Förord

Den här rapporten består av två artiklar som skrivits för två konferenser inom olika områden – Information Modelling and Knowledge Bases XI och WebNet 99. Båda artiklarna emanerar från det arbete som bedrivs inom SITI, främst hos SISU men också IMT medverkar, och som berör användandet av metadata – information om information. Artiklarna förenas inte bara av detta gemensamma problemområde och delvis samma författare utan också av RDFClient – en prototyp utvecklad på SISU som i båda fallen används för att demonstrera modellerna.

Problemområdet är den numera överväldigande mängd information som finns tillgänglig – inte minst genom Internet och World Wide Web (WWW) – som har gjort att frågor om ökad sökbarhet, bättre struktur, kategoriserat innehåll m fl aktualiserats. Detta gäller inte minst inom den nu snabbt växande marknaden för webbaserad utbildning där allt fler lärresurser görs tillgängliga på nätet. Det ökade utbytet av elektroniska dokument ger upphov till samma typ av frågeställning.

Inom flera standardiserande organisationer pågår arbete med att finna former – modeller – för hur dessa frågor skall kunna hanteras. William Song – en av författarna – har bl a deltagit i sådant arbete inom W3C (RDF Schema workinggroup) och forskargruppen följer noggrant utvecklingen inom konsortiet och andra organ som bedriver liknande arbete.

Den första artikeln **Electronic Document Exchange and Management for the Governmental Organisations** är baserad på arbetet inom MIWIS – ett samarbetsprojekt i vilket SISU och företrädare för några offentliga institutioner, bl a Riksarkivet och Statskontoret, medverkar. Artikeln beskriver de problem som är förknippade med utbytet av elektroniska dokument inom den offentliga sfären och med medborgarna. En konceptuell metadatamodell, kallad WDM, utvecklad inom projektet presenteras i artikeln och exemplifieras med hjälp av den ännu ej existerande elektroniska självdeklarationen.

Den andra artikeln **A Component-based Framework for Description and Management of Learning Objects** baseras på arbetet inom LOUIS som ingår i forskningsprogrammet FIOL i vilket SITI, Ericsson och Telia deltar. LOUIS fokuserar på Learning Objects – lärobjekt – och hur dessa skall kunna beskrivas för att underlätta dels för den lärande som har behov av att finna de relevanta resurserna i sin specifika lärsituation, dels också för producenter av lärresurser som bl a har behov av att kunna återanvända lärobjekt i olika sammanhang. I artikeln presenteras en komponentbaserad struktur för att hantera lärobjekt.

Peter Karlberg

Ansvarig för programmet IT och offentlig sektor, SITI

Abstract

The rapidly growing use of information in electronic format, i.e. the Web and other electronic document exchange creates a demand for better management of different resources. Use of metadata facilitates on way of providing structure and thereby better manageability. In this report metadata models are explored and extended within the management of electronic documents in the governmental sector and within the use of Learning Objects in the learning domain.

During the electronic document exchange processes, new information may be added, existing information may be modified, and control information has to be checked. For different purposes different formats are created and used by different organisations even if they share one database system. Our investigation in the project MIWIS (Metadata Infrastructure for the Web Information Systems in Swedish Governmental Sectors) for the Swedish governmental organisations has revealed that regional and local governmental organisations and departments usually have their own database systems and use their own customised electronic document formats. Their documents are sent out along with the localised formats. These localised formats hinder the exchange of electronic documents among organisations. The project focuses on the establishment of an intermediate metadata model, which is able to communicate with different electronic document formats, forms and styles and convert them into a set of uniquely represented schemas. A collection of criteria is suggested to measure the metadata model defined.

The growing use of the Internet and the Web has been transforming the styles of teaching and learning. The huge amount of learning resources on-line enables people to access to almost any kind of educational subjects and domains. The Internet and the Intranet are also used by organisations as technique support to train their employees and to enhance their competence. It is quite flexible for the trainees to acquire any kind of learning resources at any time and anywhere when the Internet is available. It is also flexible for the learning resource providers to supply the learning information to the Internet and the Intranet. However, due to the huge amount of information supplied and the bad-defined structures for the learning information, it is difficult for the learners to find exact learning resources and to find the information easily. In order to solve this problem, a well-defined metadata model is required. The model should be able to meet the characteristics of learning domains. In this paper, we propose a component-based metadata framework for describing learning objects and learning domain. We also discuss some necessary manipulations applied to the learning objects and try to combine them in the approach. An initial implementation for description of the learning objects and their operations is as well presented.

Electronic Document Exchange and Management for the Governmental Organisations*

W. Song, S. Hakkarainen, P. Karlberg, T. Soltesz
Swedish Research Institute for Information Technology (SITI)
Box 1243, Isafjordsgatan 22, 164 40 Kista, Sweden

1. Introduction

With the massive use of information systems, in particular, information systems being widely accessed through the World Wide Web (WWW), exchange, communication, and transfer of information among organisations is through the form of electronic documents. People submit electronic documents to and receive them from various communication media, such as the web, e-mails, etc. The electronic documents are produced by a number of word processing tools, such as Word, HTML editors, etc., and may have different formats.

During the electronic document exchange processes, new information may be added, existing information may be modified, and control information has to be checked. For different purposes different formats are created and used by different organisations even if they share one database system. Our investigation in the project MIWIS (Metadata Infrastructure for the Web Information Systems in Swedish Governmental Sectors) for the Swedish governmental organisations has revealed that regional and local governmental organisations and departments usually have their own database systems and use their own customised electronic document formats. Their documents are sent out along with the localised formats. These localised formats hinder the exchange of electronic documents among organisations.

In the WWW community, the metadata concept is introduced to provide a set of structural descriptions of information. A number of metadata models have been proposed. Some of them are developed along with their machine-readable syntax, for example, RDF (Resource Description Framework) [1]. A metadata model stipulates a set of object types, a set of relationship types to link the object types, and possibly a set of value types from which an object type can select values. An instance of such a model is called a schema. A schema is used to represent a particular domain of problems of interest in terms of the constructs defined in the metadata model. Usually, a metadata tool is also implemented along with the metadata model to interpret the schemas developed from the model.

* The work presented in this paper is supported by the SITI project MIWIS, No. 44830 and presented at the Information Modelling and Knowledge Bases XI.

The SITI Project MIWIS focuses on the establishment of such an intermediate metadata model, which is able to communicate with different electronic document formats, forms and styles and convert them into a set of uniquely represented schemas. A collection of criteria are suggested and discussed in the following sections to measure the metadata model defined in the project. The project will also produce a working prototype to implement the metadata model and its interface for formats description, editing and design.

The paper is organised as follows. In the next section, we discuss the problems on electronic document exchange and circulation among the governmental organisations as well as the supporting structures for the information management between and within the organisations. Then we present a conceptual metadata model for the description of electronic documents in Section 3 with an example to illustrate a schema using the model. In Section 4, we describe an implementation of the metadata ideas by using XML [2] and RDF based approach. Finally we conclude the paper and discuss shortly our future work.

2. Problems and the situations

In governmental organisations, there are a great number of electronic documents flowing from one authority to another. These documents are produced using different syntax and structure [3]. For example, one organisation may send an electronic document (e.g. e-mail) to its more than a hundred local administrations for a periodical report. The organisation later will receive up to hundreds of reports in electronic forms. Each report may have its own form. This will leave the administrative staff a heavy task to re-format the reports in order to have them in a standard structure.

More often than not, a document may be circularised around a group of governmental organisations with each of them adding their comments, opinions, and suggestions to the document. We can easily foresee that the metadata information and probably the document itself grow during the circulation. Some of the additional metadata information may later be included in the document whereas new descriptive information in the document may become a part of the metadata information. We are particularly interested in how to cope with the accumulative metadata information in our metadata document model and attempt to define a group of dynamic constructs in the model to describe this situation.

In practice, local administrative units usually maintain one or several federated (and/or distributed), possibly heterogeneous information systems. The main data manipulations are performed locally. The local systems maintain their own data schemas in local data forms. This situation is called information system autonomy. The information exchange with the outside world will be managed through the Internet. Because of the local data formats and local manipulations, two kinds of inconsistencies may arise. One is the inconsistent forms or data structures used by the local autonomic systems and by the outside world’s systems. The other is the conflict data formats applied within the local information systems.

Intranet-based document management systems are also frequently applied by the governmental organisations. This is to enable the organisations to easily maintain and control their electronic documents within an internal network. We can assume that partial and informal standardisation of the intranet-based document formats can be achieved. However, it is still necessary to maintain a well-defined metadata description mechanism within the intranet in order to improve the information communication and access-control, such as confidential document management.

3. Electronic Document Modelling

As previously stated, it is indispensable to build up a description framework for various electronic documents. Some basic requirements on the description framework, i.e., a metadata model, can be described as follows.

1) The metadata model is able to describe all kinds of electronic documents.

2) It can act as an intermediate data model to communicate with the various data models (formats) already defined for electronic documents.

3) It should be easy to use, i.e., simple modelling constructs for description of documents.

4) It should be expressive, to represent and exchange both the structural formats and semantic connotation of documents. Based on the requirements proposed above, we developed a conceptual metadata model for electronic documents in the MIWIS project, called WDM [4].

3.1 WDM: A document metadata model

The model WDM gives a set of centric attributes for electronic document objects to be modelled. An electronic document contains many information items *per se*. Some of the information items are more important than the others in the sense of identifying the electronic document. In general, an electronic document object mainly consists of the following four types of components:

- Object identifier, which is a global or local identity of the document, e.g., URL of a web page, which is used to locate the object in the World Wide Web;

- Object content, which contains the information that a document sender expects to convey to the document receivers;

- Object metadata or formats (or in database posters), which contain the information for describing the object, e.g. the document creator, the document creation time, etc.;

- Referencing links, that the document object provides to give further information or related information that other document objects may supply; and

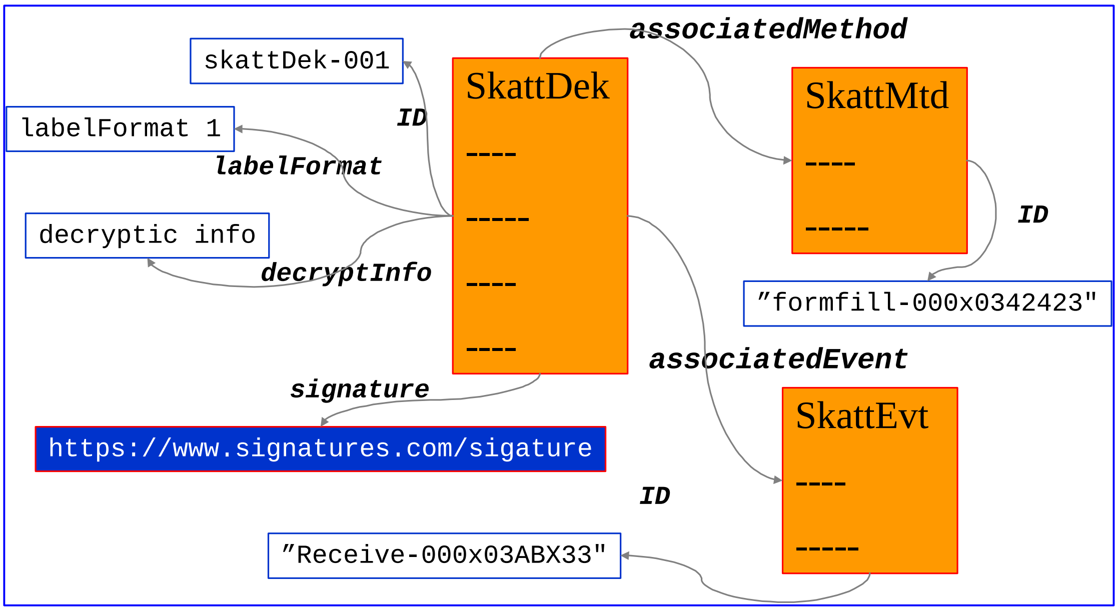
- A number of operation modules, that perform some functions on document objects for the readers' interaction.

The rationale for the WDM components lies in a sort of reasonable categorisation of a Web resource. Information content is undoubtedly a major, most important part that need to be described by a conceptual metadata model because it conveys meaning. The meta-information for information content includes for example title, leading paragraph, etc. The metadata (called format or layout sometimes) of a resource object is important because it usually provides classification information. Taking library as an example, books in a library can be classified by their authors or by their subjects. Referencing information of resource objects is particularly proposed in order to identify a resource object by its contexts.

Compared with other metadata models, such as RDF and DC [5], we consider that one of the main advantages of WDM is its capability of describing semantics of resources. RDF is able to describe objects and inter-object relations and DC can only describe objects in terms of attributes. In addition, WDM also takes into consideration the dynamic aspect of metadata. That is, WDM describes evolutionary processes of Web resources, such as versions.

3.2 Tax Declaration Document: A representation example

As an example for WDM, we use the tax declaration document issued by the Swedish Tax Board. Each tax declaration document has a unique number within Sweden. The documents vary between individuals, companies, and other organisations. Some of them may require decryption messages and signatures. Another document associated to the declaration document is the instructions (called “method document” in the metadata model) on how to fill a tax declaration form. Associated with a tax declaration document are also some events (i.e. operations), e.g., acknowledgements of receiving and sending the tax document. The following figure shows the documents used in the “Tax Declaration Process” and their interrelations, where the resource object “SkattDeK” is described by its attributes and other resource objects, such as ID – “skattDek-001” and associatedMethod – “SkattMtd”.



In terms of XML, the schema modelling the above-described tax declaration document is illustrated as follows.

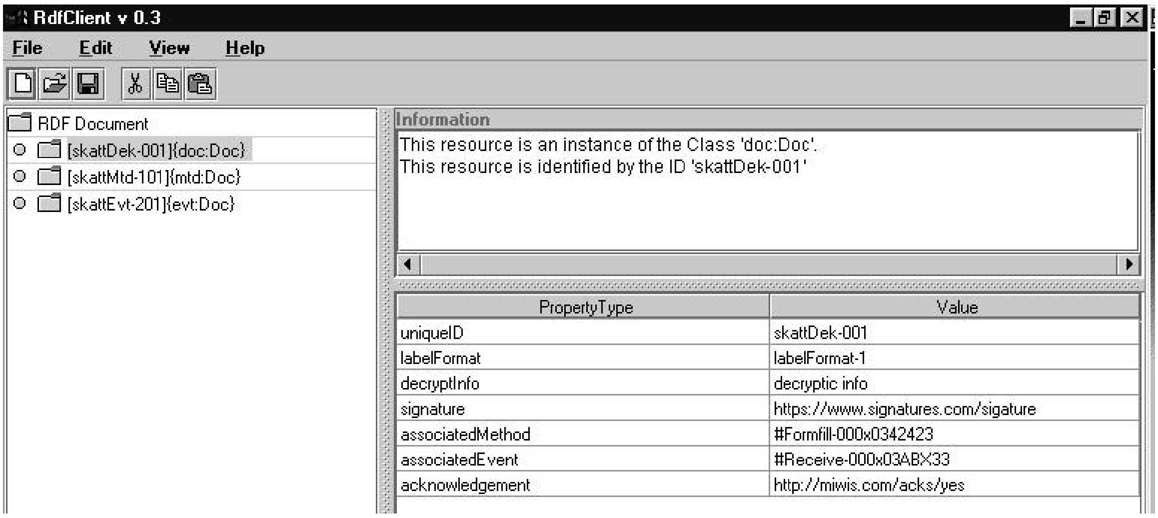
```
<wdm:WDM>
  <doc:Document ID="skattDek-001">
    <doc:uniqueID>skattDek-001</doc:uniqueID>
    <doc:labelFormat>1</doc:labelFormat>
    <doc:decryptInfo>decryptic info</doc:decryptInfo>
    <doc:signature>https://www.signatures.com/sigature</doc:signature>
    <doc:associatedMethod>#Formfill-000x0342423</doc:associatedMethod>
    <doc:associatedEvent>#Receive-000x03ABX33</doc:associatedEvent>
    <doc:acknowledgement>http://miwis.com/acks/yes</doc:acknowledgement>
  </doc:Document>
  <mtd:Method ID="formfill-000x0342423"/>
  <evt:Event ID="Receive-000x03ABX33"/>
</wdm:WDM>
```

Note that the strings starting with "#" stand for referencing links to other document objects. Objects or even object components can be embedded or indicated via their identities in the document object. These objects or components can be web pages or web documents, e.g. URLs.

4. A working prototype

A basic idea for the implementation of the metadata model, WDM, consists of three parts, a hierarchical structure for document objects (or object types), a collection of attributes for description of document objects, and a set of values for the object attributes. The hierarchical structure maintains two relationships between a parent node and its children nodes, i.e., is-a and part-of. For example, tax declaration document is (is-a) a document. The technical department is a part of (part-of) the tax declaration board.

The prototype, RdfClient, written in Java, can be run as Java Applet located in a web browser or as a stand-alone tool [6]. According to the idea based on RDF (Resource Description Framework) we just described, RdfClient has in its window frame-set three main components. To the left is the frame for the hierarchical structure where the main nodes are document objects. The nodes can be expanded and their children nodes can be viewed. This process can continue and the leave nodes are in general the attributes. The set of pairs – attributes (the leave nodes) and attribute values – is displayed at the right-down frame.



In the figure, three document objects, Tax Declaration Form, Form-fill Method Document, and Event Processing Document are represented in the left frame. The attributes of the document object skattDek (Tax Declaration) along with their values are presented at the right. In the current status of the prototype, the attribute values are editable and any changes made will be stored in the database for document objects. In the information frame, metadata information, i.e., modelling semantics and schema description, about document objects is displayed.

5. Conclusion and next step

It is extremely important to define a metadata model for description of electronic documents, which are exchanged and circularised between and in organisations. Considerations on the legal aspects such as privacy and publicity acts are also necessary when defining a metadata schema for public sector in particular. Our metadata model WDM is an intermediate metadata model with the aim to reach general appropriation, simplicity, and expressiveness. The model needs further development to include the possibility of describing hyperlinks, which are used on the web.

Our prototype has displayed its usefulness in expressing electronic documents along with their various properties and values. The next step is to develop a graphical interface for the users to facilitate definition of metadata modelling schemas (instances of WDM) and editing of the electronic documents for exchange and circulation.

In addition, when information communication among heterogeneous information systems, no matter within intranet or via the Internet, the problem regarding integration of metadata modelling schemas will become prominent, where understanding of metadata semantics will play an extremely important role. This consists in a main consideration and contribution of the WDM model towards conceptual metadata modelling and we will take this as one of our next research topics.

References

[1] **RDF (Resource Description Framework) Model and Syntax Specifications**, the W3C RDF Working Groups, http://www.w3.org/RDF/

[2] **XML (eXtensible Markup Language)**, http://www.w3.org/XML

[3] William Song, etc., **MetaContent Management of Electronic Documents**, *The SITI Project MIWIS Report*, SITI, Stockholm, Sweden, 1999

[4] William Song, **WDM: A Web Document Model and Its Supporting Web Document Analyzer**, *Proceedings of the American Conference on Information Systems – AIS'98*, Baltimore, MD, USA, 1998

[5] **DC (Dublin Core)**, http://purl.oclc.org/metadata/dublin_core/

[6] Dag Ekengren, **RDF for Multimedia Broker – Metadata on the Web**, *Master Thesis*, The Royal Institute of Technology, Stockholm, Sweden, 1998

A Component-based Framework for Description and Management of Learning Objects*

William W. Song and Mats B. Andersson

SITI, Box 1243, Isafjordsgatan 26, 164 28 Kista, Sweden

1. Introduction

The growth of the Internet and the World Wide Web is transforming teaching and learning at all levels of education in the workplace and at home [IMS]. Many training and education resources are available on the web and the users can easily educate themselves by fetching the education materials from the web. The Internet enables people to find every kind of learning information on various subjects and to access to the information at anytime and from anywhere. It is also quite convenient for the learning information suppliers to provide the educational resources and maintain them. These information pieces or resources for purposes of learning and education construct a new type of web resources. This new kind of web resources with quite distinct features is evolving on the Internet, called Learning Domain. An important characteristic of these resources (asynchronous systems) is that it can be used as a teaching medium as well as a source of knowledge about a subject area [MBA1, MBA2].

The Intranet techniques are particularly useful for large organisations to train the employees to enhance their competence. In particular, these courses and training packages are intentionally developed by large companies for their employees. The educational information is accessible through the Intranet within the organisations. Their employees can find the courses they need or they are required to take. Therefore, to find the required components and to integrate the components found from a learning object repository to form a new course are considered to be an important task for a learning information system.

Two major parts, closely related to a learning domain, are learning resource provides and consumers. The relationships between the learning information providers and the information consumers via the Internet can be described as follows.

The Internet can be considered as a resource and communication space, to which the information providers supply learning resources and from which the information consumers use the learning resources. Usually, learning information, such as courses, is stored and managed in an information system or a local web site. The supplier may give a sort of description of the structure of the information. This description is termed as metadata.

However, the information supplier usually makes such description from his/her own view of the learning domain of interest. The description of learning information by one supplier may be in conflict with that by another supplier. The description may not be a well-defined one so that it is difficult for the users to understand and hence to effectively search for it. It is indispensable to maintain a metadata model for the learning object descriptions.

* The work presented in this paper is supported by the SITI project LOUIS, No. 44850 and accepted at the International Conference WebNet'99.

From the point of view of the resource consumers, they expect to easily find the exact learning knowledge they need. They can provide a set of requirements or goals for searching for learning resources. The requirements and goals express what sort of learning materials they are seeking. The consumers may also provide their profiles as support. Indeed, profiles are sometime quite helpful. However, these requirements and goals can be vague, ambiguous, and even in conflicts. The profiles provided by the end users are very likely incomplete and bias-prone.

According to the users' needs and the knowledge about certain subject they already have, a suitable selection of learning objects is necessary in order to meet the users' requirements. Such possibility of selection depends on the granularity of learning objects. There is a relation between the degree of accessibility and what kind of information motivated to offer and how it should be structured. The granularity or refinement of learning objects in a learning system depends on the content and the carriers of the content.

No matter how the system will be used, it is necessary to handle the learning content in a modular form. Such module in a learning system is called learning object. Metadata and metadata model is used to support the organisation and structuring of learning objects.

Therefore, metadata modelling and granularity should be taken into account for the management of learning objects, while the premise for granularity is the decomposability of learning objects [WS1]. We consider breaking down a learning object into a number of sub-objects, called components. In the next section, we give some assumptions on possible structures and operations of learning objects and discuss some related work. In section 3, we propose a metadata model for learning objects through defining possible components within a learning domain. Then in section 4 and 5, we discuss architecture with two internal structures for the management system of learning domain as well as an implementation issue for learning system based on a metadata tool. In the final section we conclude the paper with some discussions of our future work.

2. Assumptions and Related Work

Assumptions

Suppose that we have a web resource, which is designed to describe an education product, for instance a training course. The course consists of a collection of textbooks, lectures, exercises, etc. The environment and style for learning and knowledge convey can be 1) within a classroom, 2) teaching on the students' requirements, 3) discussions over some particular issues, 4) self-learning. The communications may be face-to-face meetings, conversation through e-mails, or web readings.

In order to cope with such diversity of learning domain, we assume a **hierarchical structure** to represent a set of web resources and their interrelations (in the example of the education product, a learning object). Here we can further assume that an education product is **decomposable**. That is, an education product can be further decomposed into a collection of components, for instance, textbook, lecture, exercise, etc. Furthermore, these components (we will term them as objects, or web objects, resource objects, or learning objects in different circumstances) can themselves be decomposed. These components will be re-united to form a new education product based on the users' needs.

What **granularity** (i.e., how deep we shall decompose an object or a component) we should achieve relies on the users' requirements [WS2]. Let us take an example. Two students need some materials for a course. In general, the course consists of ten lectures, a number of homework, and one textbook. Assume that one student took five lectures once and did one homework, whereas another student happened to read some chapters of the textbook. Thus, the first student hopes to take the rest lectures and homework as well as reads the book while the second expects to take all the lectures and homework but the rest chapters of the textbook. Therefore, the "course" (virtual one) for the first student can be different from the “course” for the second student. In other words, textbook can be a component of a course in one case, while in the other case, chapter will be a component of a course. Different granularities employ to meet different requirements.

For each component, we assume a set of operations to be applied to it. Such operations probably include:

1) Refining (decomposing) – the component is a refinement of another component;
2) Changing – the attributes of the component have been changed;
3) Versioning – the component is a new version vs. its previous one after updating;
4) Neighbouring – the component is a neighbour of another component;
5) Updating – the component has been say renamed;
6) Integration – the component is an aggregation of a number of other components.

Related work

Many metadata models and methods proposed for the description and management of both information bases and the web resources. RDF (Resource Description Framework) [RDF], developed by the W3C working groups RDF Model and Syntax as well as RDF Schema, can be considered to be a general metadata model for describing the web resources, including electronic documents, images, sounds, movies, and other objects supplied to the web. In our implementation for learning systems, we consider a RDF based language for representing learning objects and their relationships.

DC (Dublin Core) [DC], developed within the society of library and information science, focuses on how to describe electronic publications and how to manage them. The DC model consists of fifteen elements for a publication, including subject, author, creation date, version, etc.

The IMS Project (Instructional Management Systems) aims to create recommendations for building Internet architectures for learning [IMS]. The work is divided into five groups and one of them is called Metadata. IMS propose a set of metadata to be used with RDF/XML for describing web resources.

At the level of practical research and application, a metadata prototype for training course management, developed at Ericsson, is called KNACK. A number of levels of description from curriculum to media carrier element are presented to reveal a deep infrastructure of primary components and their relationships for learning domain. This is significant for reconstruction and reuse of learning objects.

3. A Metadata Model for Learning Objects

We propose a framework of hierarchical (or tree like) metadata model, where each component (node) in the hierarchical structure consists of two major parts. One is the data description part and the other the operation description part. The data description part provides a set of building constructs for designing a metadata-modelling schema for a web document. The operation description part provides a set of operations that apply to the data of the components.

The framework consists of two models, a data model for data description and a component model for component and operation description. The data model contains three primitive constructs: objects, describing the concepts for e.g. documents, relationships, describing relations between objects, and attributes, describing object properties.

The component model describes different types of metadata, including e.g. carriers and operations. We consider that each piece of information has a sort of carrier. For example, a movie is a carrier of a story. The story can as well use another carrier, e.g. a textual document.

In the following, we in general define a collection of particular objects used in learning domain and some operations useful for management of the learning objects.

Object. An object can be any component or unit in a learning domain. A learning domain is a collection of any educational, training materials (objects) or information. The learning domain also contains one or several description structures for the organisation and use of the learning objects. Information providers are supplying learning objects for a learning domain along with some structures (metadata) for learning objects. Information consumers will obtain the learning objects, which may be reconstructed and integrated based on the consumers' requirements and needs.

An object is described by a set of attributes and relationships. The relationships in the set relate the object to other objects. A subset of the set of attributes and relationships may be able to uniquely identify the object, called identifier.

Course. A course in a learning domain is an object. It has as its attributes course number, course name, course length in numbers of lessons, course subject, course price, etc. It is related to a number of other objects: It is advised by a teacher. It is located at a (maybe virtual) classroom. It is paid through a department. The course is uniquely identified by its attribute – course number.

Lesson. A lesson is an object. A lesson is a part of a course. A lesson may inherit a number of attributes from the course as whole or part of its attributes, like course number. A lesson may inherit a number of relationships from the course as whole or part of its relationships, like being guided by a teacher.

Carrier. A carrier is in general a medium for information, e.g., in form of book, CD, tape, etc. A carrier has its attributes, including carrier type, carrier label, etc. The idea here to maintain the concept carrier is to separate the information content with the information carrier.

Atomic component. An atomic component is an object. An atomic component is the component that contains no more objects as its components.

Decomposition is a function (or operation). The decomposition function maps one object to a set of objects. Each object in the set is considered to be a part of the mapped object.

Integration is a function (or operation). This function maps a set of objects to one object.

4. A component-based structure for learning object management

As we know already, the web information is of high diversity. It is also true for the materials that are used for any education products. As we described previously, an education product can be a book, a lecture, a videotape, a CD-ROM, or a combination of these. The appearance of these materials can be a textual document, a piece of music, a fragment of video, and so on. Through the metadata model proposed in the previous section, we can describe them as components. That is, we consider each representation form for these products to one component. For instance, a textual document is a component and a tape is another component. A lecture can be an object that comprises of these two components, the document and the tape. In this section first we discuss an architecture for a learning system in order to gain an overview and then we will describe two internal structures for learning object representation.

An architecture for Learning Object components

We assume that a learning system is associated with a group of information providers and a group of information consumers. The information providers include, e.g., TV programs, department lectures, seminars, etc. The information consumers can be students, employees, TV watchers, etc. The system contains three main components: a goal-match mechanism, a data repository for managing learning objects, and an intelligent interface for the information providers and consumers.

The goal-match mechanism deals with the users' requirements on, e.g., what she or he would like to be included in a course. In general, a user will provide a number of demands, along with the user's profile and preferences. The goal-match mechanism will compare the user's requirements with the description items (metadata content) of the learning objects stored and maintained in the repository. Once a match is found, the matched course will be presented to the user.

The intelligent interface accomplishes three tasks. Firstly, it allows the information providers to easily put their learning objects to a right place in the repository, along with a set of suitable metadata descriptions of the objects. Secondly, it allows the consumers to suggest their requirements and find out what learning objects they need. Thirdly, it should be able to manage the changes, modifications, and maintenance of the metadata models, schemas, and contents.

The repository component maintains a set of metadata schemas, where learning objects are organised in the hierarchical structure and network structure. Each object or component is described by its relationships to other components. The components can be a paragraph, an image, a piece of music, a segment of movie, etc. A granularity for the elements in the repository has been defined when an information provider put its information to the repository. So the elements in the repository are assumed to be very fundamental that they are generally not allowed to decompose.

The Internal Structure for the Metadata Model

In this section, we will describe two structures for learning object components. We consider that every learning object is a set of components organised in terms of some rules or user profiles. Components are connected to each other in a hierarchical or network structure. For each structure we have a schema corresponding to a real application situation.

Hierarchical Structure

In the hierarchical structure of learning object components, we maintain three types of relationships: part-of, is-a, and sibling. These three types of relationships are used for defining basic decompositions of learning objects. The part-of relationship indicates the fact that an object is a part of another object. This relationship is also called whole-part relationship. This definition is very useful when we attempt to group a number of components from different learning objects to generate a new learning object. For instance, a table may consist of a flat and four legs.

The is-a relationship indicates that a concept is a subset of another concept. One major advantage of this relationship is that all the attributes of a super-object can be inherited by its sub-objects. For instance, cat is an animal.

The sibling relationship is introduced to consider the relationships of components at horizontal level. Two objects have a sibling relationship if both belong to a same object. E.g., in a book, the chapter 1 and the chapter 5 have a sibling relationship because they both are parts of the book. This relationship is intended to support to improve searchability.

Network structure

In the network structure, learning objects are related to each other based on their natural relationships. For example, a course may be related to teachers, locations, payment methods, etc. That means, if we take the course, the teacher, the location, and the payment method, as some nodes in the network structure, the relationships established between them will be given_by, located_at, paid_in, etc.

These relationships are not those we discussed in the hierarchical structure. They are more natural and to expose the semantics of the collocations among the objects. They provide better circumstances for the identification of the objects of interest. That is the idea that we can understand objects by their contexts. The advantages of maintaining the network structure include: a) easy capture of semantics of learning materials without well defined structures, b) easy location of the input information in the right slots in our learning domain repository, and c) together with the hierarchical structure quick and better re-composition of learning objects for users' needs.

5. RdfClient: A prototype to represent Learning Objects

RdfClient is a tool for viewing a RDF (Resource Description Framework) schema or document [DK]. It is intended to support users to authorise a metadata-modelling schema and document in terms of the RDF syntax and constructs. The RDF metadata framework is defined for the representation and communication of various information items from the Internet or Intranet. RdfClient can be considered among few tools for the metadata information description to support particularly a hierarchical structure for the description and management of digital documents.

RdfClient is implemented in Java as a stand-alone tool or a browser based Java Applet. It uses XML to interpret any RDF codes and can as well convert documents in XSL to XML or RDF.

Based on the metadata model discussed previously, we use RdfClient for illustration of the structure and the use of learning objects. We consider it very useful to include the operations such as decomposition and re-composition. In accordance with the learning object methods, we maintain a hierarchical structure for the learning objects. By using the tool, we maintain three main components, a hierarchical structure of metadata information for documents, a descriptive information of the metadata schema especially designed for the documents, and a group of binary elements of relationships (called property type in RDF) and values for the metadata information of the documents.

Each learning object, no matter it is an entity type at the schema level or a concrete instance at the instance level, can be described by a set of attribute values along with their attributes. For example, the published book “Java in a Nutshell” has a unique number (ISBN: 1-56592-262-X), is authored by the author “David Flanagan”, and is published by the publisher “O’Reilly”. In our metadata model language, this example can be written as:

```
<eduProd:book ID=”bookInternalNumber”>
        <ISBN:value>1-56592-262-X</ISBN:value>
        <dc:title>Java in a Nutshell</>
        <dc:publisher>O’Reilly</>
        <dc:author>David Flabagan</>
</eduProd:book>
```

Here dc stands for Dublin Core (a metadata description method for publications maintained in libraries).

At present, we are extending RdfClient by including a syntactic representation of the network structure, where a graphical description for the structure is necessary. Since the RDF Model provides support to the description of natural relationships among learning objects, i.e., the contextual description, the network structure in the learning system can be represented in a specific RDF schema.

6. Conclusion

The report has described a particular metadata model for learning domain and an implementation for the model in a prototype.

Out next step for the project, other than further development and refinement of the metadata model and prototype, we will tackle the following three problems.

1. Establishment of a network structure for representing learning objects, its coupling with the hierarchical structure, and mapping both structures to a repository.

2. Design of a graphical user interface for authoring learning objects and their relationships, as well as operations on the objects.

3. Design of a goal-matching mechanism. For example, if we see the users' requirements on a particular course, which need to be built out of the other course components, how to compose such course from a set of other existing course components can be considered to be a process of goal matching. This issue is extremely important when hundreds of courses are managed in the repository and about ten requirements proposed from a user for a particular course. The issue leads to what strategies can be adopted to match the user's goals and automatically or semi-automatically find the right course.

Further description of the idea of goal matching still needs more investigation not only from metadata domain but also from enterprise modelling area.

References

[1] Dublin Core http://purl.oclc.org/dc/

[2] Dag Ekengren, RDF for Multimedia Broker – Metadata on the Web, Master Thesis, The Royal Institute of Technology in Sweden (1999)

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

[4] Mats B. Andersson, Learning Domain – Education on Demand (in Swedish), CID Report, Centrum för Användarorienterad IT Design, The Royal Institute of Technology in Sweden, Sweden. (1999)

[5] Mats B. Andersson, Web Resources as Integrated part in Learning, LOUIS Report, SITI, Sweden. (1998)

[6] Resource Description Framework, http://www.w3.org/RDF/

[7] William Song, WDM: A Web Document Model and Its Supporting Web Document Analyzer, AIS'98, Baltimore, MD, USA (1998)

[8] William Song, A three Dimension Structure for Learning Object Management, LOUIS Report, SITI, Sweden (1999)

[9] William Song, OFEM: An Approach towards Detection and Reduction of Objectives Inconsistencies in the International Database Engineering and Applications Symposium, (IDEAS'97), Montreal, Canada (1997)