Towards a semantic learning model fostering learning object reusability
Emmanuel Fernandes, Hend Madhour, Maia Wentland Forte, Sami Miniaoui

To cite this version:
Emmanuel Fernandes, Hend Madhour, Maia Wentland Forte, Sami Miniaoui. Towards a semantic learning model fostering learning object reusability. International Conference on Human System Learning (ICHSL) 2005, 2005, Marrakesh, Morocco. 11 p. hal-00190205
Towards a Semantic Learning Model Fostering Learning Objects Reusability

Emmanuel Fernandes:
INFORGE-HEC
BFSH1, University of Lausanne, CH-1015 Lausanne, Switzerland
Emmanuel.fernandes@unil.ch
Tel:+41 21 692 3427

Hend Madhour:
INFORGE-HEC
BFSH1, University of Lausanne, CH-1015 Lausanne, Switzerland
hend.madhour@unil.ch
Tel:+41 21 692 3423

Sami Miniaoui:
INFORGE-HEC
BFSH1, University of Lausanne, CH-1015 Lausanne, Switzerland
sami.miniaoui@unil.ch
Tel:+41 21 692 3429

Maia Wentland Forte:
INFORGE-HEC
BFSH1, University of Lausanne, CH-1015 Lausanne, Switzerland
mwf@unil.ch
Tel:+41 21 692 3429
Fax: +41 21 692 33 05
Towards a Semantic Learning Model Fostering Learning Objects Reusability

ABSTRACT
We try in this paper to propose a domain model for both author's and learner's needs concerning learning objects reuse. First of all, we present four key criteria for an efficient authoring tool: adaptive level of granularity, flexibility, integration and interoperability. Secondly, we introduce and describe our six-level Semantic Learning Model (SLM) designed to facilitate multi-level reuse of learning materials and search by defining a multi-layer model for metadata. Finally, after mapping different learning content models with our SLM, we show how our Phoenix authoring tool can deal efficiently with share-and-reuse problem.

Keywords
Semantic Learning Model, reusability, LOM, standards, Conceptual Segmentation.

1. INTRODUCTION
In the context of elearning, document reuse can be approached both from the author's and the learner's point of view.

On one hand, authors should be allowed to build a multi-source document [11] from scratch or based on elements extracted from different sources. In all cases, the possibility to both edit new material and/or include existing items (Disk or LOR) reusing them as such or modifying them on-the-fly should be offered.

We distinguish three types of sources [11]:
- The author produces a completely new document from scratch
- The author produces a new document by recomposition. Documents can be extracted from:
  (i) a Learning Object Repository (LOR)[9] containing indexed documents (document+metadata) so that all required information can be easily found and retrieved.
  (ii) the local disk which contains documents either indexed or not. Furthermore documents can downloaded from Internet, etc.

On the other hand, learners should be delivered a document and or a course according to their individual needs, which can be achieved by having recourse to an adaptive hypermedia application.

In [34], a division into three models is proposed when developing an adaptive hypermedia application upon which many models, such as [4][34], are based:
- The domain model is composed of a set of small domain knowledge elements (DKE). Each DKE[3] represents an elementary fragment of the given domain.
- The student model represents information about the user’s preferences, interests, attitudes and goals, proficiencies, history of interactions and user’s classification.
- The adaptation model describes how Adaptive Hypermedia System performs its adaptation.

Many issues are thus raised such as: at what level of granularity should learning items be segmented into to facilitate reuse? How can search and access to learning items be improved? Is the learning object meta-data standard (LOM)[20] sufficient for indexing and querying?

In this article, we will focus on the domain model as a starting point to realize an Adaptive Hypermedia System (AHS). More precisely, we will concentrate on finding a better way to reuse learning material by sharing them in different contexts. We will show how some of the hereabove interrogations could find a solution in a new learning objects content model (or domain model) called the Semantic Learning Model (SLM).

We will then define four criteria for an effective authoring tool needed to facilitate document elaboration (in the case of the author) and delivery (in the case of the learner). Our SLM model, which is inspired from the document structure provided by our authoring tool Phoenix [11] developed in the context of ARIADNE [1], will be detailed. Finally, we will show how Phoenix fits all fo the criteria.

2. CRITERIA FOR AN EFFECTIVE AUTHORING TOOL
A number of commercial products such as Blackboard, WebCT, or Lotus Learning Space provide authoring tools based on the ADL Sharable Content Object Reference Model (SCORM). This model is based on the
assumption that the user - be it a learner or an author - interacts only with the Sharable Content Object (SCO) layer, excluding any direct access to other sources of pedagogical content. Furthermore, the design of components cannot make use of pedagogical approaches such as constructive or collaborative learning.

Conversely, many non-SCORM systems - such as ActiveMath [19], MetaLinks [22], NetCoach [30], DCG [28], Interbook [4] etc. - offer a choice of pedagogical approaches but have a number of severe drawbacks like their lack of interoperability or flexibility, not to mention the fact that they are self-contained and therefore are unable to connect to external services nor facilitate reuse or collaboration.

Therefore, we sense a real need for an authoring tool that would combine the positive aspects of both worlds avoiding at the same time their inconveniences.

2.1 Adaptable level of granularity

In one of our previous articles [10], we have identified four criteria, apart from confidence in the system, that could help in motivating users (Figure 1) to go for and accept the principle of share-and-reuse:

(i) A high degree of confidence in the quality of what is produced by others;
(ii) An existing critical mass of documents indexed in the LOR, allowing the author to find something suitable;
(iii) An adequate level of granularity, so that authors can make use of or adapt the learning objects quickly and easily;
(iv) Rules that indicate clearly how the copyright issues are dealt with.

Out of these four, we believe that the most important is the level granularity of the learning object. In principle the smaller the easier to reuse it in another context or adapt it, but one has to always bear in mind that it needs to make sense. In this framework, Phoenix offers the possibility to segment a pedagogical document on-the-fly in as many learning objects as wished, index the segments and the whole document, input and retrieve them from a LOR, in our case from the Knowledge Pool System [9].

![Figure 1. Motivation of authors to share and reuse pedagogical documents](image)

2.2 Integration

Systems should support all the required functionalities in a learning process. A learning process [17] is composed of three steps:

(i) Concept selection: A learning goal is defined over a concept hierarchy graph (e.g. A semantic network [32]).
(ii) Pedagogical design: this process is based on the use of template learning scenarios. It allows selecting pedagogical strategy.
(iii) Content selection: It depends on the learner’s known characteristics such as an established degree of knowledge or mastering of the domain concepts and historical information about related learning activities.
2.3 Flexibility
The main idea is to generate a course tailored to fit the user’s specific needs. An authoring tool is considered to be flexible if it generates an individualized course, or path in the domain knowledge, taking into account the user’s profile, meaning learning goals, pedagogical objectives and the context.

2.4 Interoperability
Typically, Learning objects and their associated metadata are located in distributed LORs. However, there is more than one approved standard used to describe the properties of learning objects [20], Dublin Core[7], SCORM[26], CanCore [12]). Interoperability between LORs is provided thanks to the metadata mapping according to some metadata standards [25].

Our SLM model regroups the main concepts proposed in each taxonomy while defining the relationships between the content objects as well.

3. LEARNING OBJECT CONTENT MODEL
Some systems such as Learnactivity [29], Scorm [7], CISCO [14], NetG [18], IMS/LD [16] and General Model [9] dealt with constructing a classification for learning items in order to facilitate learning resources reuse.

It seems obvious that the smallest the level of granularity of a resource is, the most reusable it is but the less contextualized it is.

3.1 Level of granularity and semantic elements
In the framework of the segmentation [33] of a single document, we made the assumption that a domain could be represented by a finite number of presentation chains each containing at least a concept and its definition [32]. Each concept could be further explained and detailed through a series of pedagogical elements that were graphically regrouped to constitute the above mentioned presentation chain. Segmenting a document consisted precisely in identifying and marking the concepts and their related presentation chains so as to construct meaningful and contextually pertinent pedagogical elements. We could say that the biggest granularity of a document is the document itself while the smallest is any of the identified elements.

The entities it can be composed of are the following:

- A Concept is a semantic element explicitly defined in the text. Its definition is composed of either already identified concepts or of prerequisites defined elsewhere. It is characterized by a presentation order, a label, a gender, a type, a complexity degree and content.

- An Argument is a semantic element that refers to a concept and is used to familiarize, clarify or reinforce the concept. An argument is characterized by its pedagogical function and role, according to an existing typology [32].

- A Solved problem is a special type of argument that refers to several concepts.

- A TexteSimple is a simple element used to handle unmarked text.

The resulting semantic network highlights the definitional relationships between the concepts and the links between a concept and the pedagogical entities that are related to it in order to reach a pedagogical goal [32].

3.2 Taxonomy
In order to deal with multiple source document we have further refined our approach (Figure 2) and devised a new model baptized the Semantic Learning Model (SLM).
3.2.1 Components
We divide the SLM into 6 categories:

1. **Assets**: The lower level of granularity of a document is an Asset. Assets can be pictures, illustrations, diagrams, audio and video files, animations, and also text fragments.

2. **Pedagogical information**: A pedagogical information is a group of assets expressing the same meaning. For example, a figure associated with its comment is a pedagogical entity.

3. **Pedagogical entities**: It's an information entity associated with a pedagogical role. Four roles are specified: concept, argument, solved problem and simple text but for the model to be parametrizable, the role can be anything else as long as it is previously defined by the pedagogue.

4. **Pedagogical context**: it represents the semantic structure (or network) in which active pedagogical entities are grouped. In this phase, semantic network is built before or during the pedagogical context creation. Pointers to pedagogical entities are organized following the semantic network structure.

5. **Pedagogical document**: the pedagogical document includes the pedagogical context associated with prerequisites.

6. **Pedagogical schema**: Many pedagogical documents are grouped in order to make a curriculum. This group is called pedagogical schema.

3.2.2 Mapping between different models
Table 1 compares the different e-learning Content Models including with our SLM:
Table 1. Mapping between different Metadata Models

<table>
<thead>
<tr>
<th></th>
<th>Assets</th>
<th>Pedagogical Information</th>
<th>Pedagogical Entity</th>
<th>Pedagogical context</th>
<th>Pedagogical document</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Model</strong> ([Duval02])</td>
<td>Content fragments</td>
<td>Content Objects</td>
<td>Learning object</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Learnativity</strong> ([Wagner02])</td>
<td>Raw Media</td>
<td>Information Object</td>
<td>Learning Object</td>
<td>Aggregate Assemblies</td>
<td>Collections</td>
</tr>
<tr>
<td><strong>SCORM</strong> ([Dodds01])</td>
<td>Assets</td>
<td>SCO Content</td>
<td>Aggregation</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CISCO</strong> ([Barrit99])</td>
<td>Content Items</td>
<td>RIO</td>
<td>RLO</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IMS/LD</strong> ([IMS04])</td>
<td>Learning objects</td>
<td>Environments</td>
<td>Role parts</td>
<td>Acts</td>
<td>Plays</td>
</tr>
<tr>
<td></td>
<td>IMS Simple Sequencing</td>
<td>Support/Learning Activities</td>
<td>IMS Simple Sequencing</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Netg</strong> ([L'allier97])</td>
<td>Raw content</td>
<td>Objective activity</td>
<td>Assessment</td>
<td>Topic</td>
<td>Lesson</td>
</tr>
</tbody>
</table>

- CISCO [2] identifies RIOs, assessments, overviews and summaries, which can be mapped onto pedagogical information. An RLO is an aggregation of these components and can be assimilated to the pedagogical context. RLO/RIO model defines the components of a learning object more strictly: the model specifies that a learning object (RLO) contains $7 \pm 2$ RIOs, whereas the presented model does not restrict the number of components of a learning object.

- Within the SCORM aggregation model, an asset can be associated to an asset. SCOs can be associated with pedagogical information and content aggregations can be mapped onto pedagogical context.

- The learnativity model maps easily onto the represented model. Raw media elements are associated with assets. Information objects like processes and procedures are abstract types like pedagogical information. Learning objects and aggregations fits within the represented model. The three aggregation levels of the learnativity model (learning objects, aggregate assemblies and collections) are included in our model.

- NETg uses the term learning object which comprises a learning objective, a unit of instruction that teaches the objective, and a unit of assessment that measures the objective. These are abstract types, which can be assimilated to a pedagogical information. NETg defines aggregations that fits the constraints of our model.

- The General Model uses content fragments which can be mapped onto assets. Content objects can be associated to pedagogical information. The Learning Object can be considered as an aggregate including the pedagogical entity, pedagogical context and pedagogical document.

- IMS/LD utilizes the learning object as any other resource whatever its level of granularity. So, a learning object represents the lowest level of granularity which corresponds to assets in the SLM model. An environment is composed of different resources and can thus be mapped onto pedagogical information. A role matches nicely a pedagogical entity because of the added value it contains compared to the level below.

An activity can be mapped onto a pedagogical context due to the fact that at this level the different pedagogical entities aiming at the same target are aggregated. Eventually, a unit is mapped to pedagogical document.

SLM model is a way to improve learning objects reusability. To reuse a pedagogical item, we need to search by submitting queries, find and retrieve it.

A way to facilitate this task is associating pedagogical items through different levels of granularity to a metadata.

3.3 METADATA

Traditionally, metadata is understood as a data about data. This data is formatted according to rules or concepts defined in XML [35] or eXtensible Markup Language. XML and metadata are extremely useful to describe digital media and especially 'rich media': video, audio, and streaming media.
Metadata is recognized to be a key element for discovery, interoperability, accessibility and reuse of digital educational content. It standardizes the descriptive language of digital content, provides consistent, searchable keywords through ontologies and controlled vocabularies in meta tag elements and modifiers. It offers methods for usage and access control through digital rights management systems. The development of metadata standards in response to the need to deliver digital educational material is a relatively new endeavour, driven by the network based initiatives to deliver on-line material. Learning metadata standards pose different levels of representation granularity[24].

3.3.1 Ariadne Metadata

The ARIADNE Knowledge Pool System is a distributed repository for learning objects [8]. It encourages the share and reuse of such objects. An indexation and query tool uses a set of metadata elements to describe and enable search functionality on learning objects. To ensure simplicity, understandability and adaptability for the ARIADNE community, data elements are grouped into six categories:

1. General: groups the general information that describes the learning object such as document title, document language, etc.
2. Semantics: groups elements that describe the semantic classification of the learning object like the science type, main discipline, sub discipline etc.
3. Pedagogical: groups elements that describe the pedagogic and educational characteristics of the learning object such as semantic density, interactivity level, etc.
4. Technical: groups elements that describe the technical requirements and characteristics of the learning object like OS version, required disk space, etc.
5. Indexation: groups elements that describe the general information about the metadata itself of the learning object such as the identifier of the metadata instance, metadata creation date, creator, etc.
6. Annotations: groups elements that describe people or organizations notes about learning objects like annotator, language of annotations, and date of annotation.

These specifications together with similar specifications contributed by IMS [15] served as the starting point for the IEEE LTSC LOM standard [20]. LOM has a wide set of globally agreed metadata elements.

Metadata data elements of LOM are grouped into nine descriptive categories: General, Life cycle, Metametadata, Technical, Educational, Rights, Relation, Annotation, and Classification. These specifications have been defined and agreed on by a global community to enable share and reuse.

3.3.2 Assets, Information entities, pedagogical entities

Ariadne metadata description for author, publication date and source is sufficient to preserve copyrights and for indexing.

3.3.3 Pedagogical context

Almost all standards are defined for representing content characteristic-based information. In the case of LOM for instance, metadata information can be ‘keywords’, ‘author’, ‘version’ or even ‘size’ of the document. Only a few metadata items, such as ‘learning resource type’, ‘intended end user role’, ‘typical age range’ or ‘difficulty’, are dedicated to educational purposes. These are not yet sufficient to represent different learning context. We need a context sensitive metadata for representing context, objective and semantics of a learning document. So, we propose to add a field called "context" to express this metadata information.

3.3.4 Pedagogical document

In so far a pedagogical document is concerned, information about prerequisites is needed to choose what document is the most appropriate for the learner.

3.3.5 Pedagogical schema

A pedagogical schema encompasses a program of study, a workshop, a course, a module, a lesson. Three attributes could be assigned to the global metadata: topic, creator and description.

In fact, as the schema creator is often not the author of all the documents included in the schema, this must be specified in the metadata. Figure 3 shows different metadata layers.
In this section, SLM is confronted to the different criteria presented previously for an effective authoring tool: Phoenix.

4.1 Adaptable level of granularity
SLM offers many levels of granularity for effective learning items reuse. The author need differs from the whole pedagogical context to a simple asset (an image for instance). Moreover, a learner needs a pedagogical document according to his prerequisites and not the pedagogical context.

Finally, SLM offers many levels of learning reuse according to the user needs.

4.2 Integration
SLM could ameliorate the learning process:

   (i) Concept selection: A semantic network is designed in order to reach a specific pedagogical goal. This phase is based on the pedagogical context.

   (ii) Pedagogical design: this process is based on the use of pedagogical schemas.

   (iii) Content selection: It is applied depending on the learner’s knowledge of prerequisites. This phase is based on the pedagogical documents.

4.3 Flexibility
SLM brings great flexibility because it allows to organize the same assets differently according to different semantic networks. For instance, the same pedagogical information can play different roles in different contexts: i.e. it can have the role of concept in a pedagogical context and the role of argument in another. The resulting pedagogical document will then be composed taking into consideration the context and needs.

4.4 Interoperability
Our aim in adopting the SLM is to bring a common structure to pedagogical documents (XML) and to exploit the XML schema for a semantic description of the context. Furthermore, the usage of XML as a unified data format brings interoperability between our pedagogical documents with ARIADNE metadata and other pedagogical documents with LOM metadata for example. On one hand, the Ariadne Metadata are LOM compatible and on the other hand, the mapping from one Metadata standard (expressed in XML) to another can be done with an XSLT script efficiently. For instance, the mapping between LOM and Dublin Core is yet done [21] and other works are dealing with defining mutual mappings between different metadata standards and specifications (LOM [20], Dublin Core [6], SCORM [26], CanCore [12] and IMS MD [16]). In the following
Figure 4, the figure illustrates how the transformation (Ariadne into LOM.xsl) from ARIADNE to LOM is based semantically on the conceptual mapping between the related XML Schema[23].

Furthermore, document insertion in the Knowledge Pool System (an example of LOR) is a non-trivial task. For documents with Metadata, metadata has to be specified (author, title, etc) and concepts, arguments and solved problems have to be defined. Generating an ARIADNE (LOM compatible) file header can be automated (Figure 5), but because of its semantic-based nature defining concepts, arguments and solved problems it is more complicated. Moreover, some metadata can be inferred on-the-fly, just when we select our new pedagogical document.

Globe [13] is a new federation that relies on querying. The Globe federation works by putting an agreed "wrapper"—Simple Query Interface (SQI) [27]—around the search interfaces the constituent repositories already have.

5. CONCLUSION
Our aim was to find a model for learning materials in order to facilitate their reuse by both authors and learners. In the case of learners, we dealt with adaptive hypermedia and especially as a starting point with the domain model. The domain model corresponds to learning content model when considering author's side.

We have mapped SLM with other learning content model but not to other domain model. In fact, almost works [4] [5] divide their domain model into a set of concepts (atomic and composite) linked by different relations (prerequisite, inhibitor, simple link, etc.). There is no hierarchy unless that a composite concept is a set of atomic concepts. So, it did not seem necessary to map SLM with other domain models.

Moreover, we defined four criteria to reach our aim and have an "effective" authoring tool and we showed how Phoenix suits those criteria. As the Phoenix structure was the starting point of building SLM model, it will be easy to decompose phoenix document according to our taxonomy (we don't consider here pedagogical schema).
As a result of this decomposition, we will have a set of assets pointed (directly or indirectly) by multi-level pointers.

Furthermore, after populating a LOR following our SLM model (ie domain model), we will consider user model and adaptive model:

(i) The user model will regroup two kinds of information: fixed one such as user profile and flexible ones such as proficiencies and interaction history.

(ii) The adaptive model consists of finding an algorithm allowing document delivery according to the user model and updating flexible information.

6. REFERENCES


[27] Simple Query Interface (SQI), http://www.prolearn-project.org/ori.


[29] Wagner E. D. Steps to creating a content strategy for your organization. The e-learning developers' journal (October. 2002).


