Abstract— In the last years, authoring based on e-learning standards has been consolidated as a core factor of industry and development of interoperable and effective virtual learning environments. However, there is a need for further research on abstraction to provide a more instructional view in the context of authoring tools in a variety of ways, in order to avoid being driven by Learning Technology (LT) specifications, facilitate instructional knowledge aggregation, and to provide an appropriate level of clarity and semantics in the design of collaborative activities. We propose a combination of techniques to provide this instructional abstraction in the context of the new European educational model, combining instructional layers and collaborative scripts in authoring tools, and semantic web techniques for extending e-learning material in order to harness the wealth of existing web content and semantically labeled repositories.

Keywords: E-learning content modeling, Educational authoring, collaborative learning, e-learning standards, Semantic web

I. INTRODUCTION

The generation of educational content and design of collaborative activities has always been a big effort, especially because in many occasions the objective of the author or group of authors is to reuse existing resources and develop a complete course, including complex activities with content and user interactions. If the main goal of a learning object (LO) is to be used for teaching and learning, the second one in importance should be its reuse. To this end, extensive research has been carried out in the last years to standardize learning content components and collaborative interactions [14], to make them usable in interoperable and maintainable content repositories. To organize and help in the retrieval of the right LOs, metadata labels have been defined and standardized. But this has introduced an additional burden, namely that of annotating LOs appropriately following these metadata.

In this context, gathering educational content is a matter of two factors (a) reusability, which implies to have the resource at the appropriate level of granularity, and (b) availability, which tackles the idea of actually finding the most appropriate resource using a variety of techniques. The fact is that authoring tools usually do not implement features for these two factors and generally also lack of the appropriate abstraction level to provide an efficient way to search and retrieve content and on the other hand, a suitable way to describe learning courses at an instructional level of abstraction.

To overcome these problems in authoring of educational material we propose a combination of techniques to provide instructional abstraction by means of instructional layers, collaborative scripts in authoring tools and semantic web techniques for extending e-learning material in order to harness the wealth of existing web content and semantically labeled repositories.

II. INSTRUCTIONAL LAYERS, KNOWLEDGE MODELLING AND COLLABORATION

From the instructional point of view, the notion of Learning Object has been extrapolated from a variety of computational paradigms like reusable component as a software engineering concept, providing structured reusable elements labeled with metadata, and also from knowledge engineering, allowing content organization using knowledge-based structures like ontologies or semantic web development. On the other hand, from the cognitive sciences perspective, the adoption during the 50s and 60s of some instructional theories based on cognition have obtained useful abstractions to specify appropriate methods and situations in which those are to be applied during learning process [7].

In this scenario, very rich tools are available to tackle with the problem of providing flexibility in the creation of courses based in the aggregation of LOs [1][2]. However, this authoring tools have not evolved in a parallel way to be instructional aware, and still focus strongly on LT specifications and implement a LT specifications' syntax driven approach to implement the process of authoring [15]. The authoring of learning content has similarities with the COTS software building model, as it combines creation from scratch and reuse and modification of existing content, freely
available on the network or stored in repositories. In this case, the reusing process needs some meaningful way to search and retrieve the appropriate content shaped as learning objects (LOs). Current research agrees that approaches based on a plain classification of LOs by means of a metadata labels lack a meaningful way to search, retrieve and reuse LOs from an instructional perspective. The reason for this drawback is based on the fact that (1) from an instructional perspective, retrieval of a collection of LOs by matching metadata attributes does not replicate the way a teacher operates when creating its own material, and (2) from the authors perspective, the construction of metadata instances is very costly and not as precise and consistent as desirable.

A. Using knowledge modeling to improve authoring, collaboration and virtual learning communities

It has been mentioned that people is actually a killer application of the Internet, and it is with people and some organization that impressive results are achieved at low cost (for instance, del.icio.us, Wikipedia, Twitter, or the syndication of blogs) [31]. In this sense creating educational content can also be done by users providing the appropriate environment. Several repository efforts have been launched but they have not yet taken off.

![Figure 1: Semantic annotation of educational content](image)

Sometimes the problem is one of granularity: the presentation of the learning resources is too broad as to ease the reuse of specific content. But if the granularity is too small, it is more difficult to have a good overview. Concerning this point, we agree with the approach proposed by [19] of using Topic Maps. Actually, we have already explored that from the teachers' point of view, authoring of educational content is easier to create if it is previously instructionally structured based on a pedagogical ontology (See Figure 1). This way teachers can refer to LOs like i.e. ‘I would show here examples to illustrate this concept’ or ‘exercise this concept solving this problem’ or ‘give me a hint’. We can be even more precise by saying ‘insert here some easy examples to illustrate this concept’ and so on [2].

In the current way of retrieving LOs, we look for terms that match the metadata attributes, but without asking for a precise instructional relationship. So, the retrieved elements will have to be filtered one by one. We proposed to embed a semantic query language in LT specifications and the creation of a semantic layer based on the conceptualization of a subject domain. We think educational content authoring is a 2-step process. Firstly, there is a need to create an instructional view of the learning content at a low level of granularity. In this sense, some works suggest the use of semantic layers to organize information in the repositories [19]. Our approach also proposes using a semantic layer, except that we use RDF/OWL [24] [23]. There is a very strong relationship between RDF and Topic Maps. Simply stated, one could say that RDF provides a simpler and less specific model, based on more fundamental concepts. It is thus more flexible. Moreover, the possibility of reasoning with tools such as CWM or Pychinko, provides us with a tremendous power, allowing the inference of new information from existing metadata by the use of rules [18].

Some ideas are possible in this respect. As we have already evaluated authoring can be combined with topic maps as a way to link LT-based specifications to a cognitive-based framework [8], where LOs no longer remain as isolated components and act as a part of instructional ontologies. One way for improving LT-based specifications could be using embedded languages for allowing the creation of semantic instructional queries and referring instructional content at a higher level of abstraction [15] [6]. This authoring model would allow also the creation of dynamic links in the educational content, creating on the fly content in the case the reference had been updated, obtained by a direct reference or by means of inference of semantically labeled resources using reasoning link CWM and that is partially developed in [18] and [31]. The objective in this sense is not so much the development of new content specifications based on learning technologies or e-learning standards, but to use consolidated specifications such as IMS stack to integrate on them some mechanisms that allow the use of the described techniques with the objective to validate them in a real context of virtual campus. Going beyond reasoning, semantic web techniques have also a huge potential for enhancing eLearning search and retrieval, such as selecting, recommending, repurposing and reusing learning resources or tailoring learning activities to a given social or working context. Ontologies lie in the core of these techniques as they are able to provide a formal and shared conceptualization for every aspect within a learning set-up and scenario and to infer new and unforeseen facts implied by the available knowledge. Hence, we propose using ontologies as the foundation for the aforementioned semantic layer. Besides, using ontologies could help characterizing LOs in terms of semantic entities instead of the traditional syntactic approach. It has been signaled that retrieving LOs from a repository by searching field contents fails to catch the user's intentions, which tends to make its results either inaccurate or inadequate for the user [25]. Furthermore, as metadata standards impose a number of constraints on the possible values, learners and educational authors are supposed to know the right searching terms for a given query beforehand in order to get a good precision and
A number of techniques are going to be used to achieve this goal:

- Ontological Engineering, including locating candidate ontologies to reuse, aligning and merging them and building the lacking parts by using ontology design patterns such as those described by [26] for obtaining or building set of ontologies gathering all the necessary knowledge and representing every relevant instructional entities and relationships, making it explicit and easier to (re)use as well as binding the instructional ontologies to existing educational standards (such as SCORM, LOM or IMS-LD) for enabling the interpretation of existing metadata to bootstrap the semantic annotation process.

- Inference techniques, based on Description Logics [23] [28] and rule-based engines[30].

- Semantic annotation based on the representation and efficient retrieval of terms related to the aforementioned ontologies [29] [24]. As a result, it will be facilitated the reuse and repurposing of available Learning Objects (or their meaningful components) both in an Learning Object Repository and on the web.

- Clustering techniques for discovering new knowledge gathered from the web, such as FCA [27].

All these techniques are to be applied bearing in mind the pedagogical objective of designing a recommendation mechanism based on semantic search and able to cope with the actual needs of instructional designers and learners that would create a new kind of LO, which could be named recommendation LOs. Such a LO, would consist of contents collected from the web as the result of an usage-directed query as described before.

B. Modeling collaboration

The field of e-learning, collaborative learning is a pedagogical paradigm that is getting a growing acceptance. It is articulated by arranging a set of students into a number of small groups to carry out several learning activities together. Within these settings, learning objects, rather than being provided by teachers and instructional designers, arise as a consequence of the joint work of the groups supported by collaborative tools. These new kinds of objects have been coined as ELOs (Emergent Learning Objects).

The paradigmatic shift consists in moving from the teacher-centered perspective, where the instructor delivers the appropriate contents to the students, to the learner-centered one, where students take a more active role while teachers mediate and moderate the learning process. Articulating its instructional workflow requires storing and later retrieving ELOs. These objects have to be characterized with respect to a new dimension, that of the collaborative context whence the ELOs arise. And furthermore, the collaborative context where ELOs emerge can help automating their characterization in terms of context-dependent metadata. It is quite important to determine the elements to take into account to populate the context for facing these requirements. Examples of such elements would be the collaborative activity being developed, the group or the involved tools.

Socio-collaborative context elements are fed from the collaboration-oriented virtual learning environments (VLE), and particularly, from the use students make of them and their integrated tools. Hence, VLE must be built taking into account four main requirements:

- Defining the social structure for the underlying virtual learning community (VLC) where the learning scenarios are to be unfolded and particularly, the set of groups and types of users required for this development.

- Designing the collaborative work that learners will be involved in, including defining activities according to a pedagogical method, describing their sequencing along an instructional workflow, defining the roles suiting an appropriate division of labor, etc.

- Supporting the collaborative interaction by integrating an open set of external mediating tools within a VLE. The integration mechanism must provide a sufficient interoperability level for the users to perceive a smooth sense of continuity in their learning experiences. ELOs are, in this sense, a main resource as a vehicle to provide functional connection between tools an so achieving the desired interoperability level.

- Catching all the dynamical aspects inherently bound to the collaborative learning scenarios specification, such as, creating informal learning groups within the experiences lifespan, supporting activity monitoring and control or assisting the evaluation and grading tasks.

To face these requirements, we have developed the PELICAN e-learning platform [20] [21] [22]. Different users can perceive this tool from several complementary perspectives:

- PELICAN as a design tool. PELICAN is used by instructional designers as a design tool to define collaborative learning scenarios. The platform provides a simple and flexible modeling language to allow the definition of all the aspects related with collaborative experiences (social arrangements, collaborative work, monitoring rules, collaborative evaluation strategies, etc.) Notice that these specifications could be reified as LOs to foster their reusability along different pedagogical contexts and reduce the design efforts.
• PELICAN as a VLE environment. From the students’ point of view, PELICAN is perceived as a collaboration-oriented VLE. Hence, it provides the required technological infrastructure to support the previously designed learning scenario development such as shared virtual workspaces, access to web references, interaction services and so on.

• PELICAN as an integration platform. PELICAN can also be considered an integration platform as it provides several mechanisms to incorporate external tools supplying interaction services to virtual workspaces at various integration levels.

   ![Diagram of PELICAN architecture](image)

   Figure 2: Two working levels within PELICAN

According to these three perspectives the instructional workflow in PELICAN is stated at two working levels. As it can be seen in figure 2, within the design level instructional designers specify in a formal and computational way all the aspects related with the learning scenario being undertaken. At development level, teachers deploy the design of a learning scenario along one or several workspaces bound to different socio-collaborative contexts. Between these two layers PELICAN is located as an orchestration mechanism to synchronize instructional workflows within each workspace with the prescriptions expressed in the scenario specification at design level.

III. INSTITUTIONAL POLICIES: E-LEARNING STANDARDS AND EDUCATIONAL METHODOLOGIES IN THE CONTEXT OF EUROPEAN EDUCATIONAL SPACE

In the current arena of educational standards, the development of ad-hoc standards leads to a large extent forced the re-use of existing ones. Given the new production framework being proposed here, where educational institutions need to adapt the content life cycle towards a sustainable model, they have to combine their own research and development in the area of educational standards. Given this approach it would be possible to integrate lecturers into the authoring of standardized content via a strategy of institutional production using tools that implement instructional design in the abstract sense reflected in their VLEs, and thereby avoiding the use of syntax driven tools. The UNED has since 2006 being generating open content, specifically Open Educational Resources (or OERs), marked up using various educational standards. Specifically, given the accessibility needs of an important part of our students, these resources needed to include meta data not only on the type of content but also its structure and the way in which it can be adapted and presented to students with special needs.

A. Deploying e-learning standards policies

Distance learning in virtual environments allows intensive use of new technologies, especially in the field of creating and managing multimedia content. The use of multimedia resources, either as asynchronous or synchronous learning tool in virtual environments can improve content learning with a high visual and interactive effect. But multimedia source material is diverse, ranging from video lectures to radio broadcasting, videoconferencing and slide presentations developed in many different formats.

This availability of this highly heterogenous bunch of educational objects implies that a single standardized metadata schema does not conform, as a rule, the definition of an educational repository metadata architecture, as in the case of an organization like UNED. In this context, the development of an application profile specifically designed for a LO repository with large multimedia component is highly important, being necessary to correctly define a set of metadata associated with the nature of the object which will structure and identify all elements and relations between them. In this sense and, overall, the use of standards-based content description XML allows to describe completely and extending all the elements that are part of a course.

Using XML as a standard basic language for describing content is widespread used in e-learning standards based on LOM, DublinCore and others, as well as schemas for description of multimedia (MPEG-7 is an example) or the packaging of objects (eg. SCORM). Therefore, it is perfectly feasible to full integrate the profile as a standard inside an environment that fully uses the other, providing they have a common language and a consistent set of metadata.

This new application profile will result from the integration of different standardized metadata schemas (generic, such as Dublin Core, educational like LOM, and multimedia such as MPEG-7) and represents the first step of a long process to ensure consistency and reuse of these contents in the future. Taking LOM as a starting point it may be a good choice, or at least an appropriate option but will impyla greater effort (as far as labeling is concerned) in relation to other alternatives such as DublinCore. Moreover, at present, there are many more application profiles based on DublinCore than in LOM. This does not mean in any way, that the use of DublinCore is most appropriate. The reason is based mainly on one side in the simplicity of the Dublin Core metadata (as opposed to a larger number in LOM). For another, the more "generalist" DublinCore pursued against a better "educational" search with
LOM. In any case, as shown in this document and the referenced literature, it would be more difficult making the transition from a profile based on Dublin Core to a profile based on LOM than vice versa.

The subset of LOM metadata related to technique or technology will be somehow redefine to accommodate the huge volume and type of learning objects along with a set of other standard schemes which will define more precisely the audio and visual properties of the objects. In the case, the metadata that are commonly used in the LOM standard will be integrated in the data flow described by the MPEG-7, so that the latter becomes a comprehensive description of the reusable learning objects [32].

B. Software and content accessibility conformity

ICT technologies led each student to fulfill their learning process at their time, place, pace and capacities, but if the methodology is non-accessible for people with disabilities it is completely useless. Attention to diversity must be paid continuously, as the learning process is a continuous mode.

A critical issue to improve drastically the quality of the educational process is the development of high quality on-line learning resources and educational software [33] and the accessibility level of these materials will make a difference. A precise control must be taken into account over the creation process, analyzing the existing standards for learning resources, their level of accessibility, developing new laws, directives, standards, specific guidelines and new authoring tools. The appearance of Open Source development tools for SCORM compatible materials has achieved a great goal [34].

Non-accessible web interfaces prevents and hinder right access to information and services for Internet users. Even W3C (World Wide Web Consortium) promotes since many years accessibility standards and guidelines for both content development and authoring tools, reality is highly disappointing. The problem of accessibility is not yet understood by developers and producers and also there is still a lack of development frameworks covering the whole software development life-cycle that include accessibility checkpoints.

Lecturers and learners should be provided with specific means so that they can interact with learning material regardless of disability, benefits of accessibility compliance are not only for people with disabilities, but also for elderly people, all users in general. Learning environments will be Web based in most situations [35], therefore Web materials displayed into LCMS must be available for all users. The materials will run with the same behavior in all environments, have a consistent user interface and be easily navigated so that the content can easily be accessed and understood. IMS GDALA (Guidelines for Developing Accessible Learning Applications) offers specific guidelines for design and development of e-Learning applications in all the lifecycle and W3C/WAI WCAG (Web Content Accessibility Guidelines) give general guidelines to achieve accessible content.

Lecturers create learning content using above standards. Courses will be delivered and display through an e-Learning platform - that is in fact a Web based application - therefore greater flexibility and automatic processes are desirable from the point of view of the creator, commonly the lecturer. Authoring tools are the main point in the stages in the creation of educational content. The most widely used technical development so far is based on the use of validation and verification tools for web content after development. This validation cycle model is close related to the methodology called software prototyping based on the creation of prototypes, i.e., incomplete versions of the software program being developed. A prototype typically simulates only a few aspects of the features of the eventual program, and may be completely different from the eventual implementation.

The chances of the result of the validation process vary depending on the project and the available resources, in most cases it is reduced to the validation of the use of standards and accessibility and omitting the rest of the characteristics involved. Some authors propose methodologies based on an expanded user-centered development. In this type of solutions, starting from a simple design and considering all users, accessibility requirements are considered at first, before the initial prototype is developed, along with the rest of software requirements. Therefore in this model, the prototype is evaluated in the most possible ways obtaining then the most complete error report.

In another category are those solutions based on the development of web content using authoring tools that include accessibility standards like (X)HTML and CSS. At first they represent a guarantee of compliance with W3C standards and other requirements of accessibility and usability. Often these authoring tools also include validation tools or links to them, being designed to operate in a similar way as the validation cycle model. Such tools (as the ones that comply with W3C ATAG guidelines) are one of the best alternatives. In practice, it is a limited solution, having trouble with dynamic web pages or parts that are not included in the set of standards. As in the previous case, there is an intermediate solution being used only for the development of specific content based on (X)HTML or CSS.

In a new category one could gather all those new technologies arising from technical innovation around the Web 2.0 [36]: developments that use sets of Web 2.0 standards, Rich Internet Applications (RIA), Semantic Web or Micro formats. Actually, the application of these techniques usually takes place as a whole in the project so that it is analyzed through a unique hybrid model that represent simplified solutions. The use of Web 2.0 elements benefits accessibility:

- they are standardized elements already,
- there is the possibility of reutilization of models with RIA Mashup components,
- many metadata information exists to provide Semantic Web functionality.
Regardless of the use of RIA technologies, other authors defend in any context the use of semantic information in addition to accessibility aiming the usage of standards (X)HTML, CSS, XML for data and SOAP and DOM for data exchange and interaction. This concept is at the heart of Web 2.0, and thus is implicit in the use of these technologies. Alongside, he believes that Semantic Web compiles accessibility conformity of software components because of the metadata information also easier from RIA technologies because it has been taken into account almost since its inception. Finally, the semantic information is supplemented with the use of Microforms or metadata patterns that complement those elements likely to present accessibility problems.

Regarding the scope of this solution, it is beneficial because it provides innovative technologies for web development. RIA technologies certainly give greater power and functionality to the web, besides being really closed to accessibility conformity because of the use of web-based standards and reutilization of components. Still in the area, Mashup development ensures separation of structure and content, facilitating the publication of content with a minimum and accessible version from any application or Web service. On the other hand, the semantic information through metadata is a very valuable partner to enriched components, so that that they can be interpreted by assistive technology tools and accessibility APIs for platforms.

C. The European Educational Area

In a broader context, the new European Area [9] and its convergence in education designed a model closer to what today is conducted in North America and Japan. In such systems is given greater importance to the practice load during the conduct of a subject. By providing an orientation toward more experimental tasks, and a clear direction to the working world, students develop a range of skills than in degrees with less experimentation do not have. This is an excellent testbed to carry out institutional standardization policies.

The most visible set of changes involves the abovementioned adoption of a US-like unified cycle structure involving graduate-master-doctoral cycles, as well as the adoption of a single unit of measurement, the ECTS (European Credit Transfer Systems) credit (which refers to 25-30 student hours of total effort, rather than being measured in hours of face-to-face lessons as before). In many countries (such as Spain), this involves the re-design and thus the (re)accreditation of all the degrees, under the quality certification system required by the EHEA. This massive, simultaneous redesign of all degrees presents daunting challenges but also offers unprecedented opportunities. On the one hand, since all degrees must be simultaneously redesigned, synergies among them can be effectively exploited, thus encouraging the re-utilization oriented approaches discussed in this paper (LCMS, standards like LOM, Dublin Core, IMS, SCORM, etc.). On the other hand, shifting the unit of academic measurement to student hours (through the ECTS) facilitates the seamless combination of face-to-face, distance and blended learning in academic degrees.

The other, maybe even more significant but more subtle set of changes are those aimed at shifting the focus from instructor-centered “teaching” to student-centered “active learning”. It involves methodological changes such as continuous evaluation, de-emphasizing theoretical lectures to focus more on assignments and projects, higher practical focus, allowing students higher flexibility to design their own curricula. When combined with budget limitations, this methodological shift strongly supports the introduction of effective IT based approaches to alleviate the burden on the instructor’s resources. These should facilitate the educational equivalent of the current manufacturing trend towards “mass-customization”, thus allowing individually tailored learning paths with a level of resources similar to that required by standardized education. In addition, several countries are taking this opportunity to introduce far-reaching modifications in their educational systems, which further strengthen the case for the introduction of IT based educational innovation. For example, in Spain, until now, all “official” degrees were listed in a catalogue compiled by the Education ministry (universities could also grant their own degrees on whatever they wanted, but those did not have official recognition). This catalogue included the name and the degree curriculum (structure), up to certain level of detail. The new system, however, breaks away from that closed catalogue approach and just issues some very generic guidelines to which new degrees should conform.

Within this framework, universities (both private and public) are free to propose whichever degree titles and supporting curricula they want. Once the proposal is cleared from a quality criteria point of view (general quality criteria, such as the faculty CVs, cohesiveness of the proposed degree curriculum and appropriateness of the supporting IT infrastructure) the new degree is inscribed in a national registry, and the university is free to offer it (subject, again, to periodic quality evaluations).

One last aspect worth highlighting regarding the EHEA is its emphasis on promoting mobility and the international dimension in education (through joint international degrees or through mobility in selected subjects of end term Thesis). Again, achieving this objective would be assisted by the adoption of standards-based, location independent IT-based educational solutions. These should support both distributed provision of learning services (e.g. in degrees offered by contending universities) and their consumption by distributed student groups, facilitating not just the interaction between students and instructors, but also the increasingly critical interaction among participants in distributed teams.

IV. BACKGROUND OF UNED RESEARCH GROUPS

Researchers in the UNED belong to LTCS¹ (LSI Dept.), G-Elpis (DIEEC² Dept.) or ATLAS² (Modern Languages Dept.) groups. Some of their members also belong to the CINDETEC unit, a Vice-chancellorship that coordinates the university's e-learning infrastructure.
The LTCS Group is made up of researchers from the LSI Department at UNED and external collaborators that work developing projects which apply learning and collaborative technologies to the support of human activity in distance learning scenarios. The research is based upon:

- Knowledge-based Authoring
- Collaborative modeling in educational contexts
- E-learning standards
- Software and Content Accessibility conformity

The group's research lines can be situated within the area of educational technologies, specifically learning, both individual and collaborative. The first relates to different aspects of instructional knowledge representation and its normalization. The second is centered in the problems associated with collaborative learning and the mechanisms that sustain and define the activities that configure it. The third goes from the description of the collaborative workspaces or mediational tools that facilitate the communication between direct agents in the process, to the analysis of the process based upon records of the activities undertaken. The latter reinforces accessibility issues while delivering eLearning services.


As well as working in different research projects, the group participates in other activities that promote investigation and its diffusion, taking part in summer courses, doctoral courses, and specialised seminars and congresses at both national and international levels.

Finally the ATLAS7 group has been working in the area of intelligent language learning systems for several years and has been involved in a series of funded research projects: The Virtual Verb Trainer (VVT), The Virtual Authoring Tool (VAT), I-Peter I, I-Peter II, COPPER and I-AGENT, the latest and currently ongoing project with a Ministry of Education grant, number: FFII2008-06030.

ACKNOWLEDGMENT

We thank Comunidad de Madrid for the support of E-Madrid Network of Excellence S2009 TIC-1650 and Ministerio de Ciencia e Innovación for the support of project TIN-2009-14317-C03-03.

REFERENCES


1 http://enlace.uned.es
2 http://www.coldex.info
3 http://senseilsi.uned.es/ea2c2
5 http://www.eadtu.nl/e-xcellenceqs/
6 http://www.edutubeplus.info
7 http://www.ieec.uned.es
8 http://atlas.uned.es


