

## CHAPTER 8

# A Cross-Curriculum Representation for Handling and Searching Dynamic Geometry Competencies

Paul LIBBRECHT<sup>a,1</sup> and Cyrille DESMOULINS<sup>b</sup>

<sup>a</sup>*DFKI Saarbrücken, Germany*

<sup>b</sup>*Laboratoire d'Informatique de Grenoble, France*

**Abstract.** Interactive Geometry is becoming part of the curriculum in many European countries; sharing the files of interactive geometry, the constructions, is, however difficult because of the communities are scattered between the many software systems and the many curriculum differences. The Intergeo project addresses this issue by offering a platform where cross-curriculum search and annotation can be done. The annotation language is an ontology and is made easily accessible to users; this ontology describes elementary competencies and topics and their relationships. The search functions, the management, and the access are all empowered by the semantic nature of this ontology together with the various names attached to each ontology element. This paper describes the ontology and the infrastructure that provides utility, usability, and interoperability to this knowledge corpus.

**Keywords.** Competencies, topics, ontologies, learning resources, information retrieval, cross-curriculum search, internationalisation, multilinguality, authoring.

### Introduction

Interactive geometry resources are in wide use in many educational institutions to teach mathematics. Their adoption, however, is often difficult as is often the case with information technologies at school. More convergence is required; the Intergeo project intends to approach it through three different aspects:

- Define a common file format enabling the interactive geometry constructions to cross the software borders, which, currently, often prevent neighbours to reuse each other's resources.
- Create a web-based platform where learning resources with interactive geometry constructions are visible, exchangeable, and searchable: this should cross the borders of national curriculum.
- Allow the users of the platform to annotate the resources with quality statements so that interactive geometry resources are validated, in particular, for their appropriateness in a particular educational context.

---

<sup>1</sup> Corresponding Author: Paul Libbrecht, DFKI, Stuhlsatzenhaus 3, D-66123 Saarbrücken,, Germany; Email: paul@activemath.org.

This chapter is focused on the Intergeo web-based platform where teachers can share interactive geometry resources. The sharing mechanism is based on competencies taken from European curricula. The hypothesis is that teachers and educational experts refer to competencies in their everyday work and that competencies are the best way to link their needs, both pedagogical and mathematical, and corresponding resources.

The purpose of this chapter is the cross-curriculum representation of competencies in the Intergeo web platform. With an approach mixing, on the one hand, the linguistic and mathematical semantic commitments and, on the other hand, techniques from classical database management, semantic web and information retrieval, handling and searching competencies is made at once useful, usable, scalable and interoperable.

This paper starts with motivating examples and follows with an explanation of the GeoSkills ontology that was developed to represent competencies and solve limitations of existing approaches. The Comped editor which makes it possible to edit the competencies and topics of this ontology is then described followed by the usage of competencies in management and search. We conclude with perspectives on users manipulation of GeoSkills and extension to other domains.

## **1. Motivating Examples and Related Works**

To motivate our approach, let us compare it to two other approaches upon an example. Consider the competency of constructing the division of a segment in equal parts. It is described in the French national program of study [1] as “utiliser le théorème de Thalès ou sa réciproque” whereas the English curriculum only mentions the operation of “Enlargement”. In English, the French “Théorème de Thalès” is called “Intercept theorem” (in Spanish, “Teorema de Tales”, in German “Vierstreckensatz”, in Dutch “Stelling van Thales”). However, Thales' Theorem in English or in German (“Satz des Thales”) refers to another theorem (a right triangle inscribed into a circle).

### *1.1. Keyword Approach*

When searching “Thales” competencies across European curricula with a keyword approach, for example in the GNU-EDU system [2], mismatching competencies will be retrieved, some referring to Intercept theorem and the others to Thales’s one, without means to distinguish them.

Handling competencies is also hard with this approach. There is no way to find similar competencies, except if they refer together to the same keyword, with possible mismatches. Even in the same language or country, the link between competencies is loose and browsing competencies is only possible through keyword input.

### *1.2. Thesaurus Approach*

Some existing systems propose, in addition to a keyword approach, to perform an “advanced search” based on age/level and subject classification.

Searching through “Thales” competencies on such systems (Department for Children 2009) is simply not possible as the concept of “Thales theorem” and competencies to use it are too fine-grained for these thesaurus approaches. Classification is finest at best with “Elementary Geometry”, at worst with “Mathematics”.

The competencies management in such system is very poor; lots of competencies are merged into the same generic category.

### 3. *The Intergeo Approach Mixing Linguistic and Mathematical Semantics*

With the Intergeo approach, searching through “Thales” competencies across European curricula by keyword is also possible. It provides also both types of competencies but the mathematical semantics enable to distinguish those referring to “Intercept theorem” from those referring to “Thales theorem”. Additionally, names provided in the user’s own language for competencies support the user to make the difference between them.

The search is also extended relying on inferred knowledge, thanks to an ontology of competencies and topics. For example, consider the competency of being able to *use binomial identities to solve equations* (“utiliser les identités remarquables pour résoudre une équation”).<sup>2</sup> This should of course be matched by queries using strings such as “identity”, “equation” “résoudre”, etc. But it should also be matched by queries using strings as “equality”, because *mathematically* an equality is a kind of identity. “Utiliser les identités remarquables pour résoudre une équation” matches with “use equality” because the ontology defines an equality as an identity and the platform can then infer that this competency matches with equality.

In the Intergeo approach, search can also be performed through browsing the competency hierarchy, which comes from declared or inferred knowledge, or through browsing the content of an official curriculum or a textbook content table.

Mixing linguistic and mathematical semantics provides a great flexibility and accuracy in the management of competencies. Finding similar competency is easy, using semantic relations between competencies categories. New competency can be created without a precise mathematical classification, which is performed automatically afterwards.

## 2. Representing Competencies: the GeoSkills Ontology

Representing competencies within Intergeo platform should meet two major semantic commitments taken from the previous examples: a fine-grained mathematical semantics and competency names taken from various contexts (educational regions and languages).

This sections starts with a survey of competencies representations on learning object repositories. It then presents the GeoSkills rationale and details before explaining the means offered for managing and populating the ontology.

### 2.1. *Related Works*

In order to approach the representation of competencies for the Intergeo platform, we survey the state of learning object repositories, which are closest to what the Intergeo platform should be.

Topical information forms an important part of curriculum description, the other parts being the competencies themselves, in the strict sense of the word, that is the ability to perform actions concerning a given topic. When expressed in a curriculum, a

---

<sup>2</sup> Throughout this paper we provide hyperlinks to the CompEd user-readable representation of the GeoSkills node when they are referenced.

given topic implicitly means “mastering this topic”. For this reason, in the following of this text, we will generalise the competency notion to both competencies and topics, and as well as in the Geoskill ontology.

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 some 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) by the American Mathematical Society [3]. The most fine-grained is the WebALT repository [4], which attempts to refine the MSC to a level close to a curriculum standard.

Other approaches that tend to be fine-grained are the tag-based approaches, where any person providing content can freely attribute any sequence of words (see 1.1) as annotations. While this approach works fine for statistical similarity and in communities that share a language, it does not work so well to provide similarity measures of concepts in a well managed fashion: 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 [2]: 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 generalise a query so that related keywords also matched.

The emergent repository TELOS from the LORNET research network, and its associated competency framework [5] is representing competencies. It has been considered, but rejected because of its main focus on the design and organisation 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. Curriculum Online by the British Educational Communication and Technology Agency [6] is 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 effort of Microsoft [7] and a publisher. The ExploreLearning enterprise do the same for the curricula of the USA [8]. With a smaller scope, Key Curriculum Press indexes Sketchpad dynamic geometry resources with curriculum elements [9]. 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, ready to welcome new curriculum standards as they emerge. Since the start of the Intergeo project in October 2007, we have seen the rise of at least one such repository and the halt of two.

The CALIBRATE project as explained in [10] and further detailed in private communications has appeared to be a first-class provider of annotated curriculum. Unfortunately, their intent did not seem to converge with a cross-curriculum search and their coverage intent appeared to be weak.

The arguments of [10], noticing the rise of the competencies as essential curriculum ingredient and formalising competencies by a process and topics is what we follow in this work, bringing it to the ontological world.

## 2. Rationale of GeoSkills

In order to get a precise mathematical semantics, the approach is to rely on well-defined semantics, decidable knowledge representation, and widely interoperable languages. OWL-DL meets such requirements. It is an interoperable format provided by the W3C in [11]. Its well-defined logic is the Descriptive Logic that has been proved to be decidable. Widely used OWL editors such as Protégé by [12] or Swoop described by [13] and [14]. Additionally, several inference engines such as that of [15], [16], or [17] are available.

This contrasts with the topic-maps standard. There exists a standardised language [18] for them and an editing tool [19] but this editor is less widely used than Protégé and, more importantly, there are no results about the decidability of algorithms on topic-maps.

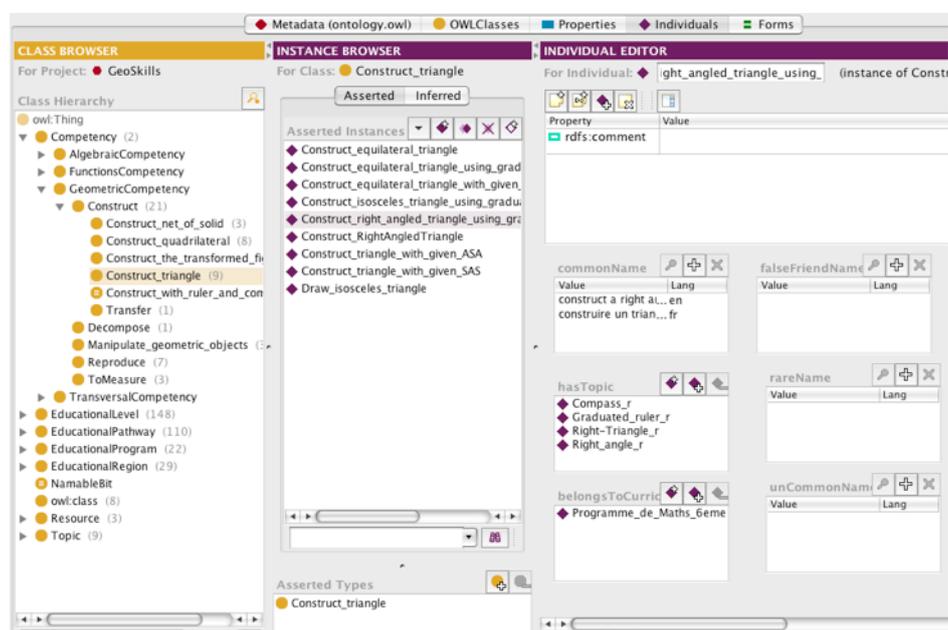


Figure 1. Protégé editing.

On the linguistic level, in order to enable users to identify ontology elements with names and descriptions they are used to, each elements of the OWL ontology (classes, instances and properties) can be described by names for each language. This is made using for instances, dedicated datatype properties, and for classes and properties, `rdfls:label` annotations.

On the practical level, the idea is to use tools providing enough affordance for non computer scientists such as curriculum experts from several countries, and to ask them to collaboratively construct the ontology and benchmark it with instances taken from the localised geometry curricula they master.

The affordance is both at the tool level and at the conceptual level: the ontology has been kept rather simple in order for curriculum encoders to be able to handle it easily. For this reason we gave up on presenting relationships between competencies and between topics beyond the is-a and subclass relationships.

Protégé has been chosen as an editing tool (see figure 1) both for the design of the ontology and for its first validation by a small group of experts. It offers OWL-DL editing, which is coherent with the theoretical requirements. To our experience and as seems to be reported by many, its stable versions are usable by non-computer-scientists or non OWL-specialists while allowing specialists to perform deep editing tasks; we had success with the majority of our curriculum expert. It is the most widely used ontology editor at this moment. The new version (March 2009) also integrates DL reasoner like Pellet and Fact++.

### 2.3. The Geoskills Ontology in Details

For each Geoskills ingredients, a set of names is provided, at least one in each language, as explained in 2.2. This allows elements of the ontology to be presented to the user but also lets her search for it following an auto-completion mechanism described below, finally it supports the search engine to search for nodes being queried for text. Because names vary in their frequency of usage, they are of four different degrees (common, uncommon, rare, false-friend), which are taken in account when a word is matched with it. These names are not being mistaken with identifiers, which are ASCII names expected to be used in such references as URIs or URLs (e.g. when browsing about a topic).

For the ontology elements, names are encoded as dedicated datatype properties: for classes and properties, `rdfs:label` annotations are used while, for the instances, the properties are `commonName`, `unCommonName`, `rareName`, and, `falseFriendName`, each describing a name with its particular commonality.

Geoskills essential ingredients are topics, competencies, pathways, levels and programs.

**Topic** is made as a taxonomy (see Figure 2), that is, a hierarchy of abstract classes each representing mathematical topics and objects. Multiple inheritance is possible thanks to OWL and is of great use in this case. Because OWL-DL properties only relate on instances, each class has a single representative individual.

Property attached to Topic (other than names):

- `belongsToCurriculum`, with range *Programme*.

Examples of topics include *isocetes triangle* or *ApproximationProcess\_for\_roots*.

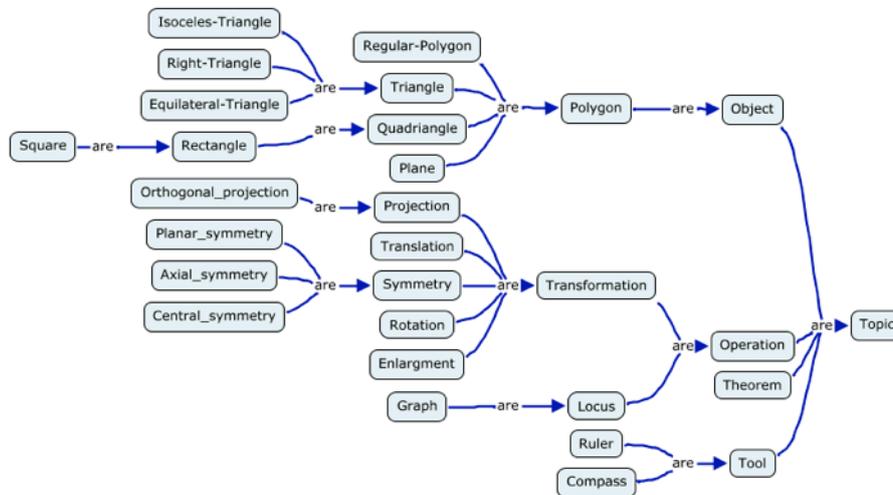


Figure 2. Extract of the topics hierarchy of GeoSkills

**Competency** is becoming the major entity of assessment and learning-plans. In GeoSkills, just as in [10] or [20], competencies are made of a verb and a set of topics. The class hierarchy of competencies represents the specialisation hierarchy of verbs, that is the *cognitive process* of the competency. Examples of competencies include *Calculate\_trigonometric\_ratio*, *Reproduce an isosceles triangle*, or *Identify\_square\_numbers*. In the first case, calculate trigonometric ratio, the OWL individual is of the class *Calculate* and contains the topic *trigonometric ratio*.

Properties attached to Competency (other than names):

- belongsToCurriculum with range *EducationalProgramme*
- hasTopic with range *Topic*.

**EducationalRegion** is an administrative educational region such as *Germany-Berlin*. or *France*.

**EducationalPathway** is a series of educational contexts such as elementary-school, or *Secondaire de Qualification Technique Artistique*.

Property attached to EducationalPathway (other than names):

- *inEducationalRegion* with range *EducationalRegion*.

**EducationalLevel** is an element of a pathway, for example one of its year, for example *Gymnasium\_Saarland\_7te*, or *Bachillerato\_Ciencias\_y\_Tecnologia\_2*.

Properties attached to EducationalLevel (others than names):

- belongsToEducationalPathway with range *EducationalPathway*
- age with range *integer*
- hasTopic with range *Topic*.

**EducationalProgramme** is the concrete plan of a level within a pathway; it is bound to curriculum standards. A programme can contain a list of competencies or the URL of an HTML where they are referenced.

Properties attached to EducationalLevel (others than names):

- hasLevel with range *EducationalLevel*
- hasSubject with range *Subject*
- hasURI with range *string*.

The ingredients of this ontology are among the ingredients of the metadata structure that the Intergeo platform is manipulating [21]. Thanks to their name-ability and mathematical semantics, they can enter an information retrieval process for both the auto-completion paradigm and the search tool paradigm (as explained below).

#### 2.4. Ontology Maintenance Practice

In order to guarantee long-term quality of the competencies and topics, the management tasks have to be taken care of seriously. Aside of the many possible handcrafted error-reporting rules that will be written as needs arise, the OWL DL ontology nature provides us several tools ready to be used:

**OWL-DL axioms to constrain properties:** several axioms can be encoded as part of the ontology constraining the properties of the individuals. For example, one axiom stipulates that at least a topic should be as an argument of a competency individual.

**OWL-DL class membership by Extension:** OWL-DL allows axioms to state sufficient or necessary and sufficient conditions for an individual to be part of a class. This allows, for example, the competency individuals that are instances of the *Construct* competency class and have the *Ruler* and *Compass* topic to be instances of the *Construct\_figure\_with\_ruler\_and\_compass* class as well. Better, this allows to automatically classify competencies as being member of a class, for example *Use\_equality* class, that is also automatically considered as a subclass of *Use\_identity* (because an equality is an identity).

**Abstraction for Similar Individuals:** OWL-DL axioms can inject automatically property values implicitly based on class-membership. This is of particular use in countries where many educational regions exist such as Germany (each of the 11 Bundesländer has its own set of educational programmes) or Switzerland (each of the 22 Cantons has its own set of educational programmes). This abstraction allows speaking of a ninth class of German Gymnasium with the same common-name (*9. Gymnasium, neunte, 9.*), which then gets specialised per *Bundesland*. All these abstractions take advantage of the ontology nature of our list of competencies and topics. The fact that the description logics formalism is used provides reasoners with a decidable algorithm. Our current usage is based on the Pellet reasoner [16] version 1.5, an open-source, java-based, OWL-DL reasoner.

#### 2.5. Populating the Ontology

Populating the ontology has two purposes. The first one is to provide a first level of validation of the ontology structure. The second one is to provide to the Intergeo web platform a first set of encoded curricula.

In the first phase, the curriculum experts of our group of the Intergeo have been involved closely followed by touches at the ontology structure, advise on the best ways to encode, and adjustments at the editing interface. The Protégé editor was the main editing tool.

We have made use of the Protégé client-server<sup>3</sup> which allowed team members to work synchronously on the ontology from remote places provided they are equipped with a very good network connection; only Universities met this challenge thus far. For other members, in particular companies involved in the Intergeo project, it was necessary to allow exclusive work on a local copy. Another limitation was met in the generic ontology-editor nature of Protégé, which makes it able to perform all sorts of changes, many of which should be reserved to ontology experts.

### 3. The Competency Editor - CompEd

For the reasons enumerated above, a platform to edit the ontology was needed, a web-based editing tool that allows curriculum experts throughout Europe to contribute by translating and editing competencies and topics.

Even before the editing actions, a first important aspect is to allow web-based navigation of nodes of the ontology to allow the annotation of curriculum texts and textbooks: both of these features are to be done by having topics, competencies, and levels addressable through URLs which can also be presented in a browser. The annotations edited in the Intergeo platforms use these links as part of their presentation as in Figure 3 aside from this paragraph.

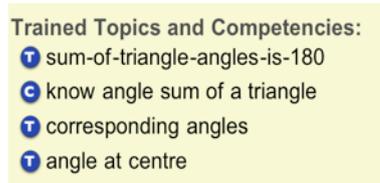


Figure 3. Rendered annotations of a resource.

The web-based editing tool is called CompEd. Its objective is to edit topic and competency individuals of GeoSkills as well as the topic and competency sub-classes and individuals. Editing includes altering names and relation properties (such as the generalisation/specialisation, instantiation relationships, or the involvement relationship of a topic in a competency).

#### 1. CompEd Features

CompEd offers the browsing and editing of individual topics, competencies, and competency processes. Individuals can be reached by tracking recent activity; by browsing the alphabetic list view or hierarchical tree view; by navigating the relationships; by keyword searching; or by an external URL.

Items are displayed in a consistent way. As depicted by Figure 4, which is an example for the "solve similarity problems" competency individual, the display is divided into three parts:

- The first part provides general information, which includes the name of the URI, the URI itself, the created and the modified dates. Below, the names in the user's language for the particular item are displayed. Names are grouped by type (common, uncommon, rare, false-friend). If wished, the user can click on the "more languages" link to get the other languages names.
- The topic part just provides a list of topics that are connected to the competency item. The list items are links, which simplifies the navigation to

---

<sup>3</sup>The Protégé client-server setting is based on Java RMI and is documented at [http://protegewiki.stanford.edu/index.php/Protege\\_Client-Server\\_Tutorial](http://protegewiki.stanford.edu/index.php/Protege_Client-Server_Tutorial).

the topic. Note there is no topic part in the view for topic items (only competencies are linked to topics).

- Finally, the structural part shows a hierarchical tree, which represents the generalisation/specialisation/instantiation path down to the competency item. In the case of competency classes (called *Competency processes* in the English GUI and *Catégories de compétences* in the French one and corresponding to the underlying cognitive process or verb), the tree will have all super-classes, subclasses, and individuals that are on the path through the competency process node.

Adding and editing of names as in Figure 4 includes the provision of a textual name, a language, and a type. The type can be one of: common, uncommon, rare, or false friend. While the latter pieces of information have a default value to be displayed in Intergeo tools (common name and the native language of the user), the validation through OWL axioms guarantees that a name is provided.

solve similarity problems

Welcome to CompEd where you can browse about the language of annotations of Intergeo resources, [GeoSkills](#).

Created: 05/10/2009, Last modified: 05/10/2009  
Uri: [http://www.inter2geo.eu/2008/ontology/ontology.owl#Solve\\_similarity\\_problems](http://www.inter2geo.eu/2008/ontology/ontology.owl#Solve_similarity_problems)

[\[Choose me\]](#)

**Names**

Common names [\[More languages\]](#)

solve similarity problems

Ähnlichkeits- Aufgaben lösen

résoudre des problèmes mettant en jeu des formes

resolver problemas de figuras semejantes

No uncommon names available  
No rare names available  
No false-friend names available

**Topics**

[similar](#)

**Structural Information**

Competencies (all processes)

Transversal Competencies

Solve

Solve in geometry

solve similarity problems

**Figure 4:** Presentation of a competency in CompEd (for curriculum encoder role)

Editing of competencies includes:

- changes, additions, and deletions of competencies
- alterations on the competencies' URI
- making connections to competency processes
- referencing to topics
- provision of a default common-name in any language.

Editing of competency classes is very similar except that connections are established to other competency classes (which denotes a subclass relation) and to competencies instances (which denotes a membership relation). CompEd supports the user in altering data as much as possible, i.e., it suggests default values and signals errors in a user-friendly way.

The remaining input that is not covered in the CompEd usage is that left for the ontology experts which includes adding or deleting extra properties, defining a class with a necessary and sufficient restriction, adding or deleting axioms about the ontology. Currently, edition of educational levels is also left to them, basically by using Protégé editor. They work informed by the curriculum encoding community based on a public forum where users of the curriculum knowledge, curriculum encoders, and ontology experts discuss.<sup>4</sup>

### *3.2. CompEd Architecture*

The CompEd server software has been designed with high-usability in mind based on web-technologies that are widely spread. Thus the AppFuse framework<sup>5</sup> is at its core and its memory management is supported by the RDBMS persistence engine MySQL through the widespread java persistence framework Hibernate.<sup>6</sup> These choices make CompEd a long-lasting responsive edition framework.

The decision not to use an OWL persistence engine is due to the apparently still lacking persistence framework for this technologies which scale long term and the ongoing need to load the complete ontology in RAM for most forms of reasoning.

## **4. Management of the Competencies and Topics**

Having described the editing framework we turn to the larger problem of maintenance of the knowledge in the GeoSkills ontology. It is done by the assertion of axioms following the OWL semantic and by its accessibility to users: either by browsing or by searching. These two aspects provide ways to make sure the utility of the knowledge of GeoSkills is maintained.

### *4.1. Ontological Management*

Two tools, CompEd and Protégé, can edit the Geoskills ontology. Protégé 3.3.1 has been the first editing tool for creating a Geoskills first version, used by two curriculum experts. It offers all the possible OWL expressivity. The normal tool to be used by curriculum experts is CompEd, but it offers an expressivity reduced to instances, hyponymy (is-a relation), links between competency and topics, and names. Because CompEd is unaware of axioms that have been expressed in OWL with Protégé, violations and new statements appear once the reasoning is invoked, nightly.

The Pellet classifier, on a dedicated server, makes these ontological consistency checks. This classifier provides also automatic classification of Competency classes

---

<sup>4</sup> The curriculum encoders' online community is being built at <http://curriculum.i2geo.net/>

<sup>5</sup> See <http://www.appfuse.org/>

<sup>6</sup> See <http://www.hibernate.org/>

between them, and of Competency instances into classes. For example, a competency can be automatically classified in the Competency class *Use formula* which is defined with a necessary and sufficient condition as “Subclasses of Use that refers to a formula” (see the example of 1.3).

We shall see below that this is done at synchronization time.

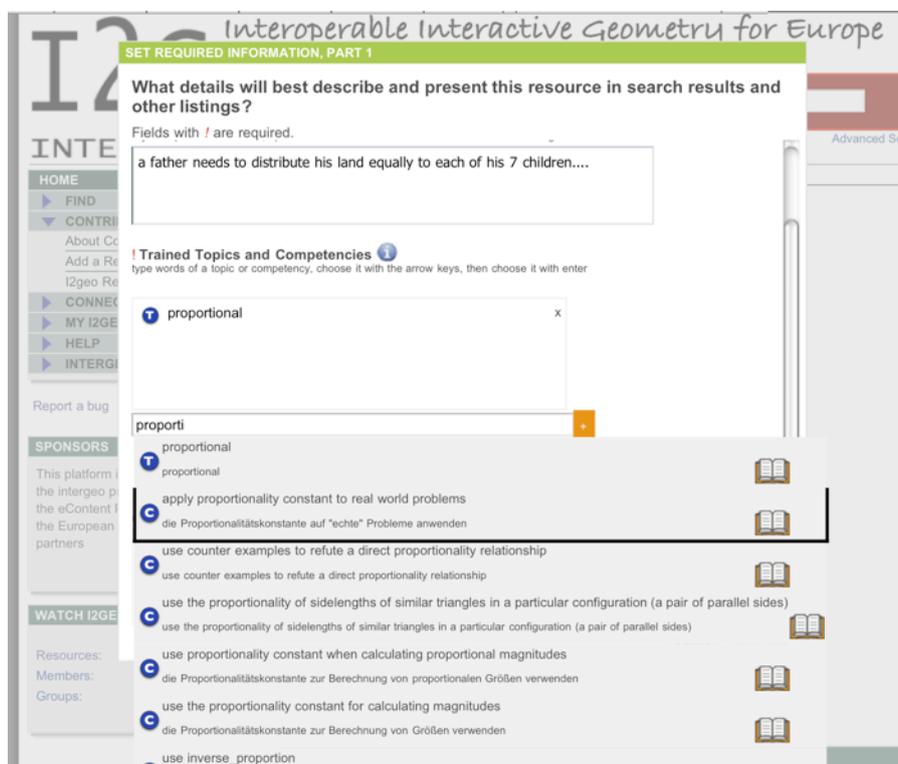


Figure 5. The second step of the wizard add-a-resource in the Intergeo platform.

#### 4.2. Access to Competencies by Typing: SkillsTextBox

To allow users to identify the competencies, topics, or levels they mean, we extend the familiar auto-completion: users can type a few words in the search field, these are matched to the terms of the names of the tokens; the auto-completion pop-up presents, as the user types, a list of matching tokens as seen on figure 5). This list presents, for each candidate, the default-common-name, the name found to match the user’s input, the number of related resources, an icon of the type, and a link to browse about the ontology at the node and around it. When chosen using either a click, or a few presses of the down key followed by the return key, the choice action either triggers a search or the addition of the node in a list, or for annotations.

SkillsTextBox uses a simple HTML form equipped with a GWT script . This script submits the fragments typed to the index on the server, which uses all the retrieval matching capabilities (stemming, fuzziness through edit distance or phonetic matching) to whose names start with the typed input, first in the languages supported by the user than in any language. The index returns the 20 best matching tokens and the script

renders as an auto-completion list. More information about it is at <http://www.activemath.org/projects/SkillsTextBox/>.

### 3. Designating by Pointing in a Text

Supplementary to letting users search for resources by explicitly typing the names of competencies and topics, we offer the possibility to do this implicitly by linking them from sections of curriculum standards or of textbooks that users know well (see for example Figure 6). Although we shall mostly not be able to offer whole textbooks to browse through, we expect it to be unproblematic to display their tables of contents and have obtained the yes of several publishers already.

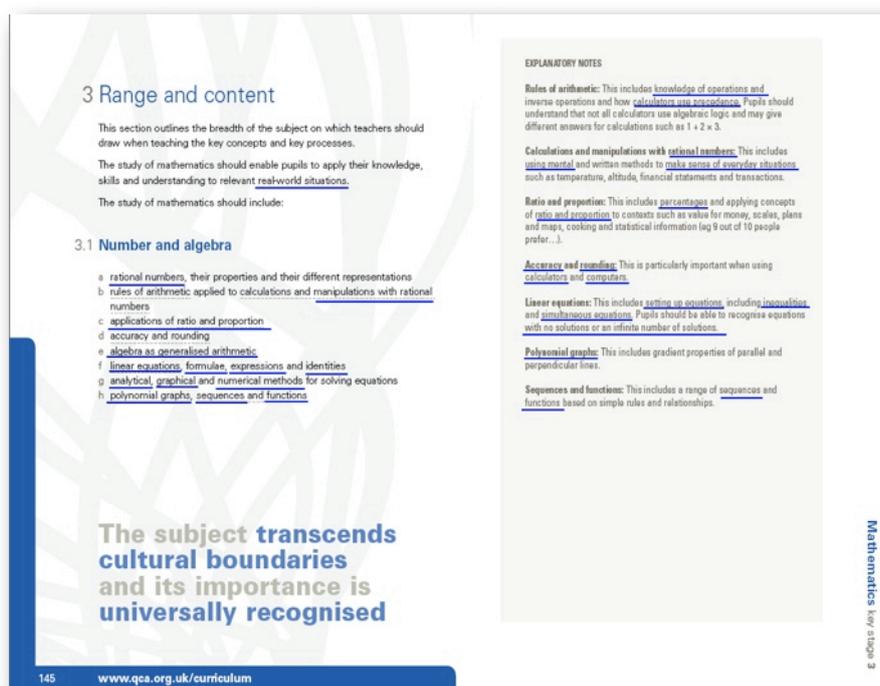


Figure 6. An annotated curriculum standard of England's KeyStage 3 maths

Linked curriculum-texts are obtained by letting curriculum experts edit HTML forms of the curriculum-texts adding, hyperlinks containing comma separated URIs of the ontology. A small process then converts these hyperlinks to javascript calls: once in context with skills-text-box instance, a click on such a links triggers a choice action which amounts to a search-query or an set of added annotations.

The idea is that a user can then browse through a table of content, through pages he is graphically familiar with, and click on sections of interest. This click triggers the selection of the competencies and topics associated with these sections, triggering the search for the related term or adding these nodes to the list.

### 4. CompEd, OWL, and the Term Index: Synchronisation

The competency-editor, the Protégé editor, the skills-text-box' term index all are places which store a representation of the GeoSkills' ontology; in this section we explain how

the OWL ontology file is at the centre of the synchronisation with incremental updates and regular resets.

The architecture of these pieces is depicted in Figure 7: CompEd stores the contents of GeoSkills in a way made for massive collaborative edition; it cannot allow edition of all facets of the ontology; on the other side, Protégé allows full edition of the OWL ontology but is not suitable for such collaborative edition; the ontology server stores the ontology in RAM and performs the reasoning but it only receives the updates done by the CompEd users through update XML documents which are then incorporated into the OWL file. Finally, the term index contains an index of the names, ready for retrieval in the auto-completion and search functions.

The communication flows between the pieces are as follows:

**CompEd updates:** following the actions of a curriculum-encoder or curriculum-translator, CompEd modifies his RDBMS storage and also sends an update document to the ontology server and to the term index. The latter update their representations following these updates.

**Regular resets:** because the intent of the competency editing process is the GeoSkills ontology, the ontology is used to replace the contents in RDBMS. This is done through a conversion from the OWL file, read through the reasoner, to the tabular format. These resets are applied every night and are the key to receive the reasoner results (such as the axioms that add properties or classes).

**Ontology adaptation:** from time to time, having concerted themselves, the ontology engineers will request to work at the ontology level, for example to add axioms, to add particular new classes or to perform clean-up operations. In order to do so, the CompEd server is taken read-only and the work on the OWL file in a text-editor or using Protégé can happen. It is followed by the regular reset, which re-imports the OWL file in the RDBMS.

## 5. Usage of the Competencies Ontology in the Intergeo Platform

We have described the representation and technical choices to allow GeoSkills to be used as a language of annotation. In this section we present how the software pieces are used in the Intergeo platform, which is publicly available at <http://i2geo.net/>.

### 1. User Roles in Intergeo Platform

The Intergeo platform's main goal is to allow sharing of interactive geometry 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. Overall, the usage of the platform is the execution of the following roles:

- **The annotator** uses the editing front-end of the community platform in order to annotate resources as referencing the given competencies or topics, and a given educational-level, as well with many other information fields (such as authorship or license).
- **The searcher** uses text-search, the ontology or curriculum-text browsing to identify the correct term so as to search through the platform's database to find relevant resources to use in teaching, to edit, or to evaluate.

- **The curriculum encoder** identifies a curriculum-text of interest that could be shared among platform users, obtains an appropriate electronic version, browses through it and creates, in the ontology, the needed competencies and topics. This may also require to declare a new learning-pathway, region, or programme. He is in charge of uploading the document into the web application for further sharing and adding hyperlinks into this text.
- **The competency translator** uses the competency-editing tool in order to add or edit titles of a competency or topic or descriptions in one's own language. Typically, they require knowledge of several languages but do not require understanding the data-model of a competency (hence cannot change it).
- **The platform translator** translates the messages of the system, a large dictionary modelled after the classical application internationalisation practice, while **the web-content translator** translates the pages of the Intergeo platform, which represent static texts. The XWiki infrastructure to this end is taken and works well.
- **The quality evaluator** role is described in [22].
- **The ontology engineer**, together with the **platform administrator**, monitor the reset processes, operate changes on the ontology for any facet not covered by the CompEd tool, such as edition of the axioms or educational levels.

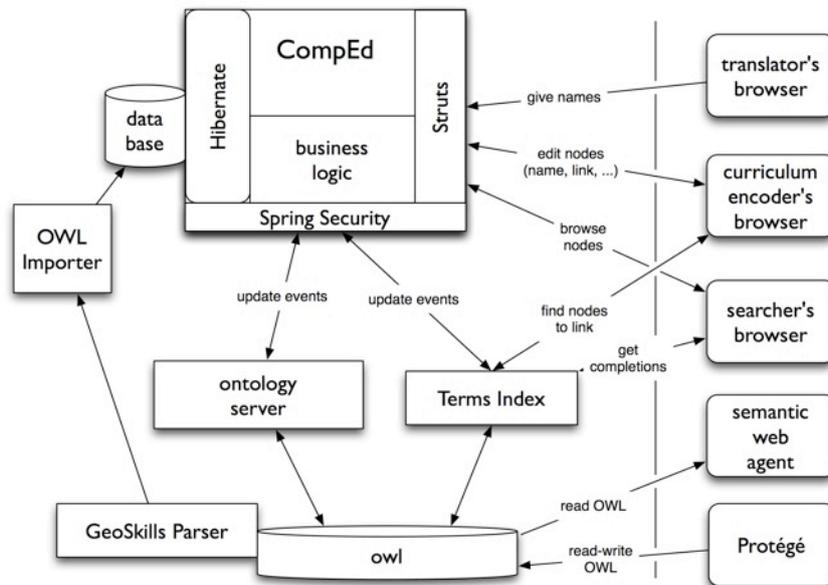
Thus the role of the annotator is to provide sufficiently detailed topical and educational context information so that all users can find resources using the language of their curriculum as well as using everyday language. For this to work, we have added two roles to this workflow: curriculum encoder and competency translators. They make sure that each competency and topic in the curriculum standard they are responsible for is properly listed and properly imputable.

## 2. *Contribute a Resource*

The annotator actions are the classical form-based editing of the metadata. The process is part of the Curriki platform whose metadata schema is relatively lightweight. Intergeo has adapted these forms to contain the fields for *trained topics and competencies* and *intended educational level* which both use skills-text-box. This is depicted in figure 5.

## 3. *Find a Resource*

We have explained in section 4.2 the retrieval process that searches through the labels of the ontology and proposes a completion list of tokens. Once chosen, the tokens can be used for the annotation or represent particular query-terms of the search tool. In this section we explain the search tools' query mechanism, which, again, relies on the knowledge of the ontology.



**Figure 7:** architecture of the CompEd and the terms-index.

In making the search tool, we rely on classical information retrieval principles which stipulate an easy query and result process with a result-list ranked by relevance as described, e.g. in [23] and using the Lucene library [24].

The query is made of a set of terms each made of a string (a set of words). Some strings represent a single node of the ontology, while others represent an arbitrary textual query. For each string, a query expansion is performed as follows:

- Each string is expanded to a query for the competencies, topics, or levels, whose names match closely the string, along with a query for the resources whose text contains the string.
- Each query to a resource annotated with a competency is expanded to a query for resources annotated with this exact competency or, less importantly, resources annotated with topics of that competency, or resources annotated with parent-competencies.
- Each query to a topic is expanded to match resources annotated by the topic or by its parent or children topics.

This query expansion mechanism, which is detailed further in [25] is the key to a tolerance of the search tool, a fundamental criterion of search tools' acceptance. This tolerance is enabled by the knowledge stored in the competency ontology, transporting the distance between ontology nodes to a distance between resources and their annotations. For example, it will allow an English-speaking teacher to search for *enlargement* as node or as text in the platform and still find what a French teacher will have annotated with the competency of *applying the intercepting lines theorem*. This goes beyond a simple *term-translation* approach because of the semantic relationships between the nodes, which are more important than the word similarity.

The query-expansion is a process that adds weights the queries it produces, this is called *boosting*. This enables, for example, a query for the textual content of the

resource (its title, description, or body) to match lower than a query for a topic or competency annotation. Similarly, the query expansion mechanism involves the user's context by preferring resources that are marked appropriate for the users' preferred educational levels.

This boosting is natively supported by the Lucene search engine used. Contrary to symbolic query engines (such as SQL or RDF storage engines), the default disjunction of a query is a weighted query where a match of several terms adds to the score of a match. Being a retrieval engine, the index structure allows Lucene to return a search result listing documents order by the score of the match, which is considered a relevance ranking.

## 6. Conclusion

Our mixed approach appears to succeed in the community-based creation of an interoperable language to annotate learning materials with a fine-grained knowledge about competencies and topics. This knowledge seems to be fundamental for the Intergeo repository of interactive constructions so that teachers, our target users, are able to publish and find the resources without being limited by the boundaries of their educational curriculum or their language. The semantic nature of the language helps its management from the validation as well as the accessibility point of view. Finally the usability of the language is supported by an easy input and browsing.

### 6.1. Implementation Status

The GeoSkills ontology is now stable in structure. The geometry parts of the topics and competencies of the curriculum of several years of the French collège and a few England's years are fully encoded, and in parts for of the curriculum of Cataluña. These parts have been done using the Protégé editor. The curriculum standards of the Czech republic, and the German state of Bayern are mostly encoded using CompEd. Further development from these Luxembourg and the Netherlands expected soon. At the time of writing, the GeoSkills ontology contains 120 competency-processes, 749 instantiated competencies, 427 topics. This ontology is available under either the Creative Commons Attributions-Sharealike License [26] or the Apache Public License 2.0 [27] from <http://i2geo.net/ontologies/dev/GeoSkills.owl>.

The Intergeo platform is available on <http://i2geo.net/> and already contains about 1500 resources with a preliminary recollection of 3000 resources, which were done in a first phase with a shallow metadata model. It is built as an adaptation of Curriki [28], both being delivered under the GNU General Public License [29]. Its installation is documented in [30]. The list of currently actionable curriculum-texts is at <http://i2geo.net/xwiki/bin/view/Main/CurriculumTexts>.

The simple user-interface of cross-curriculum search-engine can be used there, allowing search by competencies, topics, and levels and, of course, plain text.

More information about the platform can be seen in [31].

### 6.2. Perspectives

Among the avenues to be explored deeper is a more synthesised and complete exploitation of the conclusions of the reasoner. While inherited property values are easily handled by the parsing infrastructure which uses the reasoner, the automated

classification results have been ignored thus far because it would make any parent class a direct subclass of the node: at least in the competency editing process, this is wrong as it would flatten the whole tree of inheritance (e.g. as in figure 4). We have to explore such avenues as taking parent-classes inferred by the reasoners and removing the asserted ancestor parent-classes.

Beyond parsing, there should also be the possibility of the ontology server to feedback on changes done in the curriculum editing process, including indicate inconsistencies that have appeared. The XML encoding of the updates could be of use for this purposes.

The commitment to encode the curriculum standards of mathematics of many European countries seems to be novel and start on the strong basis of a usable editing tool and internationalisation infrastructure. The perspective of such a large coverage may uncover new cross-lingual issues.

Among the practical issues we have encountered is the desire of curriculum encoders to adjust URI-fragment-identifiers to be more precise or more correct especially when they display a typo or a wrong name. Such a change can break existing relationships and should, thus be discouraged. Except for a closed world where all the references can be updated, we miss management practices that would allow long term URI preservations while still allowing maintenance to bring external references up-to-date. It may be that the most adequate answer to this need is the denial of the practice of readable URIs, which removes from it any expressivity thus need to be adjusted.

Finally, a part of the GeoSkills, which seems to have a large potential for re-usability is the part about educational contexts, which catalogues educational regions, pathways, levels, and programmes within an ontology encoded in a standards-based knowledge representation. Based on reference texts, we seem to be able to provide the coverage of the full set of European schools in strongly structured way.

It has been our surprise that such an ontology is not yet available, the closest being a thesaurus with a structure which lacks a strong specification such as that offered by <http://www.eurydice.org/>.

## Acknowledgements

This work has been realised within Intergeo eContentPlus project, which was partially funded by the European Community and by participating institutions. The opinions of this paper are, however, those of the authors.

We wish to thank the members of the work-package 2 of the Intergeo project, in particular Martin Homik and Arndt Faulhaber (DFKI), Colette Laborde (CabriLog), Maxim Hendriks (TU/e), and Albert Creus-Mir (Maths4More).

## References

- [1] Ministère de l'Éducation Nationale, Programmes des classes de troisieme des colleges. *Bulletin Officiel de l'Education Nationale* **10** (1998), 108.
- [2] OFSET, GNU EDU, <http://gnuedu.ofset.org/>, Accessed 2008.
- [3] American Mathematical Society, *Mathematical Subject Classification*, <http://www.ams.org/msc/>, Accessed 2009.
- [4] J. Karhima, J. Nurmonen, M. Pauna, WebALT Metadata = LOM + CCD, in *Proceedings of the WebALT 2006 Conference*, The WebALT project, 2006.
- [5] G. Paquette, An Ontology and a Software Framework for Competency Modeling and Management. *Educational Technology & Society* **10** (2007), 1-21.

- [6] BECTA, *British Educational Communications and Technology Agency*, <http://partners.becta.org.uk/index.php?section=rh&rid=13661>, Accessed 2009.
- [7] Microsoft, *Microsoft Lesson Connection Launched At Technology + Learning Conference*, <http://www.microsoft.com/presspass/press/1999/nov99/lessonpr.msp>, Accessed 2008.
- [8] Explore learning, *Correlation of gizmos by state and textbooks*, <http://www.explorelearning.com>, Accessed 2008.
- [9] Key Curriculum Press, *Sketchpad Lesson Link*, <http://www.keypress.com/x22318.xml>, Accessed 2009.
- [10] F. van Asche, *Linking Learning Resources to Curricula by using Competencies*, in *First International Workshop on Learning Object Discovery & Exchange*, Crete, 2007.
- [11] D. L. Mc Guinness, F. van Harmelen, *OWL Web Ontology Language Overview*, W3C, <http://www.w3.org/TR/owl-features/>, Accessed 2007.
- [12] National Library of Medicine, *Protégé Editor version 3.3.1*, <http://protege.stanford.edu/>, Accessed 2008.
- [13] A. Kalyanpur, Bijan Parsia, E. Sirin, B. Cuenca-Grau, J. Hendler, Swoop: A 'Web' Ontology Editing Browser, *Journal of Web Semantics* 4 (2005).
- [14] Mindswap, *Swoop*, <http://www.mindswap.org/2004/SWOOP/>, Accessed 2009.
- [15] J. J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, K. Wilkinson, Jena: implementing the semantic web recommendations, in *International World Wide Web Conference*, ACM, New York, NY, USA, 2004, 74 - 83.
- [16] Clark&Parsia, *Pellet*, <http://clarkparsia.com/pellet>, Accessed 2008.
- [17] HP Labs, *Jena - a semantic web framework for java version 2.5.5*, <http://jena.sourceforge.net/>, Accessed 2007.
- [18] S. Pepper, G. Moore, *XML Topic Maps (XTM) 1.0 -- TopicMaps.Org Specification*, <http://www.topicmaps.org/xtm/1.0/>, Accessed 2009.
- [19] D. Dicheva, *Towards Reusable and Shareable Courseware: Topic Maps-based Digital Libraries*, <http://compsci.wssu.edu/iis/nsdl/>, Accessed 2008.
- [20] E. Melis, A. Faulhaber, A. Eichelmann, S. Narciss, *Interoperable Competencies Characterizing Learning Objects in Mathematics* In *Lecture Notes in Computer Science, Intelligent Tutoring Systems*, 5091, Springer, Berlin, 2008, 416-425.
- [21] M. Hendriks, P. Libbrecht, A. Creus-Mir, M. Dietrich, *Deliverable D2.4: Metadata specification.*, Project I, eContentPlus Program, European Community, <http://i2geo.net/files/D2.4-Metadata-Spec.pdf>, 2008.
- [22] C. Mercat, Soury-Lavergne, S., Trgalova, J., *Deliverable D6.1: Quality assessment*, Intergeo Project, eContentPlus Program, European Community, [http://www.inter2geo.eu/files/D6.1\\_060508.pdf](http://www.inter2geo.eu/files/D6.1_060508.pdf), 2008.
- [23] C. J. van Rijsbergen, *Information Retrieval*, Butterworths, 1979.
- [24] E. Hatcher, O. Gosnopedic, *Lucene in Action*, Manning, 2004.
- [25] A. Creus-Mir, C. Desmoulin, M. Dietrich, M. Hendriks, C. Laborde, P. Libbrecht, *Internationalized Ontology*, Project I, eContentPlus Program, European Community, <http://i2geo.net/files/D2.3-Intl-Ontology.pdf>, 2008, 52.
- [26] Creative Commons, *Namensnennung-weitergabe unter gleichen bedingungen 2.0 Deutschland*, Creative Commons, <http://creativecommons.org/licenses/by-sa/2.0/de/>, Accessed 2008.
- [27] Apache Software Foundation, *Apache public license 2.0*, Apache Software Foundation, <http://www.apache.org/licenses/>, Accessed 2009.
- [28] The Global Education Learning Community, *Curriki*, <http://www.curriki.org/>, Accessed 2009.
- [29] Free Software Foundation, *Gnu general public license version 2.0*, Free Software Foundation, <http://www.gnu.org/licenses/old-licenses/gpl-2.0.html>, Accessed 2009.
- [30] S. Egido, P. Libbrecht, H. Lesourd, *Platform's Administration Manual*, Intergeo Project, eContentPlus Program, European Community, <http://i2geo.net/files/D4.4-AdminManual.pdf>, 2009, 50.
- [31] P. Libbrecht, Kortenkamp, U., Mercat, C., *I2Geo: a Web-Library of Interactive Geometry*, in *Digital Mathematics Library Workshop*, Ed. Sojka P, 2008, 3-15.