Web resources for Computer-Aided Instruction: a blended classification scheme

Pacifico Cofrancesco[1], Mario Petrone[2], Filippo Bruni[2], Elena Caldirola[3]

[1] Università degli Studi di Pavia, Unità CSGI
[2] Università degli Studi del Molise
[3] Università degli Studi di Pavia, CELID

The huge quantity of data, media, applications, and services - in one word, resources - that are accumulating day after day on the Web makes it more and more difficult to search the network in an effective and helpful way. We usually spend a lot of the time trying to "filter out" what we consider "noise" added to the "good information" for which we are looking. Whatever the search domain, this messy and discouraging situation cannot be handled by general search engine, such as Google. We also face similar problems in the instructional domain. In this paper we focus on the need for creating tools that can help classifying and retrieving digital Web resources for Computer-Aided Instruction processes. We propose a strong, but simple and flexible, classification scheme, which can be easily and profitably used by a Web community (of teachers, learners, and others) to create a database of references to digital instructional resources. Our classification scheme uses a blended top-down/bottom-up approach with a Delicious-like annotation

for citations:
Cofrancesco P., Petrone M., Bruni F., Caldirola E. (2011), *Web resources for Computer-Aided Instruction: a blended classification scheme,* Journal of e-Learning and Knowledge Society, English Edition, v.7, n.3, 87-96. ISSN: 1826-6223, e-ISSN:1971-8829

and tagging system not fully free, but based upon a controlled and predefined and expandable set of metadata.

1 Introduction

Many digital instructional resources are made available every day on the Web - very often for free - but it is very difficult to know where these resources are, and how to retrieve and use them. That's the reason why the re-use of those resources is not very popular, and most people are starting the job from scratch every time.

What we need is a dedicated system that can effectively help teachers, learners, and others involved in the instructional process to find and use the most appropriate resources to reach his/her didactical and learning objectives.

The first step in our proposal is to define a general set of metadata to classify instructional resources, that must be as close as possible to the need of target users, and should also allow semantic annotations of the resources the user is interested in and finds on the Web. This classification scheme has to be simple, effective, and open to users' free annotations, in a Delicious-like style.

It has been shown (Sommaruga *et al.*, 2010) that a synergic combination of the two driving forces of the present Web evolution, that are Web 2.0 (or Social Web) (O'Reilly, 2005) and Semantic Web (Berners-Lee *et al.*, 2001), can overcome their respective limitations. In fact these two approaches are complementary, and a good basic taxonomy can be the framework in which users' defined folksonomies (i.e., users' free tagging) are better "organized" (Carcillo *et al.*, 2007). On the other hand, folksonomies reduce the excessive rigidity of taxonomies, and are very useful to choose terms and values of controlled vocabularies, as shown in the BBC website example (Loasby, 2006), in which when a tag is widely used as a resource annotation it can be promoted to a standard metadatum value. In a blended taxonomy/folksonomy system, resources can be classified and searched by using both metadata and tags, taking advantage of both these different approaches, and letting the user free to choose the method closer to his/her own mental attitude, but giving him/her the possibility to use taxonomy's metadata and folksonomy's tags at any time.

2 The context: the instructional process

Our classification scheme of Web resources for Computer-Aided Instruction is based upon the classical definition of an instructional process (Dick & Carey, 2008).

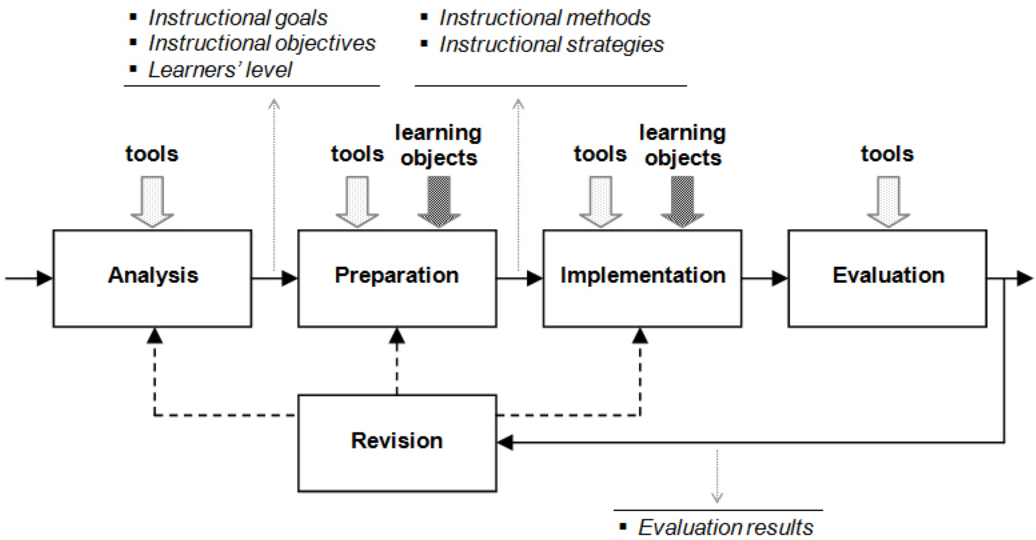

Fig. 1 – Basic block diagram of the instructional process

This is a sequence of steps (Fig. 1) starting from the analysis of learners' entry level, and the identification and definition of the learning goals and objectives, going through the *preparation* and *implementation* of *instructional methods*, *strategies*, and *materials*, and ending with the evaluation of the results of the instructional activity. As in any feedback system, a *revision* step follows, to introduce some adjustments to any of the previous steps strategies and choices, if results are not satisfactory.

In every step, several tools and learning objects, in general called instructional resources, can be used. This scheme of the instructional process is a general one, and it can be applied to a traditional instructional environment, or to a Computer-Aided Instructional (CAI) process. In this last case, tools and/or learning objects have a digital nature, so they are *digital instructional resources*.

As shown in the figure above, we have two types of digital resources, i.e. learning objects or digital learning materials, which are informational resources, and tools (hardware, applications, plans, etc.), which support the activities done at each step of the instructional process - traditional or computer-assisted.

3 Metadata: the taxonomy layer

Instructional digital resources can be classified from many different points of view. If we want to make the search of classified resources less frustrating and more effective, we need to make available descriptions (i.e. annotations) of the same resource from several different points of view, which correspond to several users' views of the same digital object.

Our classification scheme has been created after an accurate analysis of the existing proposals (such as Dublin Core, IEEE Learning Object Metadata, IMS, and existing repositories), having in mind, as fundamental requirements, essentiality and simplicity of both use and application. Some work has been done recently to define more appropriate sets of metadata for instructional resources, in particular for learning objects (Alvino *et al.*, 2007; Alvino *et al*., 2009). In our experience we found that too technical vocabularies or a too strict system of classification discourages users who are supposed to bookmark those resources. That's why we chose to define a "minimal" set of metadata, which makes the job of expanding it easier. On the other hand, special attention has been paid to the proposed values of each metadatum in the set, because we want to give the users precise and meaningful definitions of metadata.

In our classification scheme we have three different sub-sets of metadata:

1. identifying the resources
2. characterizing the resources from an instructional point of view
3. characterizing the resources from a technical point of view - because of their digital nature.

The proposed classification scheme does not pretend to be exhaustive or the most appropriate for instructional digital resources classification and retrieval. That's why we keep it "open" and expandable by free user tagging. It's not really a hierarchical scheme, every metadatum has the same importance, and it can be considered as a "different point of view" about the same resource. That is one of the many possible ways to describe that resource. Metadata can be thought as "filters" that give the possibility to other users to find resources in the Web. New categories of tags freely added by users can extend the metadata set, adding new "points of view", not considered before. The basic taxonomy is modified and updated by users' folksonomies, when needed. This is very important if we want to keep the system alive, and able to evolve as users needs and trends change.

3.1 Resources identification

A resource needs to be identified and some essential information have to be recorded about *creator*, *date* of creation, *location* on the Web, *language*, *copyrights*, and *cost*, as listed in the following table (also defined in DC and IEEE LOM).

TABLE 1

Metadata	Values
title	<tool/content title>
creator	<author or creator>
publisher	<publisher>
date	<date of creation>
location	<resource location>
language	<resource language as defined in RFC1766 ("en", "en-GB", "fr-CA", "it")>
rights	<copyright statement>
cost	{freeware, open source, commercial, other}

3.2 Instructional characterization

Many different learning theories and models can help defining metadata from instructional characterization of digital resources. Some choices and selections have to be made at this level, which eventually could be not the best one for every user. New metadata can be added, whenever they are needed, to take into account other cognitive theories or points of view.

When we start searching the Web for digital resources to be used in our instructional process, we have in mind a *topic*, a *content area*, some characteristics of our *target audience*, and the *skill* and *grade levels* of the learners. We will refer to specific instructional methods and learning theories. For example, in terms of Bloom's Taxonomy (Bloom & Krathwohl, 1956), we could search digital resources to foster learners' ability to remember (knowledge cognitive dimension) or some other to improve critical thinking (evaluation cognitive dimension). Here we explicitly refer to Bloom's Taxonomy as revised by Anderson and others (Anderson *et al.*, 2001), introducing the so-called knowledge dimension. Bloom's Taxonomy and its revision seem to be a good classification scheme for instructional digital resources, because very often instructional objectives and goals are the starting points of the search activities of such resources on the Web.

In the following table, the metadata we define for the instructional characterization of Web resources are summarized.

TABLE 2

Metadata	Values
resource type	{tool, content}
granularity	{curriculum, course, lesson, page, media}
instructional step	{analysis, preparation, implementation, evaluation, revision}
user role	{producer (course authoring, web site authoring, testing and assessment, media editor, content converter), learner (content browsing and playing), engineer (content hosting and management)}
topic	<specific topic title>
grade level	{preschool, kindergarten, primary school, secondary school, college, adult (learning and training), disabled persons}
interaction type	{teacher/learner(s), learner/computer, learner(s)/learner(s)}
target audience	{whole class, small group, couple, individual}
instructional method	{direct instruction, interactive instruction, indirect instruction, independent study, experimental learning}
cognitive dimension	{knowledge (remembering), comprehension (understanding), application (applying), analysis (analyzing), evaluation (evaluating), synthesis (creation)}
knowledge dimension	{factual, conceptual, procedural, metacognitive}
content/tool area	{art, astronomy, biology, mathematics, science, technology, history, social science, foreign languages, other}
content/tool type	{assessment, narration/description (lecture, presentation, exhibit, story telling), reference, construction, demonstration (tutorial), discussion (forum, small work group), simulation (role playing, instructional games, field trips), illustration, imagery, modeling, brainstorming, problem solving, case studies, drill and practice (apprenticeship), generative development, research project, web quest, expert system, evaluation, map, portfolio, platforms, documentation, communication, sharing tool, other}

3.3 Technical characterization

This is a very important characterization for digital resources, because we need to know which requirements are to be satisfied to be able to use them. In fact they have a digital format, and some hardware and software requirements. It's also important to know which technical background the user is supposed to have to use them, and also the level of technological and management complexity (Petrone *et al.*, 2010).

TABLE 3

Metadata	Values
digital format	<data type of the resource, see mime formats>
hardware requirements	<hardware requirements: processor type and speed, memory, display size and colors, hard disk size, CD or DVD units, audio output, audio input, video input>
software requirements	<software requirements: operating system, browser, media players (plugins), Java VM, etc.>
connection requirements	<connection requirements: connection type, connection speed>
required knowledge	{low, medium, high}
usage complexity	{low, medium, high}
technological complexity	{low, medium, high}
management complexity	{low, medium, high}

4 Users' tagging: the folksonomy layer

As shown in many papers, fully free tagging is not precise in terms of categorising and language, and tags also have no hierarchical structure, so key duplication and spelling mistakes are frequent (Carcillo *et al*., 2007). To avoid some of those problems, in our blended classification system users are free to add new tags, but the system will propose some categories in which to put them, and will give them the possibility to access the standard vocabulary associated with specific metadata or newly added user's keywords and tags, and then to make a choice among the existing keywords.

Some post-processing is implemented to get information about popular tags to be promoted to the metadata levels (Benz *et al*., 2010). So, users of the system can create themselves the more suited classification for the resources they find on the Web, which emerges quite spontaneously as a kind of "natural selection" process. In this way, users of Web 2.0 will contribute to the Web by not only adding contents to the Web but also adding a kind of "semantic" classification of other contents already available on the Web.

5 System implementation

Once defined the starting metadata set and their associated values, we create the structure of a Web repository for the references to the instructional digital resources of which users would keep trace. It is a “referatory” (database of reference) and also a true repository of resources, where users can add copies of their own learning objects or tools, with all the information needed by other users to re-use them.

The database is populated by registered users, by means of an almost free tagging of resources, in the framework of the defined basic taxonomy.

Some dedicated and intelligent tools (“smart agents”) help in browsing the database, and suggesting the most appropriate tools and materials for the various steps of the instructional process, according to users’ requests.

As said before, in the “back office” of our system, some tools analyse users’ freely added tags, and modify and update the basic taxonomy, according to the defined maintenance policies.

Conclusions and developments

The main idea of our proposal is to introduce a method to classify Web resources which reduces the rigidity of a traditional taxonomy - in a Dublin Core or IEEE-LOM style - by means of flexible, but controlled, users’ tagging of Web resources - in Delicious style. We are creating a “recommender” system that improves the possibilities of retrieving and re-using the many digital instructional materials and tools available on the Web, in a “smarter” and more effective way. A community of users feeds the database and gives new directions on how better to define the classification system itself, taking advantage of the new communication paradigms and interaction strategies of Web 2.0, together with the precision of Semantic Web.

The system will be used at the University of Pavia, Italy, in the framework of the e-learning Kiro-Maieuta project (Kiro-Maieuta), and also in some secondary schools of the Molise Region, Italy, in order to evaluate the impact and the efficacy of this approach in different instructional environments. The results of this evaluation will be described in a following paper.

REFERENCES

Alvino S., Forcheri P., Ierardi M. G., Sarti L. (2007), *Un modello per la connotazione pedagogica di LO*, Rapporto tecnico IMATI-CNR n. 13/2007.

Alvino S., Forcheri P., Ierardi M. G., Sarti L. (2009), *The pedagogical dimension of metadata for learning resources: the POEM model*, in: Spaniol M., et al. (eds.), Advances in web Based Learning – ICWL 2009, Proceedings of ICWL 2009. 58-67, Berlin/Heidelberg, Springer.

Anderson L. W., David R. Krathwohl D. R. (eds.) (2001), *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, Boston, Allyn & Bacon.

Bell S. J., Shank J. D. (2007), *Academic Librarianship by Design: A Blended Librarian's Guide to the Tools and Technologies*, Chicago, American Library Association.

Benz D., Hotho A., Stützer S. (2010), *Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge*, Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.

Berners-Lee, T., Hendler J., Lassila O. (2001), *The Semantic Web*, Scientific American, May 2001.

Bloom B. S., Krathwohl D. R. (1956), *Taxonomy of educational objectives: The classification of educational goals*. Handbook 1: Cognitive domain, New York, Longman.

Carcillo F., Rosati L. (2007), *Tags for Citizens: Integrating Top-Down and Bottom-Up Classification in the Turin Municipality Website*, in "Online Communities and Social Computing: Second International Conference", OCSC 2007, Held as Part of HCI International 2007 (Beijing, China, July 22-27, 2007), ed. Douglas Schuler, Proceedings Series: Lecture Notes in Computer Science, Vol. 4564, Springer, pp. 256-264.

Dick W., Carey L., Carey J. O. (2008), *The Systematic Design of Instruction*, (7th Edition), Boston, Allyn & Bacon.

DCMI Metadata Terms (2010), http://dublincore.org/documents/2010/10/11/dcmi-terms/ (verified June 27, 2010).

IEEE Standard for Learning Object Metadata (1484.12.1) (2002).

http://ltsc.ieee.org/wg12/par1484-12-1.html (verified on December 2010).

IMS Learning Resource Meta-Data Best Practice and Implementation Guide Version 1.2.1 Final Specification, http://www.imsglobal.org/metadata/imsmdv1p2p1/imsmd_bestv1p2p1.html (verified on December 2010).

Kiro-Maieuta, http://lotarionline.unipv.it/moodle/mod/resource/view.php?id=484 (verified on August 29, 2011).

Loasby K. (2006), *Changing approaches to metadata at bbc.co.uk: From chaos to control and then letting go again*. Bulletin of the American Society for Information Science and Technology, 33: 25–26.

O'Reilly T. (2005), *What Is Web 2.0*, Design Patterns and Business Models for the Next Generation of Software.

http://oreilly.com/web2/archive/what-is-web-20.html (verified on June 27, 2011).

Petrone M., Pasquariello E., Valente R., Brunale M. (2010), *Learning Methodologies and their Software Tools*, in: Proceedings of CSEDU 2010 - 2nd International Conference on Computer Supported Education, Valencia, Spain, April 7 - 10, 2010.

Shareable Content Object Reference Model (SCORM), Specifications, http://www.

adlnet.gov/Technologies/scorm/default.aspx (verified on December 2010).
Sommaruga L., Catenazzi N., Canetta L. (2010), *The Intelligent web: tools for knowledge management and sharing*, Proceedings of 16th International Conference on Concurrent Enterprising (ICE 2010), Collaborative Environments for Sustainable Innovation, Lugano - Switzerland, 21-23 June 2010.
Wiley D. A., (2002), *The Instructional Use of Learning Objects*, Agency for Instructional Technology. Online version here: http://www.reusability.org/read/ (verified on December 2010).