easily identify thanks to titles (such as the set of resources about the trigonometry).

She then meets another widespread issue for the central role of the learning resources platform (to make available learning resources findable to many): granularity. Because the resource that is contributed is a broad collection encompassing multiple objects (see a screenshot on the right), it cannot be finely annotated with topics and educational levels. Moreover, the description of each resource and the resources' content is not indexed. This is the reason why the best-practice guidelines of i2geo proposes alternatives (Mercat et al., 2009).

Finally, one should note that the resourcing process is also made of simple visits: our teacher has created her resource at the end, which is less expensive for her than, for example, requesting the installation of a new software. That creation has been clearly supported by the resources that she has viewed before. A *dark re-use*, as coined by Wiley (2009), has happened.

Exponential: Cross-lingual topic search

Log-book: JdB-fctexp-euler.pdf Our teacher intends to find supporting material for the introduction of the exponential function applying the Euler method.

First she searches for the *exponential function* topic. This search finds the resources that have been annotated with that concept (or a sub-concept). She only finds two resources, only one is interesting but it is in German. Nonetheless, she attempts it.

She tried to understand the presented resource (Exponentialfunktionen) but grasping it enough from the geometric aspects alone was impossible, she gave up.

She then searches in plain text in several attempts. At the beginning, the results' list is big and full of unrelated results (because such words as functions are very common and the words that match fuzzily are also included). Finally, she finds how to search for a “phrase” putting quotes around it (a fixed sequence of word). The resources she finds are, however, insufficient.

She then uses a generic web search engine, Google, and operates the same resourcing process. She finds the tenth result to be appropriate

This resource fits her needs, she is easily able to deploy it to her students and adapt it as needed. She contributes it on i2geo (Introduction...).

Metadata: topic, title, description (displayed, searched, input)

Actions: done: select, adapt, publish, deploy, rate.

It should be noted that the topic annotation is a very precise search ingredient: it allows to search with almost no error but often misses some resources which have not been tagged appropriately. Similarly the query for “fonction exponentielle” (as a phrase) is much more precise and misses results which, for example, do not contain this exact phrase. Such an ambiguity is recurrent and not fully solvable unless a considerable effort is made into polishing the annotations of the resources, for example by supporting is encoding by applying suggestions based on text analysis.

CONCLUSIONS

In this paper, we have proposed a coarse model of computer functions and of didactical functions of metadata. We have applied this model to the log-books reporting early activity of the i2geo platform successfully: the resourcing process described there is entirely dependent on the quality of the metadata records.

These log-books have shown the tricky role of the metadata: When read, it must be expressive enough for resources to be easily identifiable, still it must be easy to input. They also have shown that the criteria to choose a given resource to be applied in a teaching situation include all fields of the metadata that can be searched or displayed; these fields also include the didactical facets of the resource (in particular, the available documentation) and the compatibility to the technical environment.

These log-books have also shown a premise of the metadata that is often forgotten: its goal is to form a catalogue, and this catalogue should be informative. If a search result shows information that does not allow recognising the resources contents, it is likely to require extensive manual skim through all the results. This implies that a person that inputs a good metadata input is one that knows the available content well.

This study has also shown a role of metadata which is completely different than that of enabling the *automatic assembly of learning resources* (as expressed in early visions such as those quoted in Wiley (2001)): the metadata display forms a step in the selection process, where the teachers’ expertise plays an important role.

Finally, these log-books have shown us interconnections between select, publish, adapt, and deploy actions: all teachers' log books demonstrate that previous actions have influenced the next ones, even if they were a selection that lead to a rejection.

NOTES

1. http://opendiscoveryspace.eu.

2. The name *application profile* is generally used to describe a structure of metadata that partially follows and extends a previous metadata structure.

3. Note to the reader: this paper is written a bit more than two weeks before Sep 30th when several important public deliverables of the OpenDiscoverySpace projects will appear. Among others, one will describe the architecture and another will survey in details the metadata and the possible mappings.

ACKNOWLEDGEMENTS

This work has been partially funded by the European Union’s ICT-PSP Programme under the project Open Discovery Space. The opinions expressed in this paper are the sole responsibility of the authors.

The authors would like to thank the team of teachers associated in this entreprise, including: Frédérique Bourgeat, Anne Calpe, Marina Digeon, Esmaël Esfahani, Isabelle Leyraud, Sophie Soury-Lavergne, René Thomas, Olivier Touraille, and Jana Trgalová.

REFERENCES

Gueudet, G., Pepin, B., & Trouche, L. (2012) *From Text to 'Lived' Resources: Mathematics Curriculum Material and Teacher Development*, Springer Verlag, Berlin. http://www.springer.com/book/978-94-007-1965-1.

Kortenkamp U., Blessing A.M., Dohrmann C., Kreis Y., Libbrecht P., Mercat C., (2009) Interoperable Interactive Geometry for Europe – First technological and educational results and future challenges of the Intergeo project, *Proceedings of the Sixth CERME congress*, Lyon (France). http://ife.ens-lyon.fr/editions/editions-electroniques/cerme6/working-group-7.

Libbrecht, P. (2011) Re-Use? Is this Re-Use? *ZDM – The International Journal on Mathematics Education* 43(3), 337-351

Mercat, C., Hasek, R., Leistenschneider, M., et al. (2009) *Report on best practises for the use of DGS content*, Deliverable of the Intergeo Project. http://i2geo.net/xwiki/bin/view/Main/Publication.

Megalou, E., et al. (2012) *Review of Open Learning Content Requirements*, Open Discovery Space deliverable D7.1, September 2012.

Trgalová, J., Soury-Lavergne, S., Jahn, A. P. (2011) Quality assessment process for dynamic geometry resources in Intergeo project: rationale and experiments. *ZDM – The International Journal on Mathematics Education* 43(3), 337-351

Wiley, D. (2001) 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* (pp. 1-35), Association for Educational Communications and Technology. See http://reusability.org/.

Wiley, D. (2009) *Dark matter, dark reuse, and the irrational zeal of a believer*. Retrieved from http://opencontent.org/blog/archives/905.