Cross-Curriculum Search for Intergeo

Paul Libbrecht[1], Cyrille Desmoulins,[2] Christian Mercat,[3]
Colette Laborde,[2] Michael Dietrich,[1] and Maxim Hendriks[4]

[1] DFKI GmbH, Saarbrücken, Germany
[2] LIG, Université Joseph Fourier, Grenoble, France
[3] I3M, Université Montpellier 2, France
[4] Technische Universiteit Eindhoven, The Netherlands

Abstract. Intergeo is a European project dedicated to the sharing of
interactive geometry constructions. This project is setting up an anno-
tation and search web platform which will offer and provide access to
thousands of interactive geometry constructions and resources using them.
The search platform should cross the boundaries of the curriculum stan-
dards of Europe. A topics and competency based approach to retrieval for
interactive geometry with designation of the semantic entities has been
adopted: it requests the contributor of an interactive geometry resource
to input the competencies and topics involved in a construction, and
allows the searcher to find it by the input of competencies and topics
close to them; both rely on plain-text-input.
This paper describes the current prototypes, the input-methods, the
workflows used, and the integration into the Intergeo platform.

1 Introduction

The last decade has seen a bloom in tools that allow teachers to enrich their
teaching with interactive data, whether in face-to-face or distant mode. This
wealth has its drawbacks and teachers need support to navigate through this
diversity: which software should I use, where can I find resources, will this
resource work for my class? Indeed, apart from pioneer work by dedicated
teachers, the actual practices in the classroom have not evolved much. The rea-
sons are manifold. Here are the three main ones:

- All the communities that have grown around the different technical solutions
  and software available have produced resources that they share in one way
  or another. They have all thought about their practice and produced diffe-
  rent approaches. Currently these cannot be merged, because the data they
  produce is scattered, both physically and semantically. The resources need
  to be *centrally visible and exchangeable.*
- As well as being difficult to find and analyze, the resources are usually diverse
  in quality and relevance to a specific need. Teachers are unsure in which
  situation a given resource, even if apparently interesting, could actually be
  used, and whether it adds pedagogical value to the learning experience [1,2].
  They wait for a bolder colleague to report on her attempt. The *resources
  need to be tested*, and *published reports need to reflect these tests.*

– Mastering a piece of software is time-consuming, and very few teachers grow to become power-users of their tool. The resources need to be *easy to use, share and adapt*, in spite of software choices.

In order to solve these issues at least for one specific subject, interactive geometry, we propose to centralize educational resources from this field on the Intergeo web platform. All resources will have clear Intellectual Property Rights, promoting open licences. And they will be there in an interoperable file format we are going to create, based on OpenMath [3]. This format will be supported by the most common software programs for interactive geometry, so teachers can keep on using their own. In this article we will detail the way in which resources are annotated and how our search tool works with the competencies of many curriculum standards.

1.1 Outline

Sub-section 1.2 provides a very short description of what interactive geometry software is. Section 2 is a survey of learning object repositories comparable to what the Intergeo platform should be and presents some of the rationale behind the choice of platform. Section 3 deals with the preliminary phase of the development of the search tool while analyzing the workflow for the search and annotation tasks. Section 4 describes the design process of the ontology used for both representing the various European curricula and the resources we want to see shared among all users. Section 5 describes the two methods of inputting queries, by typing and explicitly selecting competencies and topics or by pointing in a curriculum or a textbook. The search process, from the query to a list of resources ordered by matching scores, is described in section 6. The paper ends with a vision towards the dynamic evolution of the ranking algorithm (section 8).

1.2 What is Interactive Geometry?

The Intergeo project is driven by European leaders in interactive geometry software. We are going to explain what is understood by *interactive* or *dynamic geometry*, a way of doing geometry which is required of math and science teachers more and more often. Interactive geometry allows for the manipulation and the visualization of a construction (a figure) on a computer. The construction depends on some free parameters, like the position of one or several control points. The user manipulates the figure through the keyboard, the mouse or a tracking device, by changing one or more of these free parameters. The construction then changes accordingly.

Of course, the main entities and relations in interactive geometry are of geometrical type. You will find triangles, circles, lines and points, barycentres, tangents, secants with given angles and distances [4]. But it is much more general than antique Greek geometry – you can have functions, derivatives, colors, random variables, all sorts of constructs that allow you to visualize and manipulate concepts that arise in all sorts of contexts, inside mathematics as well as outside [5,6,7].

2 Survey of current repositories

In order to approach the realization of the intergeo platform for sharing interactive geometric constructions across curriculum boundaries, we give a brief survey of the state of learning object repositories which are closest to what the intergeo platform should be.

2.1 Annotations and Retrieval in Learning Object Repositories

As far as we could observe, learning object repositories all classify learning *objects* of a highly variable nature using a certain amount of bibliographic information augmented by pedagogical and topical information. Unfortunately, there is rarely enough information to allow fine-grained search. Topical information is, at most, encoded in broad taxonomies such as the Mathematical Science Classification (MSC)[8]. The most fine-grained taxonomy for mathematics seems to be the WebALT repository [9] which attempts to refine the MSC to a level close to a curriculum standard but seems to stick to a single organization.

Other approaches that tend to be fine-grained are the tag-based approaches, where any tag can be attributed freely by any person providing content. While this approach works fine for statistical similarity and in communities that share a single language, it could only offer translation capabilities if mostly used by multilingual users and users that bridge several communities; we have not found, yet, such users to be common.

A learning object repository that provides topical information directly within the curriculum is GNU Edu [10]: this platform catalogues learning objects according to the skills described in a curriculum, split into years and chapters. GNU Edu allows the skills to be annotated with keywords which can be used to access the skills directly. The keywords are translated and this is how GNU Edu achieves cross-curriculum search: a query matches a set of keywords, each matching skills from each curriculum. GNU Edu does not, however, rank the results or generalize a query so that related keywords also matched.

The emergent repository TELOS from the LORNET research network, and its associated competency framework [11] have been considered, but rejected for their too generic approach. We are not concerned with the design and organization of coherent courses or evaluations; on the contrary, Intergeo resources will be aimed at being used as building blocks by more elaborate Learning Content Management Systems.

Several approaches to link resources to curricula are available. England's Curriculum Online [12], a concerted effort between the Education Board of England and several publishers to present the curriculum standard of England associated with resources that schools may purchase. Microsoft Lesson Connection is a joint of effort of Microsoft and a publisher to do the same for the curricula of the USA [13]. Most of these approaches seem to be based on directly and manually associating resources to lines in curricula, something which is clearly not an avenue for us, since we want the resources to *cross the curriculum barriers*, even being available for a freshly encoded curriculum.

The American commercial project ExploreLearning [14] has a similar view on cross-curriculum and textbook search: They propose interactive figures that are associated to both curricula of the different states of the United States, and to standard textbooks. We don't know whether this association relies on skills and topics ontology or (more probably) is performed manually.

The analysis above leads us to the belief that text search engines, based on information retrieval principles, still tend to be the most used approach for learning object identification. Information retrieval, the science of search engines, is a very mature field with pioneer works such as [15]. Software tools such as Apache Lucene [16] provide a sturdy basis to apply the theories of this field with good performance expectations. Indeed, we shall exploit partial search queries as often as 100 times a second for the purpose of *designating* the topics. Information retrieval is mostly for word matches. It has taught us the fundamental approach to quantify the *relevance* of a document matching a query: this yields search results that are ranked from most to least relevant and expects users to read only the most relevant results.

One way to generalize a query is to make it tolerant to typos or to match phonetically. Another way is to generalize the search by including *semantically close* words. An example is the Compass tool [17], which uses an ontology of all concepts to generalize queries using concepts related to the query words. But even the Compass approach needs to be complemented for cross-curriculum search of interactive geometry, since we wish that a search in French for the topic *théorème de Thalès* should match (at least mildly) a construction contributed by an English speaker who has annotated it with the competency of *recognizing an enlargement.* As a result, the Intergeo project needed an approach that imitates the query-expansion mechanism found in Compass and others but that performs this expansion with the mathematical relationships. Hence we need to tackle the work of encoding the geometric parts of curriculum standards of Europe in a way that identifies the common topics and their relationships. In particular, the search engine that associates topics and competencies to queries will be able to help annotate forthcoming curricula quickly and resources matching its entries will appear instantly.

2.2 Choice of Repository Platform

Learning object repositories can be compared by the services they offer. We shall tackle services which are relevant for interactive geometry constructions, on the authoring side, in order to upload, version, preview, convert, encapsulate into easily edited web-content, deliver and annotate the resources, and on the user's side communicate and report within a chosen community, especially in order to promote enhancement quality cycles through quality evaluation of resources and more casual forums.

In order to obtain all these objectives, we settled on building on the foundation of the Curriki learning object portal [18], an open-source extension of the XWiki platform, which provides textschtml-editing and communication services,

and appeared easy enough to be developed further to accomodate needed extensions such as the search tool.

Having documented the general problematics of applicable technologies, we now turn to a more precise description of the user workflows, and afterwards shall cover the literature relevant to curriculum encoding.

3 User workflow for searching and annotating

The Intergeo platform's main goal is to allow sharing of interactive geometric constructions and related materials. This material can take on the form of interactive geometric constructions, with or without concrete learner tasks attached to them, as well as web-based materials that encompass these. We shall use the term *resource* here, as has been done often on the web, to denote any of these data types. What does the sharing mean? Overall, it is the execution of the following *roles*:

- the *annotator role*: provision of authoring, licensing, topical, and pedagogical information about a resource contributed to the Intergeo platform;
- the *searcher role*: navigation and search through the platform's database to find relevant resources to use in teaching, to edit, or to evaluate.

The roles described here shall be complemented with the *curriculum encoder* role (described below) and the *quality evaluator* role described in section 7.

A crucial condition for the annotator's and searcher's roles to work is that together, they use a similar vocabulary to input the information about the resources and to search for the resources. A fundamental aspect of Intergeo is to solve this in a cross curriculum fashion, so that the annotator and searcher can express themselves in vocabularies that may be in different human languages and in different environments.

A simple example of a matching that crosses curriculum boundaries is the construction of the division of a segment in n equal parts. This should be matched by queries using strings such as "divide in equal parts", "diviser en parties de même longueur", etc. Curriculum standards, however, do not all speak about this topic in the very same way. The English curriculum only mentions the operation of *enlargement*, whereas the French national program

Figure 1: Théorème de Thalès. $AB/BC = A'B'/B'C' = 1.62$

of study mentions "connaître et utiliser dans une situation donnée les deux théorèmes suivants" and provides the formulation of the "Théorème de Thalès" and its converse [19].
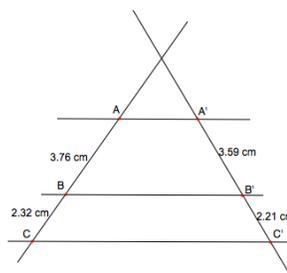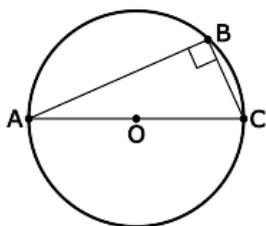
Figure 2: Satz des Thales.

A simple example of a mismatching across some of the curriculum boundaries is the name "Thales' Theorem". In French (*théorème de Thalès*) and Spanish (*teorema de Tales*) it indicates the intercepting lines theorem, concluding proportionalities of segments, as in figure 1. However, Thales' Theorem in English or in German (Satz des Thales) refers to the theorem that if one takes a point on a circle and draws segments to the two endpoints of a diameter of the circle, these segments will be perpendicular, as in figure 2.

Thus the role of the annotator is to provide sufficiently detailed topical and educational context information so that users in other curricula can find resources using the language of their curriculum as well as using everyday language. For this to work, we have added a third role to this workflow, that of curriculum encoder. This person makes sure that every competency and topic in a curriculum standard he or she is responsible for, is encoded in our ontology. The curriculum standards come from different sources, mainly official ones such as ministries of education publications, and often go by a different name. But everyday practices of teachers lead us to consider more practical implementations of curricula: textbooks that teachers ordinarily use in the classroom. We will ask editors to provide at least the table of contents of their textbooks and manuals to annotate them similarly.

4 GeoSkills: a cross-curriculum ontology for geometric constructions

4.1 Ontologies and sub-ontologies

The Intergeo project defines an ontology called GeoSkills [20], consisting of several sub-ontologies (to which we also simply refer as ontologies in themselves). These contain classes for *competencies*, *topics* and *educational level*, respectively. The first two reflect an agreement of the community as to what is actually being taught. The third one will be discussed a little more in the next section. The ontology mostly describes mathematics learned at the secondary school level, but could of course be extended to cover much more. Let's make clear what we mean by competencies, topics and educational levels:

- a topic is an object of knowledge such as *isosceles triangle* or *Thales theorem*;
- a competency is the compound of an ability (a verb) and a topic such as *identify parallel lines*;
- an educational level is a stage in the development of a learner, in the context of a specific educational region and educational pathway (school type). For example "Eerste klas" of the pathway "secundair onderwijs" in "Vlaanderen", an educational region within Belgium.

The competency ontology makes it possible to represent that the competency "use of scale" taken from the English national program of study [21] is related to "intercept theorem", itself linked to triangle, enlargement, similarity of triangles, measuring segments and measuring angles. And it enables us to capture the fact that the resource depicted in figure 2 refers to this competency.

One thing the ontology aims at is providing European curriculum experts the means to encode localized geometry curricula with a common semantics. Another is to enable searching, which will be discussed in the next section. The goal of this section is to describe the approach and the design decisions made in order to provide the ontology for curriculum encoding by experts. It presents the tool we used to create the ontology collaboratively. The methodology we followed is to rely on mature and widespread tools and practices both at the theoretical and practical level. On the theoretical level, the approach is to rely on well-defined semantics, decidable knowledge representation and widely interoperable languages. On the practical level, the idea is to use tools providing enough affordance for non computer scientists like curriculum experts from several countries, and to ask them to collaboratively construct the ontology and benchmark it with instances.

4.2 Other projects

To design the competency ontology, we first surveyed tools providing curriculum mappings, encodings, and cross-curriculum search, especially in Europe.

Dragan Gasevic and Marek Hatala [22] developed a curriculum mapping using SKOS between the ACM (Association for Computing Machinery) classification [23] and the IEEE Information Management Course Curriculum Recommendation [24].

The CALIBRATE EU project has been working on curriculum encoding of competencies for curriculum-based resource browsing [25]. Their ontology is composed of an Action Verb Taxonomy and a Topic Taxonomy. A text fragment is tagged with a specific competency described with an Action Verb and a set of Topics taken from the ontology, together called a Tuple. The Tuple approach is well-suited for curriculum indexing. They also developed TopicMapper, a tool enabling curriculum expert to encode curriculum texts in HTML format into this ontology. It is a tool based on the XTM language providing an easy-to-use Graphical User Interface. However, Topic Mapper is a standalone application. It does not work on the web and uses local files.

We chose to use a similar approach to CALIBRATE's one (a verb plus a set of topics) to design our competency ontology.

4.3 Editing GeoSkills with Protégé OWL

The Protégé tool [26] has been chosen, both to design and edit the curriculum ontology and to provide an ontology-based curriculum encoding facility for the national experts. It corresponds to our need of a widespread tool. Protégé is the most widely used ontology editor at this moment. It also provides a simple

Graphical User Interface for designing the concepts and properties of the ontology and for encoding curriculum competencies as instances of this ontology.

Protégé offers two major ontological representations (and a corresponding interface): frame based language [27] or OWL. We chose OWL, because it is an interoperable format provided by the W3C [28], and because it has a well-defined semantics. OWL-DL has been proven to be decidable, which was therefore our final choice. Additionally, several inference engines are available that could help searching [29,30,31]. This contrasts with the previously mentioned topic maps. There exists a standardised language [32] for them and an editing tool [33]. But this editor is less widely used than Protégé and, more importantly, there are no results about the decidability of algorithms on topic maps.

In order to collaboratively design and populate the ontology with our *sample* curriculum, we first tried to use the WEB versioning system (LibreSource so6 synchronizer [34]). After some files were corrupted, we switched to the use of Protégé Server. It saves changes of each concurrent co-author in real time, thanks to JAVA RMI, thus providing a truly collaborative tool. The limits we encountered are twofold. Firstly, as RMI saves each change on the server, the network bandwidth is critical and sometimes not enough. Secondly, Protégé's user interface is not editable in the server version, despite it really being important to let curriculum experts work. It was solved by stopping the server when performing changes to the user interface, which was quite rarely.

At this moment, a first version of the curriculum and resource ontology have been designed and the parts of curricula around the intercepting lines theorem have been encoded for four countries (Great Britain, France, Germany, Spain) without major difficulties.

4.4 Design of the ontology

The competency ontology contains two main concepts. The first is the class **Competency**, a hierarchy of action verbs divided in two main subclasses : **TransversalCompetency** (such as **Apply**, **Calculate**, **Explore**) and **GeometricCompetency** (such as **Construct**, **Infer**, **ToMeasure**). An instance of one of the Competency class is described with:

- a set of instances of the topic ontology (at least one);
- a set of curricula it belongs to;
- names (strings) that can be common names, uncommon names, rare names or false friends.

The various **names** properties provide an easy way to qualify the type of names related to a competency and consequently serve as a basis to implement a fuzzy search among names (a common name is more probable to be matched than any other type). They also provide a way to manage localized names in a simple way, as Protégé OWL provides a user-interface requesting the language of each string.
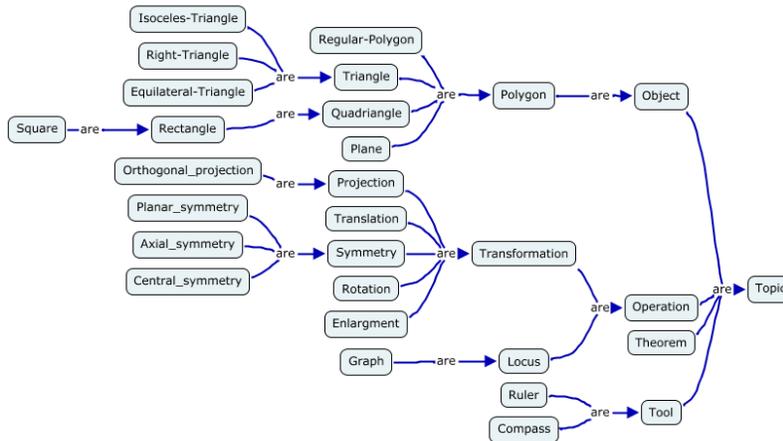
Figure 3: An extract of the **Topics** branch of the GeoSkills ontology

The second main class is **Topic**, a hierarchy of geometry topics mainly divided into the following sub-classes : **Object**, **Operation**, **Proof**, **Theorem** and **Tool**. Part of the topic ontology is shown in figure 3.

Educational levels are encoded following [35]. Similar to competencies and topics, levels are named. The levels branch of the ontology encodes the pathways of learning in the countries of the EU, and the particular country and age that are associated with a certain context. The latter two pieces of information then allow the system to find resources of the appropriate educational level with ordering and distance provided by the pathway and age range.

To test the ontology with concrete resources, a second simple ontology has been developed representing them. It contains a three class hierarchy: Resource, LearningSubject and SchoolLevel. A resource is described by a title and an URI. It is linked to school levels, learning subjects, and competencies (taken from the competency ontology). This represents the objects which shall be indexed well.

5   Using the ontology in searching

Search engines are a crucial part of everyday internet usage, they are the applications that power information retrieval (see [15]). Both the comprehensive nature of the major search engines on the web and their simple query mechanism are extremely attractive. This simplicity is created on the one hand by simple text input and on the other hand by the responsiveness of results. These stimulate numerous search attempts and refinements to attain the right set of documents.

But because search engines are generally text-based, they are improper to search for conceptual entities such as described in the previous section, which can be made of several (overlapping) words. Therefore we designed two means to let the users easily designate nodes of the ontology.
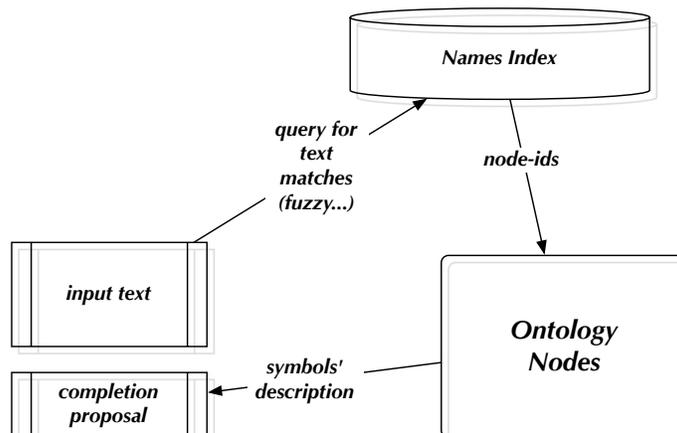
Figure 5: Token designation using SkillsTextBox

5.1 Designating by typing: SkillsTextBox

To let users designate a node of the ontology, we extend the familiar auto-completion: they can type a few words in the search field, these are matched to the terms of the names of the individual nodes; the auto-completion pop-up presents, as the user types, a list of matching nodes similar to figure 4. This list presents, for each candidate ontology node, the full name of the node, the number of related resources, an icon of the type, and a link to browse the ontology around that node. When chosen using either a click, or a few presses of the down key followed by the return key, the sequence of words is replaced by the name of the node, surrounded by square brackets to indicate an exact reference to a conceptual entity in our ontology.

This process is used not only to search but also when annotating a resource: individual competencies, topics, and educational usage are then provided.

SkillsTextBox uses a simple HTML form equipped with a GWT script [36]. SkillsTextBox also uses the Rocket GWT library. This script submits the fragments typed to the index on the server which uses all the retrieval matching capabili-

Figure 4: Choosing among competencies about "Thales conf"

ties (stemming, fuzziness through edit distance or phonetic matching) to provide an object description of the best matching 20 nodes of the ontology, which the script renders as an auto-completion list. This process is depicted in figure 5. More information about it is at http://www.activemath.org/projects/SkillsTextBox/.
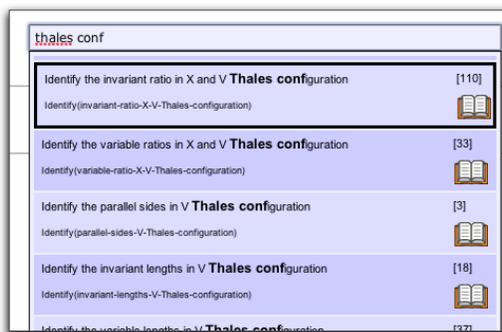
5.2   Which Names to Match?

For SkillsTextBox to come up with the right resources, it is also vital that it knows the educational context in which a query is submitted. For one, it is a basic necessity that the system works transparently for the user: when typing a query, the user should be able to use his or her own language.

The resulting suggestions for competencies and topics should also be in this language, and the returned resources probably suitable for students native to that language should be higher ranked. So letting the person who inputs a query select their language is a first step (done once in the user's preferences). This measure will solve the ambiguity of the name "Thales" either in German or in French. But the problem runs deeper, for there is a bigger cultural context, that of the educational pathway in which the user is working. Not regarding this aspect can create false friends when searching.

One example of a false friend would be when a French-speaking teacher searches for a resource suitable for grade *5ème*. The intergeo system could match the educational programme where the typical age 10–11 of Western Switzerland or that of France where the typical age 12–13. The context of use can exploited to disambiguate.

5.3   Designating by Pointing in a Book

Supplementary to letting users search for resources by explicitly selecting competencies and topics, we will offer the possibility to do this implicitly by letting the system infer these automatically from specific sections in curriculum standards or in text books they know well and that include geometry. Although we shall mostly not be able to offer whole text books to browse through, we expect it to be unproblematic to display tables of contents.

A user can then browse through a table of content and click on sections of interest. This click will trigger the selection of the competencies and topics associated to these sections, adding the necessary queries in the search field.

6   The search tool at work

We have described, in the previous section how a set of words is used to identify interactively a node in the ontology, such as a competency or an educational level. In this section we turn to the actual search, from a query as simple-text to a list of documents, ordered by matching score.

Once a query is launched by the user, it is decomposed into a boolean combination of search terms. Fragments of texts between square brackets indicate queries to individual node names in the ontology whereas isolated words indicate the generic word-query. Consider the example `[Identify parallel lines] [Enlargement] keystage 3`, which includes a reference to a competency `[Identify...]`, a topic `[Enlargement]` and two words `keystage` and `3` that do not yet designate an educational level.

First, the plain words appearing in the query are matched with names of the nodes of the ontology. The query is then expanded to include queries for the competencies that include these words, with a low boosting. Next, the query is transformed as follows:

– query competencies and topics pointed to with high boost;
– for each competency queried, expand with a query for a competency with the parent competency-verb with lower-boost;
– for each competency and topics queried, add weakly boosted queries for the ingredient nodes.

This last expansion step is where a query for the word *Thalès*, not identified with a topic by the auto-completion mechanism, is expanded to a query for Thales-theorem (with high boost) or for enlargement (with lower boost), or for the task of dividing a segment in equal parts or for parallel lines (both with low boost). The isolated words appearing in this query are also matched outside of the ontology (in order of preference):

– in the title of the construction;
– in the author names;
– in the names (along their varying degree of commonality) of the nodes associated to each resource;
– in the plain-text content of the resources.

The expanded query for the example above might look like:

$$\text{comp:identify\_parallel\_lines} \cdot 100 + \text{top:Identify} \cdot 30 + \text{top:Parallel\_r} \cdot 30$$
$$+ \text{top:Enlargement\_r} \cdot 100 + \text{top:Amplification\_r} \cdot 100$$
$$+ \text{top:Reduction\_r} \cdot 100 + \text{top:Operation\_r} \cdot 100 + \text{txt-en:keystage} \cdot 20$$
$$+ \text{txt:keystage} \cdot 5 + \text{txt-en:3} \cdot 20 + \text{txt:3} \cdot 20 + \text{lvl:keystage\_3} \cdot 5$$

This expanded query has now taken full advantage of the ontology, it is passed to the resources' index, also a Lucene index, which returns the first few matching documents with the highest overall score. A presentation similar to the prototype of figure 6 is being implemented.

The combination described yields a search engine with the following characteristics:

– Most importantly, the nodes of the GeoSkills ontology, encoded in the query and in the annotation, are matched against each other. The queries are generalized using the relationships in the ontology. This is multilingual and multicultural, e.g., through the use of topics and competencies, but is expressed using a language-dependent and culture-dependent vocabulary.
– Less importantly, the query words are matched to the resources information and contents. This match is mostly single-language: e.g. queries in english search contents in english and maybe contents in another language
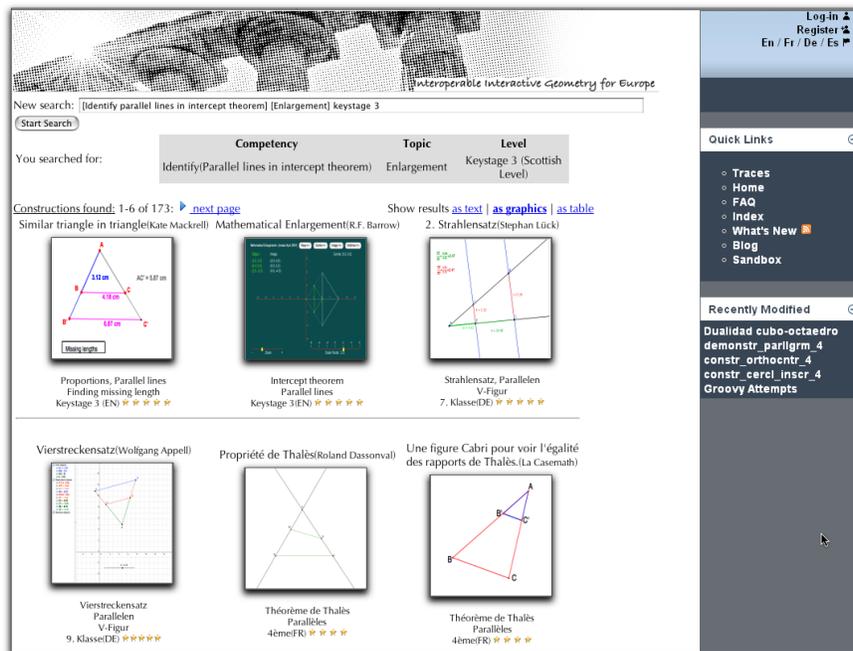
Figure 6: Prototype of the result of searching for `[Identify parallel lines in intercept theorem] [Enlargement] keystage 3`

7 Enhancing quality of retrieval and resources

In this section we will present our vision towards a dynamic evolution of the ranking algorithm based on social network behavior and quality evaluation by peers.

First, the quality asserted by peer review is going to play an important role [37]: the eLearning objects that we gather on our website will be used in the classroom, and we will organize and collect data about this use. We will collect automatic server data like the number of downloads of the resource but also user reports on the usage: on informal forums and chats where user's opinions will be expressed in their own words, but as well through a short questionnaire regarding the adequacy of the resource for the advertised purpose. This questionnaire will be available online to every identified user, for a priori quality assessment for a teacher enrolling for a teaching experience, and a posteriori quality assessment after the experimentation in the classroom has taken place.

Therefore a resource with many positive users feedback will be ranked before a resource with fewer or absent or negative feedback. To achieve this, we shall enrich the index, during nightly updates, with the results of the quality statements, and queries will be expanded to take it in account. A forum will be attached to resources to promote the evolution of the resource that should not always be used "as is" but should be periodically revised, adapted and improved. This quality

assessment as well as the sense of community (see below) will promote responsibility and we expect that authors will syndicate around subjects to organize the evolution and production of quality educational content.

Second, the pedagogical and personal context has to be taken into account, with items such as the country and the language in which the teaching is going to happen, the circle of friends a user has or pioneer experts she tries to emulate; we might want that a resource deemed relevant with regard to the internationalized query to be ranked before another one, despite its lower query-matching score, because it was previously used and appreciated by fellow teachers belonging in the same real or virtual community. This shall be realized by enriching quality results in the index with contextual informations (per educational context and per named-user), the query expansion will then favour resources validated in a similar educational level or by a user tagged as a *friend* [38].

The paradigm behind this is philosophical as well as practical: *the users know better*. Therefore the social interactions and the actual use of the resources should dictate the distances and the scoring, not the reverse.

8   Conclusion

8.1   Implementation Status

The GeoSkills ontology is reaching completeness in structure, it can be seen at `http://i2geo.net/ontologies/dev/GeoSkills.owl` and a rendering can be seen at `http://i2geo.net/ontologies/dev/index.html`. The ontology is being completed for the most learning pathways of Germany, France, UK, The Netherlands and Spain, before October 2008. Curriculum-encoding will then be done by contributing curriculum-experts during the remaining two years of the intergeo project.

The SkillsTextBox GWT project can be enjoyed and downloaded from its project page `http://ls.activemath.org/projects/SkillsTextBox`. It is made available under the Apache Public License. The Search Tool is under active development and will be made available to the public in the summer.

The intergeo platform is under a first harvesting phase where interested parties report about intent to contribute interactive geometry constructions, with license. Since its launch several hundreds of reports have been submitted. The platform is accessible on `http://i2geo.net/`. The second phase will be activated at the end of Spring, based on Curriki, the annotations system will be incorporated in August. Finally, the quality framework will be embedded in the platform and in the search engine at the end of the year.

8.2   Summary

In this paper we have presented an approach to cross-curriculum search relying on a multinational and domain-aware ontology. The ontology basis is the major ingredient for both helping annotating and searching the constructions, of which

we expect to receive several thousands. It will be the result of the coordinated work of curriculum encoders, which we expect to be done by curriculum experts in their community aware of the language of others cultures. This ontology is the key to enable the multinationality of the seach and annotation process tool.

8.3 Acknowledgements

References

1. Guin, D., Trouche, L.: Intégration des tice: concevoir, expérimenter et mutualiser des ressources pédagogiques. Repères (55) (2004) 81–100
2. Ruthven, K., Hennessy, S., Brindley, S.: Teacher representations of the successful use of computer based and resources in secondary school english, mathematics and science. Teaching and Teacher Education **20**(3) (2004) 259–275
3. Buswell, S., Caprotti, O., Carlisle, D., Dewar, M., Gaëtano, M., Kohlhase, M.: The openmath standard, version 2.0. Technical report, The OpenMath Society (2004) Available at `http://www.openmath.org/`.
4. Philippe, J.: Exploiter les logiciels de géométrie dynamique. 4 constructions géométriques avec Géoplan. Les Dossiers de l'ingénierie éducative (54) (2006) 35–37
5. Ait Ouassarah, A.: Cabri-géomètre et systèmes dynamiques. Bulletin de l'APMEP (433) (2001) 223–232
6. Falcade, R., Laborde, C., Mariotti., A.: Approaching functions : Cabri tools as instruments of semiotic mediation. Educational Studies in Mathematics (66.3) (2007) 317–333
7. Hohenwarter, M.: GeoGebra: Dynamische Geometrie, Algebra und Analysis für die Schule. Computeralgebra-Rundbrief (35) (2004) 16–20
8. American Mathematical Society: Mathematical Subject Classfication (2000) `http://www.ams.org/msc/`.
9. Karhima, J., Nurmonen, J., Pauna, M.: WebALT Metadata = LOM + CCD. In: Proceedings of the WebALT 2006 Conference, The WebALT project (2006)
10. OFSET: GNU Edu (2008) `http://gnuedu.ofset.org/`.
11. Paquette, G.: An ontology and a software framework for competency modeling and management. Educational Technology and Society (10(3)) (2007) 1–21
12. British Educational Communication and Technology Agency: Curriculum online (2008) `http://www.curriculumonline.gov.uk/`.
13. Microsoft: Microsoft Lesson Connection Launched At Technology + Learning Conference (1999) `http://www.microsoft.com/presspass/press/1999/nov99/lessonpr.mspx`.
14. ExploreLearning: Correlation of gizmos by state and textbooks (2005) `http://www.explorelearning.com`.
15. van Rijsbergen, C.: Information Retrieval. Butterworths (1979) `http://www.dcs.gla.ac.uk/~iain/keith/`.

16. Hatcher, E., Gosnopedic, O.: Lucene in Action. Manning (2004)
17. Graupmann, J., Biwer, M., Zimmer, C., Zimmer, P., Bender, M., Theobald, M., Weikum, G.: COMPASS: A Concept-based Web Search Engine for HTML, XML, and Deep Web Data. In: Proceedings of the Thirtieth International Conference on Very Large Data Bases. (2004) 1313–1316 `http://citeseer.ist.psu.edu/graupmann04compass.html`.
18. The Global Education Learning Community: Curriki (2008) `http://www.curriki.org/`.
19. Ministère de l'Éducation Nationale: Programmes des classes de troisieme des colleges. Bulletin Officiel de l'Education Nationale (10) (1998) 108
20. The Intergeo Consortium: D2.1 internationalized ontology (2008) `http://www.inter2geo.eu/en/deliverables`.
21. Qualifications and Curriculum Authority: National Curriculum Mathematics Keystage 3. Technical report, Qualifications and Curriculum Authority (2007) `http://curriculum.qca.org.uk/subjects/mathematics/keystage3/`.
22. Gasevic, D., Hatala, M.: Ontology mappings to improve learning resource search. British Journal of Educational Technology **37**(3) (2006) 375–389
23. Association of Computer Machinery: ACM Classification (2008) `http://www.acm.org/class/1998/`.
24. IEEE: ACM/IEEE Computer Society Computing Curricula (2008) `http://www.computer.org/education/cc2001/final/`.
25. Van Asche, F.: Linking learning resources to curricula by using competencies. In: First International Workshop on Learning Object Discovery and Exchange, Crete (2007)
26. Stanford Medical Informatics: Protégé ontology editor version 3.3.1 (2007) `http://protege.standford.edu`.
27. Kifer, M., Lausen, G., Wu, J.: Logical foundations of object-oriented and frame-based languages. Journal of the ACM **42**(4) (1995) 741–843
28. McGuinness, D.L., van Harmelen, F.: OWL Web Ontology Language Overview (2004) `http://www.w3.org/TR/owl-features/`.
29. Racer Systems GmbH and Co. KG: Racerpro owl reasoner version 1.9.2 (2007) `http://www.racer-systems.com/`.
30. HP Labs: Jena - a semantic web framework for java version 2.5.5 (2007) `http://jena.sourceforge.net/`.
31. Parsia, C..: Pellet owl dl reasoner version 1.5.1 (2007) `http://pellet.owldl.com/`.
32. Pepper, S., Moore, G.: Xml topic maps (xtm) 1.0 – topicmaps.org specification (2001) `http://www.topicmaps.org/xtm/1.0/`.
33. Dicheva, D.: Towards reusable and shareable courseware: Topic maps-based digital libraries (2005) `http://compsci.wssu.edu/iis/nsdl/`.
34. Skaf-Molli, H., Molli, P., Marjanovic, O., Godart, C.: Libresource : Web based platform for supporting collaborative activities. In: 2nd IEEE Conference on Information and Communication Technologies: from Theory to Applications. (2006)
35. Eurydice European Unit: Key data on education in Europe 2005. Eurydice European Unit (2005) `http://www.eurydice.org/portal/page/portal/Eurydice/showPresentation?pubid=052EN`.
36. Google Inc.: Google Web Toolkit (GWT), a java to javascript compiler and toolkit (2008) `http://code.google.com/webtoolkit/`.
37. The Intergeo Consortium: D6.1 quality assessment plan (2008) `http://www.inter2geo.eu/en/deliverables`.
38. Xerox Research: Knowledge pump (2006) `http://www.xrce.xerox.com/competencies/past-projects/km/kp/home.html`.