Publicado en Actas de la Conferencia: *Current research in Information Sciences and Technologies: Multidisciplinary approaches to Global Information Systems. InSciT2006.* Badajoz: Open Institute of Knowledge, 2006, pp. 309-314

Metadata and ontologies for organizing students' memories and learning: standards and convergence models for context awareness

Eva M. Méndez[a] and Jane Greenberg[b]
[a] *University Carlos III of Madrid, Librarianship and Information Science Department, C/ Madrid, 128, Getafe (Madrid), SPAIN. (emendez@bib.uc3m.es)*
[b] *School of Information and Library Sciences, University of North Carolina at Chapel Hill, CB #3360, 205 Manning Hall, Chapel Hill, NC 27599-3360, USA. (janeg@ils.unc.edu)*

This paper focuses on ontologies supporting context awareness and Personal Information Management (PIM) and their applicability in Memex Metadata (M^2) project. M^2 is a research project of the University of North Carolina at Chapel Hill to improve student digital memories using the tablet PC, Microsoft's SenseCam technology, and other mobile technologies (e.g., a GPS device) to capture context. The M^2 project offers new opportunities studying students' learning with digital technologies. This paper introduces the M^2 project; discusses E-portfolios and current educational trends related to pervasive computing; reviews relevant ontologies and their relationship to the projects' CAF (context awareness framework), and concludes by identifying future research directions.

Keywords: Context Awareness, Ontologies, Metadata, e-Portfolios, Pervasive Technologies, Learning evaluation systems, Higher Education, Standardization.

1 INTRODUCTION

New models in higher education are more learner-centered than teacher/educator-centered. Although North American universities and colleges appear to have a long tradition on student-centered education, new European Higher Education Area (EHEA) together with the ECTS (European Credit Transfer System) [6] are make student-center approaches more universal. These developments support learning processes that focus on learning outcomes. Information and Communication Technologies (ICT) play a key new role in addressing new educational and have the capacity to improve education and to promote *life long learning.* This new model stimulates a wide variety of ICT research, both in devices development and use, and in standards application.

The development of Teaching and Learning technologies and the rapid growth of e-Learning environments require a set of student-centered information systems that need new methods and model for organizing, processing and retrieving information—authored or collected by the student. All together, this forms a new personal information environment (learner's information or learner's e-portfolios) dependent on user-centered file management options [2]. Discipline specific information and needs, however, require ontologies and metadata sets that can be used in personal information systems and facilitate interoperability with other personal information systems or shared information system. In short, robust personal information systems are needed to support individual educational needs, but that can interoperate with other more formal and shared systems.

The need for robust personal information systems in the e-learning environment is complicated by the wide variety of information bearing objects that students produce (e.g. lectures, whiteboard notes, slides, handouts, Web pages, emails, personal notes, assignments, etc.) in digital formats of different nature (pictures, texts, sounds in their iPods, audio files recorded with the mobile, etc.) organized in e-portfolios or webfolios. Descriptive information about information (metadata) is a core issue when managing digital resources. Metadata management gets more complex as the variety and complexity of information objects grows. E-portfolio management needs metainformation at different levels, ranging from simple descriptions and metadata schemas to advanced metadata content-oriented schemes (ontologies), with semantics to express relations and inferences between and within those digital memories. To this end, we will be able to have intelligent student portfolios that take advantage of new technologies, and facilitate the learning process

2 STUDENT MEMORIES AND THE M^2 PROJECT

Digital memory involves the use of digital technology to record the experiences, activities, or events in which one is involved. Digital recording is known to be more accurate and long lasting than the memories a person stores in their brain. Digital technologies used for memory present a host of new challenges, particularly as an individual's digital information store evolves and grow in size, topicality and complexity. One of the most pressing challenges is retrieving digital memories a person has stored. Studies have shown that people rely on contextual cues to retrieve items based on partial information [8].

The Memex Metadata (M^2) for Personal Educational Portfolios project (hereafter, M^2) at the University of North Carolina is exploring this important research question. M^2 is a partnership between the University of North Carolina at Chapel Hill (UNC/CH) School of Information and Library Science (SILS), Metadata Research Center (MRC), Information Technology Services (ITS), and the Department of Biology. M^2 was launched as part of Microsoft Research's Digital Memories (Memex) project[1], testing the utility of SenseCam technology and MyLifeBits software. A main goal of M^2 is to build a contextual retrieval environment for educational information recorded via the use pervasive technologies. Metadata and ontologies are a key component of the M^2 project and the utilization of the Memex Research kit, which include SenseCam technology (image capture technology that records images approximately every 90 seconds, or when different activities are motions are detected via sensors), and MyLifeBits (MLB) PIM software. The Microsoft Research has been initiative inspired by Vannevaur Bush's[2] notion of the Memex, and includes mobile devices [12] (e.g., annotating personal information with contextual descriptive metadata such as subject, date, and time) and University of North Carolina is exploring the Memex notion within the academy to see if it can help improve student learning

M^2 work has included the development of a modular metadata schema that we would like to integrate with the MLB Memex software so as to enable more detailed contextual metadata. MLB currently supports annotation of captured personal educational information with and other basic metadata elements, such as the date a resource was created or a file type. The annotation feature allows personal educational information to be tagged with contextual information and facilitates retrieval in a rudimentary way, but MLB requires and is planning to implement more robust metadata functionalities. To focus our research, we are working with the Biology Department and undergraduates engaged fieldwork learning about plant identification and scientific taxonomy, although our methods are applicable to different disciplines where field activities are present (e.g., Anthropology. The remainder of this section discusses 1) E-Portfolios and student's personal information; 2) Pervasive Computing and context awareness; and 3) Metadata and Ontologies for reflective learning.

2.1 E-Portfolios, memories and student's personal information

In today's technologically rich environment, personal educational information includes a wide array of information with different origins, formats and scopes. The wide array of personal information is digital in pure e-Learning context, but it is increasingly electronic in traditional learning environments where education increasingly depends on technology, and is experimenting with virtual approaches throughout the term. Among they type of digital resources a student may have are:

- Information disseminated in class sessions, such as lectures, lesson plans, whiteboard notes, slides, handouts, Web pages, etc. stored in different profiles of different C/LMS (Course/Learning Management Systems).
- Personal notes created both in traditional notebooks and in electronic format, through laptops, tablet PCs and other computer devices that the students could have.
- Work products, such as quizzes, papers with multiple versions and revisions, evaluations, and Instant Message conversations with classmates.

Researchers and academic institutions are beginning to develop fully digital personal educational portfolios, containing students' documentation, with usually integrated access or links to "official" course materials. E-portfolios, ePortfolios or Web-portfolios are becoming a key component of new educational models[3] [4]. According to one of the most cited e-portfolio definition (the EDUCASE National Learning Infrastructure Initiative [14]) electronic portfolio is *a collection of authentic and diverse evidence, drawn from a larger archive, that represents what a person or organization has learned over time, on which the person or organization has reflected* […].They offer advantages for student reflection, collaboration and skill

[1] Microsoft Research Digital Memories (Memex): http://research.microsoft.com/ur/us/fundingopps/RFPs/DigitalMemories_Memex_RFP_Awards.aspx

[2] Vannevar Bush. As We May Think. *The Atlantic Monthly*, July 1945. A digital version of this landmark article could be found at: <http://www.theatlantic.com/doc/194507/bush> or <http://ccat.sas.upenn.edu/~jod/texts/vannevar.bush.html>

[3] Even the Eife-L (European Institute for E-Learning) has created a consortium called EuroPortfolio <http://www.europortfolio.org/> which organizes a specific conference on ePortfolios to study standardization and interoperability between different e-portfolio systems (See: http://www.eife-l.org/news/ep2006).

assessment and serve as a rich content personal information profiles useful both for learners and teachers. There number of available e-portfolios systems parallel the number and diversity of students and learners. Additionally, there are extensive amounts of educational information created and disseminated outside of the formal digital workflows —for example, class discussions, notes and conversations in study groups, brainstorming on whiteboards, or images and artifacts from field trips. This is where Vannevar Bush's notion of the Memex is very applicable and serves as a key motivator for the M^2 project. That is, the idea to capture much of this "informal" information, and to use metadata for contextual retrieval of students' memories, deepening and enriching not only their portfolios but their education and their consciousness about learning and knowledge.

2.2 Pervasive computing and Context Awareness

In the last five years we have seen the expedient acceptance of different kinds of ICT, is nearly all sectors of society. Pervasive technologies, such as cell phones and digital cameras were first called "new" information technologies, the ubiquity of ITC in our society. Now "new" technologies are the catalyst of a trend called *pervasive computing*, which is the natural extension of the existing computing paradigm, where computers, computer devices and information systems build on them will seamlessly integrate into the life of everyday users, providing them with services and information "anywhere-anytime" [3]. Pervasive computing encompasses all kind of handheld devices such us laptops, tablet PCs, PDA, mobile phones, etc. M^2 is experimenting with pervasive technologies in addition to SenseCams technology (Fig. 1), a memory technology developed by Microsoft Research that records images of your whereabouts every 90 seconds or more frequently when it senses changes via sensors.

Fig. 1. SenseCam connected to a laptop (M^2).

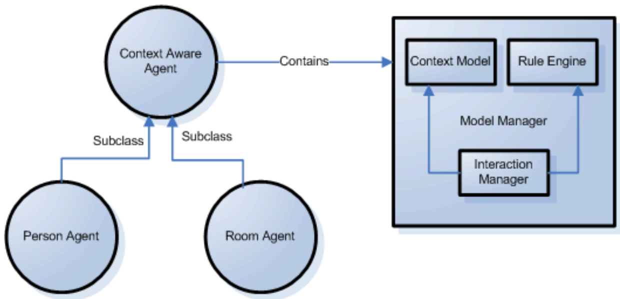

Fig. 2. Content Awareness Framework (M^2).

Context awareness is increasingly being explored in relation to ubiquitous and wearable technologies [7]. Context awareness can apply to SenseCams technology, which records information about the circumstances under which this device operates and records information, making assumptions about the user's current situation. Mobile technologies are becoming more embedded, ubiquitous and networked, with enhanced capabilities for social interactions, context awareness and internet connectivity [16]. Context awareness increases the potential of pervasive technologies based on devices mobility [9, 10, 11]. Mobile devices are provide new options for context aware applications because they are available in different locations, and so can draw upon different context and locations to enhance learning activity. M^2 is integrated with UNC's Context Awareness Framework (CAF) developed by ITS. UNC's CAF has two key features (Fig. 2):

- An integrated framework with different ontologies applicable to a range of objects (people, classes, events, etc.) active in the University's pervasive computing environment.
- Software agents that will be installed on students' computers to communicate with CAF ontologies and other rules sets specific to the university environment.

Context aware technologies and applications are visible in museums, and have been used for some time [10]. Such applications provide additional information about exhibits based on the user location. As universities and other higher educational institutions increasingly require students to own/have laptops, and course work becomes increasingly digital, it makes sense that researchers explore the current value and potential of context aware technologies. Thus contextual retrieval, based on the metadata and ontologies is particularly promising to capture and retrieve learning experiences from students' e-Portfolios in our M^2 project.

2.3 Metadata and Ontologies for M^2

Metadata and ontologies are core components of the M^2. These standards are required to improve learners' portfolios and learning activities. Both, ePortfolios and context awareness technologies are based on metadata [18, 17, 9, 10, 11]. We also need to consider shared ontologies to provide for an interoperable environment. The next section of our paper addresses metadata and ontologies supporting the M^2 research and development.

3 STANDARDS AND MODELS FOR CONTEXT AWARENESS APPLIED IN HIGHER EDUCATION

A recent review of context awareness research [9] identifies two main aspects of the learner's context: 1) the learners "setting" (including physical location, objects and people in close proximity, and available resources), and 2) the "learner" themselves (including their current activities, goals, and learner profile. On M^2 we identify three kinds of metadata that cut across these two aspects:

- Content metadata: those descriptive metadata intended to point out, for example the subject of the SenseCam images, documents, emails, photos, and audio objects[4] related to the topic of the class.
- Structural metadata: the format and architectural composition of the SenseCam images, documents, emails, photos, and audio objects (e.g., gif or tiff files, mp3, etc.).
- Context metadata: When (date and time) and where (location and address) the objects are produced, manipulated, and used.

These context metadata should be, by definition: automatic, standardized, and based on an ontology defining learning scenarios. Current contextual metadata efforts are, however, limited in that they fail to take advantage of the distinctive characteristics of "context" in particular domains. Moreover, they do not extend into an "awareness" framework and so fail to gather important data about the occurrence of events. Research in this sense needs to develop richer and more granular contextual awareness metadata schemas to enable more powerful retrieval and secure memories.

3.1 Relevant metadata schemas and ontologies for describing, organizing and retrieval student digital memories

It is impossible to find an e-learning information system, where metadata is not a core component. Among those schemas known to have a major impact on the design of current e-learning systems are IEEE LOM[5], IMS[6], SCORM[7], DCMI-Ed[8], and different application profiles related to them, which are created to describe the properties of learning objects. A limitation of these standardized metadata models is that they focused primarily on descriptive and administrative metadata for describing learning objects properties. As a result, they do not sufficiently help to organize, describe and retrieval learning memories[9] (where learning resources are produced within the working process) richer contextual metainformation is needed. The M^2 project is addressing this limitation by analyzing metadata schemas and ontologies to describe context awareness.

In late 90's, pervasive computing –then called only "mobile" computing– started studying metadata schemas to be applied for context awareness, like MCFE (Mobile Computing in a Fieldwork Environment) Metadata Elements [12]. MCFE was intended to enhance higher education learning describing (using a DC based metadata schema), virtual reality field courses and more practical scenarios focusing on spatial and temporal objects coverage. Now, GPS or RFID location information could be extracted automatically and validated through ontologies. Schemas like MCFE can not be used for content and structural metadata. Despite this limitation, more research on automatic metadata extraction is needed, and it needs to link more with ontological developments.

M^2 work on ontologies leverages existing works and incorporates Semantic Web developments, such as OWL (Web Ontology Language). Among existing influential work are:

---

[4] "Object" here is used in a generic sense as the traditional designation for electronic resources (DLO, Document Like Object), some authors identified this digital information unit as "artifacts" or simple "resources". It is considered like a minimum unit to be described by metadata (e.g., a word file, an image, an e-document), as well as happenings or experiences that might be recorded via the SenseCam or other technology.

[5] IEEE Learning Object Metadata: http://ltsc.ieee.org/wg12/

[6] The IMS Global Consortium has its metadata specification: http://www.imsproject.org/metadata/

[7] Shareable Content Object Reference Model): http://www.adlnet.gov/scorm/index.cfm

[8] Dublin Core Metadata Initiative-Application Profile on Education: http://dublincore.org/groups/education/; http://www.ischool.washington.edu/sasutton/dcmi/ed/04-05/DC-Education_AP_11-30-04.html

[9] Smith [11] called these: "context awareness learning objects".

- Ontologies to manage user/learner properties:
  - GUMO (General User Model Ontology, described in [19]) to define user features including psychological state, personality or demographics, which can be interesting to consider to evaluate learning behaviour.
  - IMS LIP (Learner Information Profile) [20], a XML schema to represent personal data, including accessibilities; activities; affiliations, competencies, interest or goals.
  - FOAF (Friend of A Friend), an RDF vocabulary to define people features and their relationships. The use of FOAF here is pretty interesting, since it will allow us to define the influence of teachers and other classmate or students, who give influence on learning process.
- Specific ontologies related to the information object generation with mobile technologies. SOUPA (Standard Ontology for Ubiquitous and Pervasive Applications) ontologies described in [3]) and other ontologies like FOAF, DAMLTime, the spatial ontologies in OpenCyc, COBRA-ONT, RCC (Regional Connection Calculus), MoGATU BDI, and the Rei policy ontology are important for this category of work.
- Specific ontologies used in higher education. At this level it is common in the literature to switch the strict concept of ontology as a set of statements and rules to represent a logical theory and/or a knowledge field, with the concept of metadata schema or set of elements, like LOM, DC, etc. But we have considered the omnipresence of ontologies in Education, as well as the value of ontologies themselves as a cognitive tool for learning [14].

3.2 Development, interoperability and integration with Context Awareness Framework (CAF)

The M^2 team has developed and evaluated a modular metadata schema. This schema extends UNC's Context Awareness Framework (CAF), which links a software agent on a student's computer with ontologies and rules to the university environment (remember Fig. 2). The schema is more elaborate that what MLB software currently supports, but the software may be enhanced over time to support more robust metadata functionalities. Initially, the M^2 schema has been constructed with a focus on supporting biology students that engage in courses where field study and lab elements are involved. The main objectives of our schema are:

a) Annotate captured personal education information (including these data types: lecture, conversation, audio and video, slides, handouts, student notes, quizzes, exams, papers, homework, projects, lab assignments, field exercises, images, and user profiles)

b) Determine education related context of students. Answering questions such as:
  - Where they are at (physical location. From country to campus room number)
  - What they are doing (in a meeting, study group, class, lecture, lab, etc.)
  - Who they are with (students, faculty, staff)
  - Student's cognitive state and personal behavioral characteristics.

Our schema can be generalized to other disciplines, particularly those involving field work as part of the educational experience. Preliminary testing demonstrates that our schema supports context retrieval, which can in turn aid memory, and ultimately contribute to richer learning experience. At the time of writing this paper, we are beginning an in-depth study to test the Memex research kits and the effectiveness of our metadata schema.

4 CONCLUSIONS

The 21st century provides many new learning options. The global education community encompasses students of all ages and nationalities, and needs to embrace pervasive computing. By taking this step, the educational community can learn how to effectively connect new technologies and learning science knowledge to improve educational centered on students. The M^2 project focuses on metadata needs for contextual retrieval, ultimately to improve learning. This paper has introduced the M^2 project, discussed E-portfolios and current educational trends related to pervasive computing; and reviewed relevant ontologies and their relationship to the projects' CAF (Context Awareness Framework). Future research plans include testing the usability of the Memex Research kits (SenseCam technology and MLB), enhancing MLB's metadata functionalities, and testing learning supported by pervasive technologies. All oft his work will inform the development of the projects modular metadata scheme and supporting ontologies, and allow identify how the University and students can take advantage of and effectively use new technologies to enhance learning.

ACKNOWLEDGMENT

Thank you to Microsoft Research, ITS Technology Assessment and Planning Department of the University of North Carolina at Chapel Hill, and the Metadata Research Center (MRC) for supporting the M^2 project. Thank you also to the Council for International Exchange of Scholars (CIES) and Fulbright Program for supporting Dr. Eva Méndez as an EU Research Scholar at MRC, School on Information and Library Science of the UNC, which allowed her be a part of the M^2 project.

REFERENCES

[1] Memex Metadata (M^2) Project: http://ils.unc.edu/mrc/memex

[2] Guo, Zinan and Greer, Jim. Connecting E-portfolios and Learner Models. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005*, Chesapeake, VA: AACE, pp. 2075-2080.

[3] Chen, Harry, et al. SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications. In *International Conference on Mobile and Ubiquitous Systems: Networking and Services*, August, 22 2004. Available at: http://ebiquity.umbc.edu/_file_directory_/papers/105.pdf

[4] Greenberg, Gary. 2004. The Digital Convergence: Extending the Portfolio Model. *EDUCAUSE Review*, July/August 2004, 39 (4), pp. 28–37.

[5] European Commission: Education and Training 2010: http://ec.europa.eu/education/index_en.html

[6] ECTS: European Credit Transfer and Accumulation System: http://ec.europa.eu./education/programmes/socrates/ects/index_en.html

[7] Context Awareness (Wikipedia): http://en.wikipedia.org/wiki/Context_awareness

[8] Teevan, Jaime. The Perfect Search Engine is not Enough: A Study of Orienteering Behavior in Directed Search. In *Proceedings of the Conference on Human Factors in Computing Systems*, 2004 , Vienna, Austria, pp. 415-422.

[9] Lonsdale, Peter, et al. A context awareness architecture for facilitating mobile learning. In *Proceedings of MLEARN 2003, London, LSDA.* Learning with Mobile Devices. London: Learning and Skills Development Agency, pp. 79-85

[10] Lonsdale, Peter, et al. Context awareness for MOBIlearn: creating an engaging learning experience in an art museum. In *Proceedings of MLEARN 2004*. Bracciano, Rome: LSDA, 2004.

[11] Schmidt, Andreas. Potentials and Challenges of Context-Awareness for Learning Solutions. In, *Proceedings of the 13th Annual Workshop of the SIG Adaptivity and User Modeling in Interactive Systems (ABIS 2005) Saarbrücken, October 2005.* Available online at: http://www.andreas-p-schmidt.de/publications/abis05_aschmidt.pdf

[12] *MCFE (Mobile Computing in a Fieldwork Environment) Metadata Elements ver 0.2.* (1997): http://www.cs.kent.ac.uk/projects/mobicomp/Fieldwork/Notes/mcfemeta.html

[13] Davis, Marc, et al. MMM2: Mobile Media Metadata for Media Sharing. In Extended Abstracts of the Conference on Human Factors in Computing Systems, *Portland, Oregon*, April 5-7, 2005, pp. 1335-1338.

[14] Dicheva, Darina, et al. Ontological Web Portal for Educational Ontologies. In *International Workshop on Applications of Semantic Web technologies for E-Learning (SW-EL)* 2005. Available online at: http://www.win.tue.nl/SW-EL/2005/swel05-aied05/proceedings/4-Dicheva-final-full.pdf

[15] EDUCASE NLII: http://www.educause.edu/nlii/

[16] Naismith, Laura, et al. *Report 11: Literature Review in Mobile Technologies and Learning.* Bristol: Future Lab, 2004. Available at: http://www.futurelab.org.uk/download/pdfs/research/lit_reviews/futurelab_review_11.pdf

[17] Capra, Licia, Emmerich, Wolfgang and Mascolo, Cecilia. Exploiting Reflection and Metadata to build Mobile Computing Middleware. In *Advanced Topic Workshop Middleware for Mobile Computing November 16, 2001, Heidelberg, Germany*. Available online at: http://www.cs.arizona.edu/mmc/16%20Capra.pdf

[18] IMS ePortfolio binding: version 1.0 final specification (June, 2 2005): http://www.imsglobal.org/ep/epv1p0/imsep_bindv1p0.html (See: IMS ePortfolio specification: http://www.imsglobal.org/ep/)

[19] Heckmann, Dominik, et al. GUMO: The General User Model Ontology, In *Proceedings of UM 2005: International Conference on User Modeling, July 24 -30, 2005, Edinburgh, UK.* Available online at: http://w5.cs.uni-sb.de/~schmitz/publications/UM05_Gumo.pdf

[20] IMS LIP (Learner Information Profile): http://www.imsproject.org/profiles/index.html