

Document-centered Learning Object Authoring

1. Introduction

Learning Object (LO) open standards (e.g., SCORM [ADL 2004]) have achieved stable specifications, providing reliable ways for exchanging digital content and fostering environments and tools to offer support for LO storage, indexing and presentation. Relevant examples are Learning Management Systems, like Moodle (<http://moodle.org>) and Sakai (<http://sakaiproject.org>), which accept SCORM packages. This scenario motivated our initiative to build a digital content repository for storage and sharing e-learning content, fostering reuse at our university.

The process of e-learning content production at our university is guided by a workflow detailed in Figure 1. It starts up when a professor/author submits written documents by using a popular text-processing tool (content authoring). After content and style review a course goes towards the course authoring process. Course authoring addresses tasks of: systematizing the course content according to a learning design, course description using learning metadata and preparing the content to be reused.

Following this workflow the author acts in two distinct stages: content production and content authoring. These two stages can be also seen in another perspective, the former concerns mainly *data production* and the latter mainly *metadata production*. Therefore, we here consider the content authoring a kind of annotation task of the produced content. We have remarked three main practical difficulties related to the annotation process:

1. The annotation task appears as a disjoint *additional step* after the content production process.
2. The *metaphors* adopted by annotation standards and tools are distant from those adopted by content production tools to which authors are used.
3. When production+annotation tools are adopted for developing specialized content, the *lack of homogeneous content format* has hampered the content reviewing process.

The remarks above have led us to a research that issued in a methodology for content annotation, which we call *In Loco Semantics*.

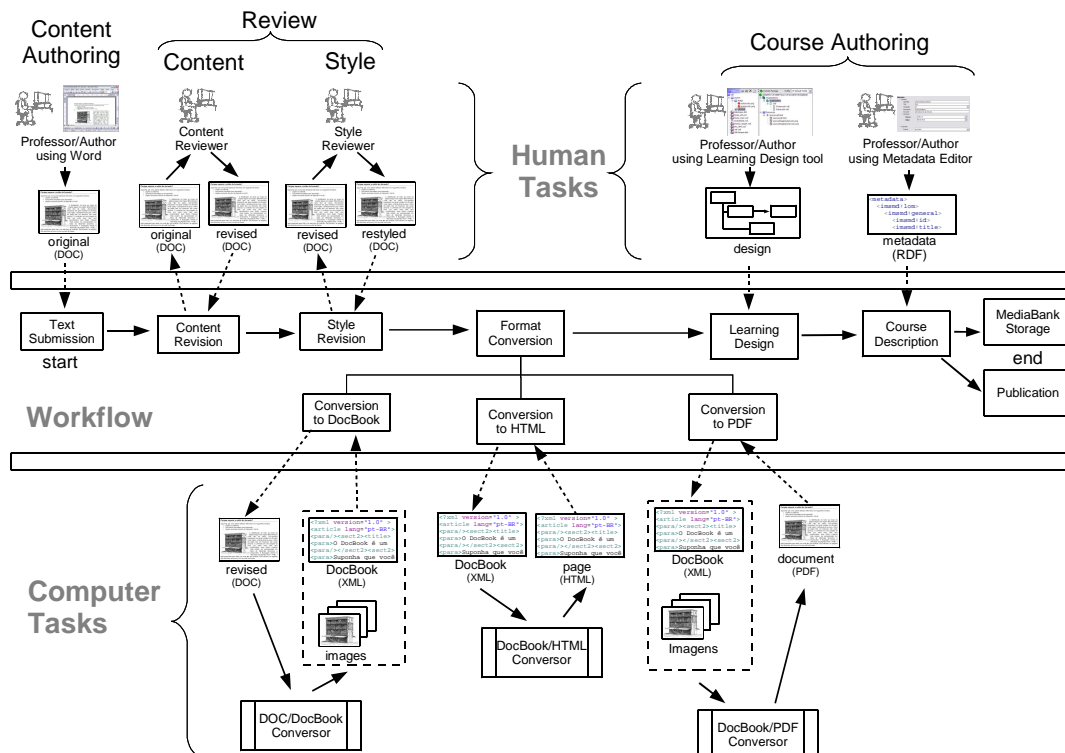


Figure 1. Workflow diagram of an e-learning course production.

2. In Loco Semantics

In Loco Semantics is a methodology for producing annotations linked to digital content by using interoperable semantics and ensuring semantic persistence, with the following principles:

1. *In Loco Annotation*: the annotation process occurs concomitant to the content production (in loco).
2. *Metaphor Integration*: the metaphors and models adopted in content annotation are aligned with those adopted for content production.
3. *Interoperability*: in loco annotation strategies are designed to enable automatic information extraction and conversion to Semantic Web open standards.
4. *Semantic Persistence*: in loco annotation elements are connected to unification ontologies, which will guarantee their equivalent interpretations in different contexts, subsidizing semantic persistence among transformations.

The diagram in Figure 2 synthesizes a typical scenario of the methodology application. This scenario can be organized into three distinct activities: (1) **design** – a content profile is designed including its templates or schemas; (2) **production** – a content is produced and annotated following an *annotation pattern* and based on a template or schema; (3) **CO transformation** – the original content is transformed by an information extraction tool, which can map the content to a domain neutral complex object (CO), (4) **Domain transformation** – the domain neutral CO is mapped by a conversion tool to a domain specific representation, a LO in this context.

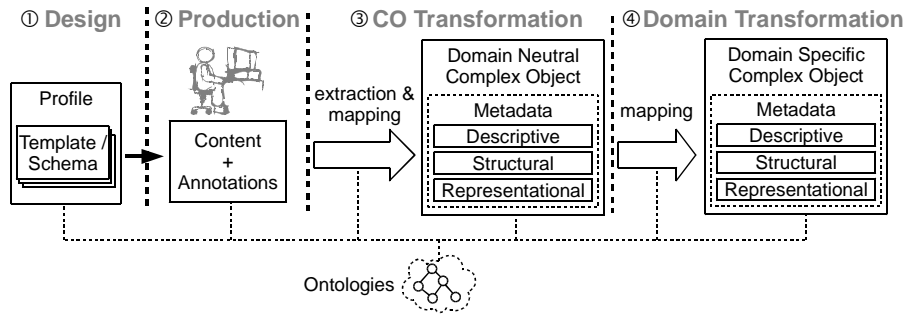


Figure 2. In Loco Semantics process diagram.

3. Document-centric Authoring

Even though in loco semantics can be applied to many content producing contexts, the most popular content producing tools available to end-users adopt documents as metaphors for content producing and as units of representation and distribution.

Modern word processors have the “style” feature, which enables to assign a name to a specific format configuration. Our *annotation pattern* extends its interpretation to our purposes. As shown in Figure 3, besides its formatting concerns, styles work as a semantic markup. Therefore, by marking fragments of documents using a {Title} style, for example, authors perform two operations: they apply a preconfigured format to the text and simultaneously attach a semantic annotation to the fragment.

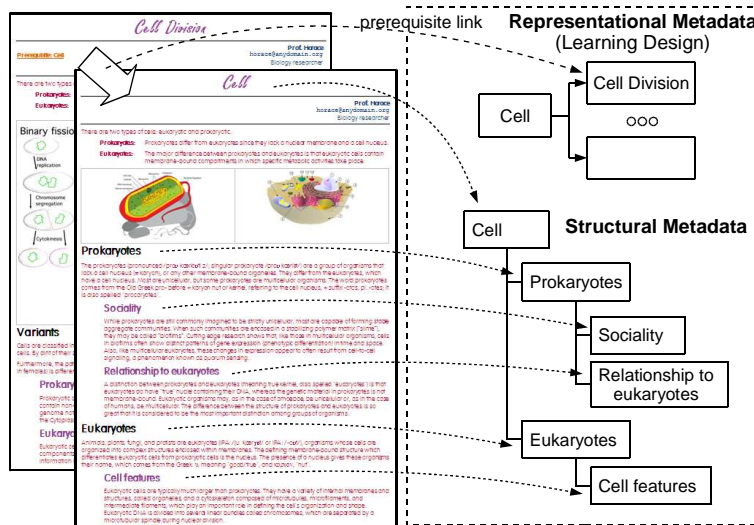


Figure 3. Converting document annotations in a LO comprising its learning design.

Figure 3 illustrates a summary index automatically extracted from e-learning content. In this case, the titles and subtitles can be straight converted to SCORM [ADL 2004] internal organization structure. Furthermore, the figure shows a link between the document entitled “Cell Division” and the one entitled “Cell” formatted using the {Prerequisite} style, indicating a prerequisite lesson. This example shows how annotation patterns are explored to infer the learning design behind the content following the IMS LD [Koper 2004].

4. Related Work

There are other solutions to produce LOs by using text processing tools, like GTK Komposer Suite (<http://www.gtkpress.com>) and Question Based Learning System (QBLs) [Dehors et al. 2006]. The main advantages of our solution lie in:

1. It is not limited to a monolithic tool: In Loco Semantics is a methodology and its tools are implementations which can be adapted and expanded.
2. Mappings from annotations to a LO structure are based on customizable XML open documents.
3. In Loco Semantics is not restricted to a text processing tool or a specific annotation pattern (e.g. style-based).

Concluding Remarks

This paper shows our document-centric approach to automatically produce LOs by means of In Loco Semantics methodology. It has been successfully applied at our University. We are expanding the methodology to other tools – e.g., spreadsheets and slide presentations have already been implemented – and other annotation patterns. Thus annotation patterns become a semantic communication strategy between humans and machines. Instead of requiring humans to learn ever increasing number of complex tools, our strategy proposes to enhance the language adopted in widespread and well known tools.

References

- ADL (2004). Sharable Content Object Reference Model (SCORM) 2004 – 2nd Edition – Overview, accessed on 11/2004.
- Dehors, S. et al. (2006). Reusing learning resources base on semantic web technologies. In: Proc. of the VI International Conference on Advanced Learning Technologies.
- Koper, R.; Olivier, B. (2004). Representing the Learning Design of Units of Learning. Educational Technology & Society, v. 7, p. 97-111.

André Santanchè

UNIFACS – Salvador University
Salvador – BA, Brazil
santanche@unifacs.br

Luiz Augusto Matos da Silva

Acre Federal University
Rio Branco – AC, Brazil
luiz.matos@ufac.br