Specifying Semantic Conformance Profiles in Reusable Learning Object Metadata

Miguel-Angel Sicilia, Elena García, Carmen Pagés
Computer Science Department
Polytechnic School, University of Alcalá
Ctra. Barcelona km 33.6 – 28871 Alcalá de Henares (Madrid)
Spain
{msicilia,elena.garciab,carmina.pages}@uah.es

Salvador Sánchez-Alonso
Computer Languages and Systems Department
Pontifical University of Salamanca (Madrid campus)
Paseo de Juan XXIII 3 - 28040 Madrid
Spain
salvador.sanchez@upsam.net

Àngels Rius
Computer Science Studies
Open University of Catalonia
Av. Tibidabo 39-43
08035 Barcelona
Spain
mriusg@uoc.edu

***Abstract*—Current learning technology standards and recommendations have defined common languages for describing and sequencing learning resources. But further work is needed to provide common and consistent means for many processes internal to *Learning Management Systems*, which are essential to achieve a high level of automation. This paper describes the concept of *semantic conformance profile* as a way to specify complex run-time behaviors for such processes in a flexible way. These profiles use a contract-based specification to make run-time semantics clear, and are intended to be used in conjunction with specialized ontologies for the steps that require a complex representation. Five basic profiles are sketched as use cases to drive the specification approach.**

I. INTRODUCTION

The growing interest in Web-based learning has fostered the process of standardization of learning contents. As a result, several initiatives have proposed specifications for diverse aspects of Web learning [1], and the LOM metadata specification [2] has reached the status of international standard. The concept of *reusable learning object* (RLO) [3] – representing "independent and self standing units of learning content predisposed to reuse in multiple instructional contexts" [4] – is the central structuring notion underlying standards, specifications and modern reusable content design methods. The description of RLOs in standardized form is achieved by associating *metadata records* to the Web contents that comprise the RLO, and in consequence, the quality of these records becomes critical to achieve reusability [5]. The first step to characterize quality metadata records is that of defining what a *complete* metadata record is, with respect to the prospective usage scenarios of the RLO, i.e. it's necessary to define the metadata elements that are required for each automated functionality, along with their unambiguous semantic interpretation.

Recent studies have pointed out that current unstructured metadata annotation practices produce metadata records that are mostly useless from the perspective of automated processing [6]. To overcome such barriers to automated reusability, *semantic conformance profiles* (SCPs) are required for specific functionalities like RLO location, trading, aggregation or device-adaptation. Such metadata specifications are oriented to be processed by software modules or agents, and thus, they should ideally be formal or semi-formal, and have an unambiguous interpretation. For example, the *design by contract* philosophy has been recently proposed as a possible technique to express pre- and post-conditions on RLO usage [7, 8]. Conformance requirements in SCORM [9] can be considered as specific cases of SCPs, but currently they only cover basic (although important) processing, oriented towards course launching, sequencing and packaging.

This paper describes how SCPs can be specified in terms of required metadata elements, metadata idioms, and run-time commitments. Required metadata elements are the meta-information items that are required for the given functionality, idioms are requirements for its specification, and run-time commitments are the actions that are expected to be carried out by the system(s) supporting the functionality. In addition, it is described in which points such definitions can be integrated with *Semantic Web* ontologies [10], enabling richer semantic descriptions and eventually, inference on metadata descriptions.

The rest of the paper is structured as follows. The second section describes the concept of semantic conformance profile, and the essential elements of their description. Sections three to five provide details about five basic profiles that are essential to any automated process of learning object selection, composition and delivery. Finally, conclusions and future research directions are provided in the sixth section.

II. THE CONCEPT OF SEMANTIC CONFORMANCE PROFILE

A SCP can be defined as "a contract-based specification of a basic LMS process oriented towards its automation". The contractual approach is intended to specify the prerequisites or *pre-conditions* required for the process to take place, as long as the expected outcomes or *post-conditions* resulting from its execution. Such approach clearly delineates the responsibilities the LMS assumes if the required preconditions are satisfied, and thus forms a basis for normative conformance with regards to the effects of the process being carried out. Table 1 sketches some example SCPs. Each of them is specified in terms of three aspects, namely, "Required metadata elements", "Description idioms", and "Run-time commitments entailed by the profile".

In the column "Required Elements", the metadata that needs to be accessible for the profile is listed. LOM items are listed with their associated standard number, and additional aggregated elements (i.e. that need to be described with further levels of detail) are put into square brackets. Such list is a set of data preconditions that is complemented by additional constraints of any kind described in the "Idioms" column. The effects or post-conditions and run-time requisites are listed as assertions that must be satisfied after successful completion of the process described in the profile, or during its execution. Curly braces are used to denote effects that are complex to specify and thus open to different degrees of conformance, due to their inherent vague or multifaceted nature.

TABLE I
FIVE BASIC SEMANTIC CONFORMANCE PROFILES

Profile	Required Elements	Idioms	Run-time commitments
`ACQ`	6.1. Cost. 6.2. Copyright and other. [Buying conditions] [Seller System]	a) Localized cost and copyright. b) [Seller System] available through P	a) Charge_Unit validated. b)[Seller System] functioning. c) [Buying conditions] attainable d) Audit enabled. e) Buy {justified}
`CMP-1`	9. Classifications [Content spec] [Content separation]	a) [Domain ontology connection] b) Independence	a) Appearance merging. b) Semantic coherence c) {Metadata coherence}
`U-SEL`	5. Educational [Pedagogical contexts]	a) [Pedagogical ontology]	a) Selection {justified}
`P-SEL`	4. Technical	a) Include detailed device required capabilities	a) Device self-description.
`PUB-1`	1.1. Identifier 1.2. Title 1.3. Language 2.2. Status [Appropriate content packaging]	a) Global identification scheme. b) Status ≥ final c) Standard physical packaging (e.g. IMS)	a) Basic retrieval protocols according to the global identification scheme. b) Learning object available

It should be emphasized that these profiles are deliberately protocol – or technology – neutral, since they are intended to become canonical representations of abstract processes (the way OAGIS[1] processes are in the area of B2B), that would eventually be mapped to one or several enabling technologies.

The `ACQ` (Acquisition) profile is intended to describe the automated or semi -automated buy of a RLO to fulfil a given learning objective inside a *Learning Management System* (LMS). The cost, buying condi tions and copyright must be specified in the metadata record to enable the automated transaction. Moreover, such items must be localized or "localizable" to the conditions of the buyer, and the seller system(s) require a specified protocol `P` to carry out t he transaction (e.g. using an e -commerce infrastructure like `ebXML`). The LMS can be expected to validate the account to be charged and the proper functioning of the seller, and it should check the conditions, and audit the transaction. In addition, the tra nsaction must be justified from the viewpoint of the stakeholder. This latter commitment is largely system - dependant (as denoted by the braces) and may involve complex decision procedures.

The basic Composition (`CMP-1`) profile is intended to situations in which a LMS decides to automatically aggregate two or more learning objects into the same learning-oriented structure. RLOs can be joined together in sequences or other structures when they contribute to different aspects of the same topic. Such content similarity can only be achieved by detailed domain ontologies, more elaborated than current LOM `keywords` element, which could be embedded into the LOM `Classification`s category. Independence (or the presence of only explicitly and machine-readable dependencies) is also required, but this is considered a definitional condition for RLO. In addition, it is required that the contents are physically separated from presentation [11] – e.g. by using stylesheets –, since overall usability design requires a consistent appearance in the aggregated learning objects.

The User-Selection (`U-SEL`) profile is aimed to capture the semantics of targeted search of a RLO for a given need. It requires the provision of `Educational` LOM metadata, enhanced with specific separation of prospective pedagogical contexts [5]. Advanced consideration for learning styles and learning theories or approaches could be achieved by shared pedagogical ontologies. The system must be able to justify its decision in terms of the learning objectives (this again is largely system-dependant).

The Platform Selection (`P-SEL`) profile requires a detailed specification of `Technical` metadata, including detailed required device capabilities that are necessary to cope with the wide heterogeneity of mobile devices. The FIPA ontology[2] can be used to attain such level of detail. To be able to select RLOs according to their technical requirements, the LMS

[1] http://www.openapplications.org/
[2] http://www.fipa.org/specs/fipa00091/XC00091C.pdf

should be able to self-describe the devices it uses to deliver learning contents.

The basic publication (PUB-1) profile requires at least basic identification data, and a global and widely accepted identification scheme, and such location scheme is required at the side of the LMS.

Several profiles would be commonly required to carry out everyday's operation in an automated LMS. For example, in knowledge gap analysis, U-SEL based functionalities will be constrained by P-SEL restrictions, and ACQ conditions.

It should be noted that profiles as those described above could be used as criteria for *completeness* of metadata records. In other words, conformance with SCPs determines the functionalities that fulfil the preconditions to enable certain kinds of functionality. Nonetheless, some of the requirements stated in profiles are fairly difficult to validate by automated means, e.g. pedagogical adequacy requires a consideration of many dimensions of the learning experience that calls for rich knowledge representation strategies or the intervention of human experts. At these points, the use of logics-based knowledge representation provides a richer support than existing simple attribute-value schemas.

In what follows, more details are provided about the five basic profiles in Table 1, using a syntax PROFILE(PARAMS) to denote the elements involved in the profile as parameters of the process.

III. Publication and Acquisition

The PUB-1(O1, LOR1) profile groups a number of basic requisites for learning object repositories, with regards to accessibility of learning objects. The minimal set of metadata required for a learning object O1 to be accessible is a globally unique identifier, and a title, language and status information. The status of the learning object needs to be greater than "final" according to the LOM vocabulary (or any equivalent status in other schemes), since PUB-1 is intended to result in learning object O1 ready to be used at least through the repository LOR1 (and perhaps in other, federated repositories). In addition, the learning object must be encoded in a declared, public physical format (e.g. the SCORM content packaging format, based in the IMS one), so that LMSs are able to decide if the contents are "legible" for them. This PUB-1 profile does not consider relationships from O1 to other learning objects, so that it would require additional extensions in the presence of each kind of relationship [15].

The implementation of this profile could be done through *Web Services* according to common interoperability protocols, but such communication is outside the scope of the profile itself, since it belongs to profiles of "register and query" for repositories or mediated by repositories, using interfaces provided for that purpose, e.g. [12].

The ACQ(O1, SS1, LMS1) profile is a typical example of LMS-initiated process that is very close to current specifications for B2B e-commerce like *OAGIS* or *RosettaNet*. Basic information needed about the learning object being bought comprises localized cost (not only the fact that it is subject to payment, but its amount), and also copyright and other buying conditions. Note that such specification is complex n the general case, involving rights transfer and legal regulation, as addressed, for example, by the XrML language[3]. In addition, the seller system SS1 must be available, including complete binding information.

The minimal commitments for the ACQ profile include the following:

- A "Charge Unit" at the buyer (LMS1) should be validated for permission for the transaction.
- Buying conditions must be attainable according to the criteria of LMS1. This entails consideration of available budget.
- The operation must be audited both at LMS1 and SS1 sides, to support traceability of business operations.
- The buy must be "justified" according to some kind of individual or organizational need. This "explainability" of the decision to buy LO1 could be simple or complex, depending on the system, and it ideally connects a "knowledge gap" identified to the knowledge the learning object is supposed to facilitate.

This last consideration of learning objects as commodities require an explicit account of learning objects outcomes, that could be expressed in terms of categorizations or as "post-conditions" as described in learning object contracts [8]. This should be reflected in the profile as part of the {justified} verb. ACQ processes could be the result of learning object selection processes, in which case, the process is explainable in terms of the associated SEL process(es).

IV. User and Platform Selection

Platform selection can be informed by the Technical metadata category in LOM, so that effective adaptation can be done by comparing the metadata of the learning object with a description of the target context of the user, i.e. P-SEL(O1, CTX1). Nonetheless, the description of "context" of use no longer is determined by the physical capabilities of the interaction device, but also with environmental conditions that would become more common with pervasive technology. Examples are dynamic descriptions of illumination or noise conditions. Such detail in platform descriptions is not covered in LOM currently, and it would be necessary to update it continuously. The required commitment for P-SEL is that the devices used in learning provide self-description capabilities through standardized protocols.

The selection of learning objects targeted to the user is a complex problem that has been addressed by various previous research efforts in the areas of intelligent tutoring and adaptive hypermedia. A basic realization of the profile would only consider metadata elements included in LOM

[3] http://www.xrml.org

Educational category, but richer schemas can be used instead. A tradeoff between ease of implementation and description and richness can be obtained through the use of learning object contracts [7, 8], which allows for the specification of pre-requisites on the user regarding states of “knowledge”. These states of knowledge can be expressed in terms of domain ontologies that help in clarifying the relationships of knowledge items of an arbitrary granularity.

In addition to such ontology, an additional representation is required so that the (educational) *context* of the user (i.e. his/her characterization as a learner) matches to one of the possible pedagogical contexts in which the learning object was declared to be usable – see [5] for the rationale of such contexts. In consequence, the profile can be specified as `U-SEL(USER1, {LO})` where `{LO}` represents a collection of available objects that are candidate to be selected for the user.

The main problem of this approach is that the characterization of “educational usage context” is difficult to represent through metadata due to the number of factors influencing the usability of a learning object in a given situation (age, style, density, social factors and the like).

The main run-time commitment of `U-SEL` is that the selection should be justifiable in terms of the actual and intended educational contexts. This required “explainability” complements the one present in `ACQ`, which is oriented to organizational and not individual justification. In fact, `U-SEL` processes would in some cases trigger `ACQ` ones.

A straightforward extension to `U-SEL` could be that of targeting learning objects to groups (`G-SEL`), which essentially provides the difference that delivery is multi-cast, and group descriptions using the same schemas of `U-SEL`.

V. Composition

The basic composition profile `CMP-1` is oriented towards the automation of learning object aggregation into higher levels of instruction. Concretely, it uses `Classification`s of the contents of the learning object as the information that drives the composition. The profile can be expressed in terms of candidate objects, and a specification of the objectives of the aggregate, which in its simplest form can be a set of terms inside a classification describing the intended outcomes, i.e. `CMP-1({LO}, {C})`.

Domain ontologies as those described for `U-SEL` can be used to annotate objects with the “knowledge they provide”, so that a system composing a learning experience with objective `C` is able to obtain through the ontology pre-required knowledge items `Ci`, thus initiating search for objects covering each `Ci`. The actual organizations of the selected objects is only constrained by the relationships in the ontology, so that extended `CMP` profiles could be devised to make room to adapting with specific organizations, as those enabled by SCORM sequencing.

Context separation and independence are required for the LMS to be able to compose learning objects from the technical and semantic viewpoints, respectively. In consequence, it is required that the resulting aggregate is seamlessly merged both in appearance and content. In addition, the metadata records should be *coherent* in a general sense, that is, they should not contain contradictions as described in [14]. For example, the technical requirements of the aggregate should cover at least the technical requirements of its parts.

The composition of learning objects would entail additional processes except in trivial cases. The declared dependencies of objects `Ci` being aggregated are required to recursively verify the coherence of metadata records to them, and they eventually would result in chained `ACQ` processes. In fact, dependencies between learning objects entail the propagation of run-time actions that should be carefully studied for each type of profile.

It should be noted that `PUB-1` is necessary but not sufficient to provide the pre-requisites of `CMP-1`, since it only enforces the provision of the title, language and identifier information. In consequence, an additional profile (say `PUB-C`) would be required to close the path from publication to composition that includes the provision of at least minimal classification information regarding the contents of the learning object(s) involved. Such new profile could be declared as a **requisite** for `CMP-1` and as an **extension** to `PUB-1`, so that any learning object published through `PUB-C` is prepared to be used in `CMP-1`. This kind of connections between profiles makes convenient the storage of the history of processes carried out in the past, as a source of information.

VI. Conclusions and Future Research Directions

The concept of *semantic conformance profiles* has been described as a way to specify internal processes required or enacted by LMSs. These profiles are intended to complement existing standards, broadening their scope to processes that are internal to *Learning Management Systems*, and also providing a contract-based specification that clarifies their run-time semantics.

Five basic conformance profiles have been sketched as an advance for more detailed specifications. As a result of the process of specification of profiles, learning object metadata can be classified according to the profiles that can be fulfilled with its current metadata record. This way, for example, a learning object with no cost information does not fulfill the criteria of *completeness* for `ACQ`.

Future work should add detail or refine the profiles sketched here, and it should also provide additional profiles for a variety of automation processes. Basic profiles can be used to define more complex ones, and the cause-effect relationships between processes should also be subject to further inquiry.

In addition, formal languages and knowledge representations like *Cyc* [13] should become an integral part of the approach, enabling the construction of *Semantic Web* applications. A promising technology in that direction is that of the *Web Service Modeling Ontology* (WSMO)[4], which provides a model to describe needs and capabilities that could be used for the implementation of profiles.

REFERENCES

[1] Anido, L. E., Fernández, M. J., Caeiro, M., Santos, J. M., Rodríguez, J. S., and Llamas, M. (2002). Educational metadata and brokerage for learning resources. Computers & Education, 38(4), 351 – 374.

[2] IEEE Learning Technology Standards Committee. Learning Object Metadata (LOM). IEEE 1484.12.1, 2002.

[3] Wiley, D. A. (2001). The Instructional Use of Learning Objects. Association for Educational Communications and Technology, Bloomington.

[4] Polsani, P.R. (2002) The Use and Abuse of Reusable Learning Objects. Journal of Digital information, volume 3 issue 4. Available at http://jodi.ecs.soton.ac.uk/Articles/v03/i04/Polsani

[5] Sicilia, M.A., García, E. On the Concepts of Usability and Reusability of Learning Objects. International Review of Research in Open and Distance Learning 4(2)

[6] Pagés, C., Sicilia, M.A., García, E., Martínez, J.J., Gutiérrez, J.M. On The Evaluation Of Completeness Of Learning Object Metadata In Open Repositories. In: Proceedings of the Second International Conference on Multimedia and Information & Communication Technologies in Education (m-ICTE 2003), 1760-1764.

[7] Sánchez-Alonso, S. and Sicilia, M. A. "Expressing preconditions in learning object contracts", in Proceedings of the Second International Conference on Multimedia and Information & Communication Technologies in Education (m-ICTE 2003), pp 1734-1738. Available at: http://ssanchez.colimbo.net/papers_en.htm

[8] Sicilia, M.A. and Sánchez-Alonso, S. Learning Object "Design by Contract". WSEAS Transactions on Systems, 2003. Available at: http://ssanchez.colimbo.net/papers_en.htm

[9] Advanced Distributed Learning (2002). Sharable Content Object Reference Model Version 1.2 Conformance Requirements, available at http://www.adlnet.org.

[10] Sicilia, M.A., García, E. (2003). On the integration of IEEE-LOM metadata instantes and ontologies. IEEE LTTF Learning Technology Newsletter, 5(1) (Special Issue on "Implementation(s) of learning objects metadata standard").

[11] Bohl, O., Schellhase, J., Sengler, R., and Winand, U. (2002). The Sharable Content Object Reference Model (SCORM) – A Critical Review. In Proceedings of the International Conference on Computers in Education (ICCE02), Auckland, New Zealand, 950 – 951.

[12] Hatala, M., Richards, G., Eap, T. Willms, J. "The eduSource Communication Language: Implementing an Open Network for Learning object Repositories and Services", in Proceeding of the special track on Engineering e-Learning Systems held at ACM Symposium on Applied Computing (SAC) 2004, Nicosia, Cyprus, March 2004.

[13] Sicilia, M. A., García, E. and Sánchez-Alonso, S. "On integrating learning object metadata inside the OpenCyc knowledge base". In Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies - ICALT 2004. Joensuu, Finland. (to appear).

[14] Sánchez-Alonso, S. and Sicilia, M. A. "How learning object relationships affect learning object contracts: commitments and implications of aggregation". In Proceedings of ED-MEDIA 2004 - World conference on educational multimedia, hypermedia and telecommunications. Lugano, Switzerland (to appear).

[15] Sánchez-Alonso, S. and Sicilia, M. A. "On the semantics of aggregation and generalization in learning object contracts". In Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies - ICALT 2004. Joensuu, Finland. (to appear)

[4] http://www.nextwebgeneration.org/projects/wsmo/