GRAPPLE

D7.2a Version: 1.0

Data models and related documentation

Document Type	Deliverable
Editor(s):	Lucia Oneto (GILABS)
Author(s):	Lucia Oneto (GILABS), Dominikus Heckmann (DFKI)
Internal Reviewers:	Olga de Troyer (VUB), Martin Harrigan (TCD)
Work Package:	WP7
Due Date:	31-10-2008
Version:	1.0
Version Date:	23-02-2009
Total number of pages:	19

Abstract: This document contains the first version for a set of harmonized data models to be exchanged within the GRAPPLE system. The focus is set on syntactical and semantic aspects of user model exchange. The central (syntactical) RDF application is called "GrappleStatements". The central (semantic) OWL vocabulary is called "GrappleOntology". The idea is to represent and externalize the knowledge of the individual systems with modern, reusable Semantic Web technologies.

Keyword list: Knowledge representation, data model, semantic web, ontologies, GrappleStatement, Unique Identification, Syntax

Summary

This document presents a first proposal for a set of harmonized data models to be exchanged within the GRAPPLE system, able to guarantee interoperability within the operational infrastructure and to harmonize the data flows between the participating systems. The main question that we look at in this deliverable is "how do we identify and represent partial information units that are developed decentralized by a variety of systems"?

This first version of the data models introduces how the new approach of the semantic web layers can be applied to the GRAPPLE data modelling task. The three important parts "identification", "representation" and "semantics" are collected into the three challenges: 1) GrappleNames, 2) GrappleStatement and 3) GrappleOntology.

The *GrappleNames* face the problem of the non existing unique name assumption in GRAPPLE: everything, every concept, every user can have several names and identifiers by different subsystems. Since the necessary mapping in order to share information can technically not be done by the system itself, we need to form a group of consensus for naming resources under the GRAPPLE project partners. The ideas behind the GrappleNames are discussed in section 2.

The *GrappleStatements* face the problem of the knowledge representation task in a distributed heterogeneous environment of independent systems where stability and formality is needed. Formality allows for sharing of information. GrappleStatements form a compromise of uniformity and flexibility, thus several versions of GrappleStatements are defined in section 3.

The *GrappleOntology* is an attempt to define the semantics of the concepts that occur within statements, and that are not defined in any other existing ontology in an appropriate form. Since the GRAPPLE project tries to solve real world problems rather than develop a light weight test-tube demonstrator, we are confronted with a difficult "openness" for knowledge representation and data modelling: we may be confronted with conflicting statements, and we have to apply the "Open World Assumption" and not the "Closed World Assumption". This means that missing statements are just unknown rather than false. All these issues lead to the need of reusing existing ontologies and gradually building a new ontology that is presented in section 4.

The great challenge of data modelling for sharing knowledge between LMS and ALE completely under the influence of the semantic web is relatively new in its completeness and not straight forward at all, since several tools for the semantic web are still under development. We need to do pioneer work which however promises to come up with a better solution for integration and interoperability. The task of this work package is to carefully select and collect data models for GRAPPLE and apply the "Semantic Web idea" to them.

Authors

Person	Email	Partner code
Lucia Oneto	l.oneto@giuntilabs.com	GILABS
Dominikus Heckmann	heckmann@dfki.de	DFKI

Table of Contents

Tables and Figures

List of Acronyms and Abbreviations

ALE	Adaptive Learning Environment
CAM	Conceptual Adaptation Model
CRT	Concept Relationship Type
DM	Domain Model
GAL	Grapple Adaptation Language
GQ	Grapple Query
GR	Grapple Rule
GRAPPLE	Generic Responsive Adaptive Personalized Learning Environment
GS	Grapple Statement
GUMO	General User Model Ontology
IMS TI	IMS Tools Interoperability
IT	Information Technology
KIF	Knowledge Interchange Format
LMS	Learning Management System
LOM	Learning Object Metadata
OWL	Web Ontology Language
PALCS	Personal and Adaptive Learning Control Service
RDF	Resource Description Framework
SCORM	Sharable Content Object Reference Model
SPARQL	Simple Protocol and RDF Query Language
SUMO	Suggested Upper Merged Ontology
SWRL	Semantic Web Rule Language
UM	User Model
UML	Unified Modelling Language
URI	Uniform Resource Identifier
UTF	Unicode Transformation Format
WP	Work Package
XML	eXtensible Markup Language

1 Introduction

The GRAPPLE project will enable the cooperation of adaptive learning environments (ALE) and learning management systems (LMS). In order to reach this challenging task, suitable data models for interoperability are needed.

1.1 Task and Deliverable Description

T 7.2 Data modelling for interoperability *(DFKI, GILABS, LUH)*

In order to guarantee interoperability within the operational infrastructure and to harmonize the data flows in the GRAPPLE system suitable data models need to be defined and implemented, to be used by the subsystems to exchange information through their interfaces (e.g. Partial-User-Model-Exchange-Format). The design of data models will be carried out by means of a UML-based approach. For a rapid prototyping of the models and their early evaluation, suitable UML tools will be used, which allows for the semi-automatic generation of XML-like (e.g. UserML) schemas from class diagrams. This task will strongly rely on the outcomes from WP6.

D7.2.a Data models and related documentation - first version (GILABS, M9): *This document will contain a first proposal for a set of harmonized data models to be exchanged within the GRAPPLE system.*

1.2 Design decisions and Course of action

This deliverable describes the current status of our attempts in this early stage of the GRAPPLE project. This "first version" D7.2a) will be followed by two planned updates D7.2.b) and D7.2.c). For a more detailed description of the described data models we refer to the deliverable D2.1: "Definition of an appropriate User profile format", and deliverable D6.1:"Design of a distributed user model (storage) platform and retrieval service", where they mostly stem from.

One important issue with adaptivity, personalization and user modelling is the knowledge representation of information about the users: the so called user model. In learning environments the users are often called learners, while the models are called learner models. An early partial-user-model-exchange format has been defined in XML by (Heckmann D. , Ubiquitous User Modeling, 2006) which is called UserML.

The developed data model of "Grapple Statements" is based on UserML. Our research is also based on (Carmagnola & Dimitrova, 2008), (Berkovsky, et al., 2007) and (Heckmann, Schwartz, Brandherm, Schmitz, & Wilamowitz-Moellendorff, 2005). These papers gradually develop the idea to apply Semantic Web technologies to the user modelling domain.

Having started with the task and deliverable description of data modelling and interoperability, we came to the insight that it could be an advantage to add two topics in addition to the initial task T7.2 and deliverable description:

- One is the topic of a uniform naming and identification guideline for all GRAPPLE modules and partner systems.

And instead of using a UML-based approach, we decided to apply the semantic web approach, which leads to a complete decoupling of the syntax and the semantics of the data models.

- The content semantics is moved into an externalized, commonly agreed set of ontologies.

The management, collection, selection, or even new development of ontologies for GRAPPLE can be settled within Task 7.2 since these ontologies are the most important part of the data models in the semantic web era.

1.3 Overview of the approach

The semantic web technologies as well as the service oriented architectures are about to become commonly accepted standards in academic and industrial settings. In Figure 1 we present the individual layers like

"Unicode + URI", "XML + XML schema", or "RDF + RDF schema" with the corresponding conceptual modules of the deliverable D7.2a) that form the initial main subtasks to solve:

- "GrappleNames" → uniform naming, identification and special character handling
- "GrappleStatements" → uniform knowledge representation
- "GrappleOntologies" → uniform semantics

Two further identified subtasks to solve that are not yet elaborated:

- "GrappleQueries" → uniform user model requests
- "GrappleRules" → uniform inference rules

These five terms are introduced to denote the modules of the data models for interoperability. More specifically, they describe the current separation of the whole deliverable into independent components.

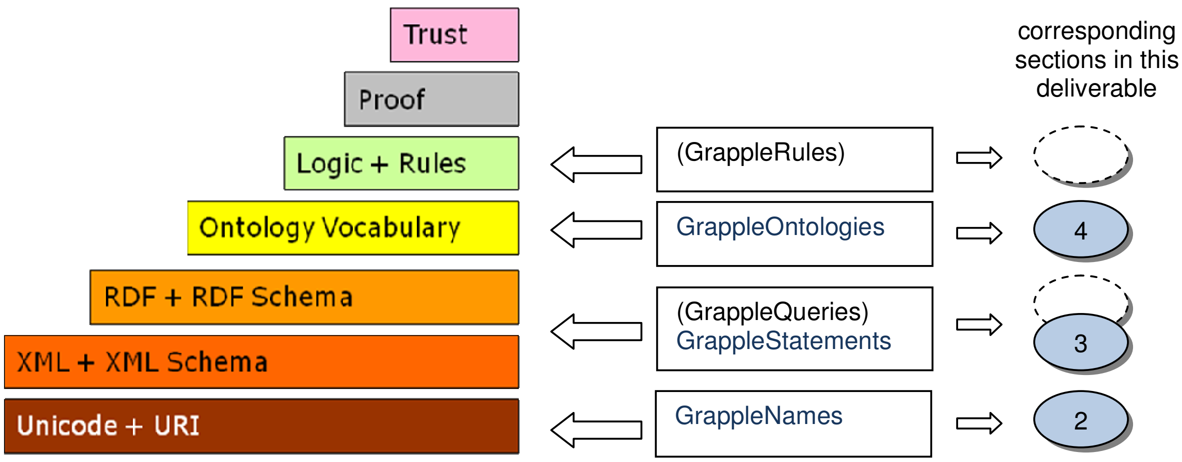

Figure 1: Layered Architecture of the Semantic Web according to Tim Berners-Lee together with the conceptual modules of deliverable D7.2a).

In section two the GrappleNames are defined and introduced. In section three we look at the syntax of GrappleStatements and GrappleQueries. The semantics of the concepts will be defined in a variety of ontologies that are available for all systems world -wide via the semantic web. They are discussed in section four.

2 Identification & Interoperability

In order to enable data models for interoperability, we first have to ensure that all systems (LMS and ALE) identify the entities and resources with the same identifiers or names. We need to form a group of consensus for naming resources under the GRAPPLE project partners. The data models can also play a central role or the other way round are influenced by the identification alignment.

2.1 GrappleNames: Identification mechanism in GRAPPLE's data models

If we want to combine several so far independent systems that adapt to their users, we have to ensure that these systems don't use same names for different things and - vice versa - that they don't use different names for same things. We are confronted with the task of distributed and decentralized naming and identification.

2.1.1 Identification of users

We can apply the single sign-on as described in D7.1a) to identify the users uniquely within the LMS and the ALE. The Shibboleth System (Internet2, 2008) is a standards-based, open source software package for web single sign-on across or within organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner.

2.1.2 Identification of learning objects, courses and general concepts

The identification of learning objects, courses and general concepts is recommended to be done in general URIs.

2.1.3 Identification of GrappleStatements and GrappleQueries

We suggest using the following form to identify GrappleStatements, GrappleQueries or GrappleRules:

A string consisting of the following four parts:

Prefix-Date-Time-Random

With the corresponding syntax:

Prefix: XX (consisting of two letters)

Date: YYMMDD (consisting of six digits)

Time: HHMMSS (consisting of six digits)

Random: NNNNNN (consisting of six digits)

If we all agree to calculate the Date and Time according to a commonly agreed time-zone we come up with a temporal ordering of all new elements. One advantage of this temporal ordering of the created identifiers is that is later allows to simply apply the recency-filter (prefer newer statements before older statements) in the conflict resolution phase (inspired by deliverable D6.3). A second advantage is the indirect support for a distributed uncoordinated creation of new identifiers.

We define the prefix “GS” for GrappleStatements, the prefix “GQ” for GrappleQueries, and the prefix “GR” for GrappleRules.

Examples of such GRAPPLE identifiers (or GrappleNames):

GS-081031-124402-123456

GQ-081031-124635-048012

GR-081031-204133-404402

Two possible alternatives that we do not recommend for GRAPPLE:

1) **encode the creator of each identifier into the identifier itself**
   Because of the distributed setting with the GRAPPLE project, one could think of an alternative to encode the creator of each identifier into the identifier itself, in order not to create same identifiers twice. However, due to privacy regulations it is better not to do so. Otherwise the creator of a piece of information will not be able to hide his role any more.
2) **use the existing GUID standard**
   The second alternative is the already existing and implemented GUID[1] approach (Globally Unique Identifier). In this identification approach a huge random number is generated that is almost always globally unique. The advantage of this alternative is that the implementation already exists. However, the temporal ordering would be lost. And further more the GUID identifiers are much longer than the suggested new GRAPPLE identifiers.

[1] http://en.wikipedia.org/wiki/Globally_Unique_Identifier

2.2 GrappleEncoding : Special character encoding in GRAPPLE

Since there are partners from several European countries in our EU-founded project, we also have to look at other languages than English. Especially the learning material and the domain models are expected to be in the native languages of the learners. This leads to an increased need for handling special characters. The existing solution is "Unicode". However, Unicode is complex and offers a variety of different representations.

Percent-encoding for example, see (Berners-Lee, 2005) is a mechanism for encoding information in a Uniform Resource Identifier (URI) under certain circumstances. It is widely used on the Web and it is a good standard to start with. However, one disadvantage of percent-encoding is for example the fact that it is based on UTF8 which could become very long.

A possible alternative could be the so called "underscore encoding" which is currently under development in the UbisWorld.org project[2]. However, since this is still in a too early stage to be used in GRAPPLE, we currently recommend using percent-encoding for all special characters in URIs.

3 Syntax & Interoperability

The two main data model types that we envision so far for GRAPPLE are statements and queries that lead to "GrappleStatements" and to "GrappleQueries"[3]. Each of them will have several variations to remain flexible for different applications within the complex GRAPPLE system.

3.1 GrappleStatements

Grapple statements have been developed in work package WP2 in task 2.1 with the title "Definition of an ontology-based user-model format". We describe the current version which is still under discussion in GRAPPLE. Grapple statements are based on the resource description framework (RDF)[4] and situational statements (Heckmann D. , Introducing "Situational Statements" as an integrating Data Structure for User Modeling, Context-Awareness and Resource-Adaptive Computing, 2003).

We define the namespace "grapple" and the namespace "gc" in the following equal way:

xmlns:grapple="http://www.grapple-project.org/core/"

xmlns:gc="http://www.grapple-project.org/core/"

In the subsequent definitions all V_x denote variables and placeholders for the concrete values. These values will be pointers to ontologies (realized by URIs), the local identifiers as used by the participating partners, or simply textual strings if their semantics is obvious and commonly agreed on.

3.1.1 GrappleStatements (basic triple form)

The syntactical structure of GrappleStatements in basic triple form look very similar to RDF triples:

<grapple:statement>

<grapple:subject> $V_{subject}$ </grapple:subject>

<grapple:predicate> $V_{predicate}$ </grapple:predicate>

[2] http://www.ubisworld.org

[3] GrappleQueries are only mentioned for completeness here. They are not yet analyzed and defined in this deliverable.

[4] Resource Description Framework (RDF): http://www.w3.org/RDF/

```
        <grapple:object> V_object </grapple:object>
</grapple:statement>
```

The difference between RDF triples and GrappleStatements lies in the values of the V_x since in RDF the V_x always needs to be URIs for the subject and the predicate, while in GrappleStatement local identifiers are also allowed.

The basic triple form in short-cut XML is defined in the following way: (Note that for user model data we introduced the variation “user” to name and differentiate the “subject” explicitly):

```
<statement>
        <user> V_user </user>
        <predicate> V_predicate </predicate>
        <object> V_object </object>
</statement>
```

3.1.1.1 Examples

Some examples to show the basic usage of GrappleStatements: , with the three following synonyms defined: s := subject/user, p:=predicate, o=object:

3.1.1.1.1 „Mary is female“

```
<statement>
        <s> Mary </s>
        <p> hasGender </p>
        <o> female </o>
</statement>
```

3.1.1.1.2 „Mary is a teacher“

```
<statement>
        <s> Mary </s>
        <p> hasRole </p>
        <o> teacher </o>
</statement>
```

3.1.1.1.3 „Mary likes chemistry“

```
<statement>
        <s> Mary </s>
        <p> likes </p>
        <o> chemistry </o>
</statement>
```

With the information from these simple statements, GRAPPLE could already start adaptation and present the learning material adapted to gender, adapted to the current role and the preferences of Mary. However, the role and interests of a person changes over time. Thus we need to handle temporal aspects in a notion of expiracy of information. Apart from temporal aspects, we also have to handle privacy aspects. Furthermore, statements in triple form are always meant to be facts, which means that their semantic is true.

3.1.2 GrappleStatements (extended form)

If we look at the statement: *"Mary likes chemistry"*, we could ask ourselves which meta data is important for interoperability to be solved with GrappleStatements?

- When and where is the statement valid?
- Who claims this and which explanation is given?
- What is the evidence and the confidence?
- What will be done with this personal information?
- When will this information be deleted?
- Who is the owner of this information?
- What are the privacy settings?
- How can the statement be uniquely identified?
- Can the statement be grouped with others?

All this information can be seen as one meta-level above the actual content of the statement. Thus we could roughly divide the information into the "main" part of the "statement" and into the "meta" part of the statement. This view is similar to reification in RDF. However, the new meta-attributes can directly be added to the XML-like syntax.

The so identified statements with possible meta-attributes have the following form:

<statement>

  <subject> $V_{subject}$ **</subject>**

  <predicate> $V_{predicate}$ **</predicate>**

  <object> V_{object} **</object>**

  <$V_{\text{meta-attribute-1}}$> $V_{meta\text{-}value\text{-}1}$ </$V_{\text{meta-attribute-1}}$>

  …

  <$V_{\text{meta-predicate-n}}$> $V_{meta\text{-}object\text{-}n}$ </$V_{\text{meta-predicate-n}}$>

</statement>

Remark: In task D2.1 of work package D2 we are currently collecting a preselected set of so called meta-attributes. One prominent source for this "GrappleCore" metadata set will be the Dublin Core[5] metadata set. A member of the Dublin core metadata set is for example "dc:creator" with the URI http://purl.org/dc/terms/creator which defines the information of "who has created this statement".

3.1.2.1 Examples

3.1.2.1.1 „Peter says that Mary likes chemistry and he is highly confident that this is true"

<statement>

  <s> Mary </s>

  <p> likes </p>

  <o> chemistry </o>

  <dc:creator> Peter </dc:creator>

[5] Dublin Core homepage: http://dublincore.org/

```
        <dc:identifier> GS-081031-123820-149503 </dc:identifier>
        <gc:confidence> high </gc:confidence>
</statement>
```

Note, in analogy to the namespace “dc:” for Dublin core, we define the namespace “gc:” for grapple core with

```
xmlns:gc=“http://www.grapple-preoject.org/core/”
```

In order to refer to such meta-statements, we often need to identify the statement uniquely. Therefore, we introduce the new tag <i> for the identifier and define the semantics <i> := <dc:identifier>. A complete list of semantically defined GrappleCore elements is expected for the deliverable D6.1b) – Distributed user model platform & retrieval service implementation.

4 Semantics & Interoperability

The semantic web approach – that we decided to apply to the GRAPPLE system – defines the semantic of the data models in ontologies.

Ontologies provide a shared and common understanding of a domain that can be communicated between people and heterogeneous and widely spread application systems, as pointed out in (Fensel, 2001). Since ontologies have been developed and investigated in artificial intelligence to facilitate knowledge sharing and reuse, they should form the central point of interest for the task of exchanging data models for interoperability.

Ontology development is a difficult task, often highly underestimated. Therefore, we recommend to re-use as much as possible from already existing ontologies. However, missing concepts and instances have to be defined and could be arranged into a so called new “GrappleOntology”. A problem that is inherent in semantic web and that we also have to cope with is that most concepts are defined in several different ontologies at the same time. Thus, a selection of ontologies has to be undertaken. In the following subsections we discuss some preferred examples.

The RDF standard has been proposed as a data model for representing meta data. Nonetheless, the web ontology language OWL has more facilities for expressing semantics, (McGuinness & van Harmelen, 2003), and it has a greater machine interpretability than XML and RDF. It adds more vocabulary for describing properties and classes. OWL can be used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. Thus, OWL is our choice for the representation of user model terms and their interrelationships.

4.1 Related Work: existing ontologies for possible re-use

In this section candidate ontologies are discussed that define top-level and mid-level ontologies that could be interesting for partial re-use in the GRAPPLE system.

4.1.1 Sumo = Suggested Upper Merged Ontology

The SUMO[6] ontology (Suggested Upper Merged Ontology) was created as part of the *IEEE Standard Upper Ontology Working Group*. The goal of this working group is to develop a standard upper ontology that will promote data interoperability, information search and retrieval, automated inferencing, and natural language processing. SUMO has been translated into various representation formats, but the language of development is a variant of *Knowledge Interchange Format* KIF, which is a version of the first-order predicate calculus. An upper ontology is limited to concepts that are meta, generic, abstract or philosophical, and hence are general enough to address at a high level a broad range of domain areas. Concepts specific to particular domains are not included in an upper ontology, but such an ontology does provide a structure upon which ontologies for specific domains can be constructed. Figure 2 shows the top level concepts in SUMO.

[6] SUMO: http://www.ontologyportal.org/

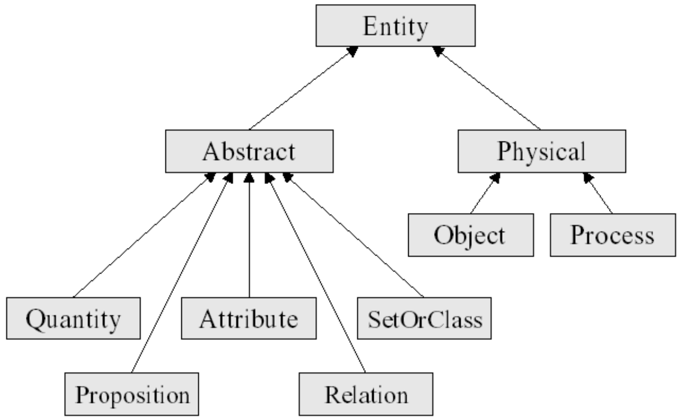

Figure 2: Top-level concepts in SUMO.

It uses the term entity as root which is well established in database research. The main distinction is drawn between abstract and physical entities. Physical entities are further distinguished as objects and processes, etc. SUMO is a collection of well-defined and well-documented concepts, interconnected into semantic network and accompanied by a number of axioms. The axioms mostly reflect common-sense notions that are generally recognized among the concepts. The concepts range from very general ones, such as Quantity as shown in Figure 2, to very specific, such as Emotional State as shown (together with its superclasses) in Figure 3(a) . Subclasses of a class are usually mutually exclusive, i.e. they do not share common instances. For example, nothing can be both an abstract entity and a physical entity, neither both an object entity and a process entity, which is explicitly specified in SUMO. However, some classes can have multiple superclasses. For example, a Human is both an Organism (which is itself a subclass of Organic Object) and a subclass of Agent and a Cognitive Agent which is defined as an entity with the ability to reason as shown in Figure 3(b).

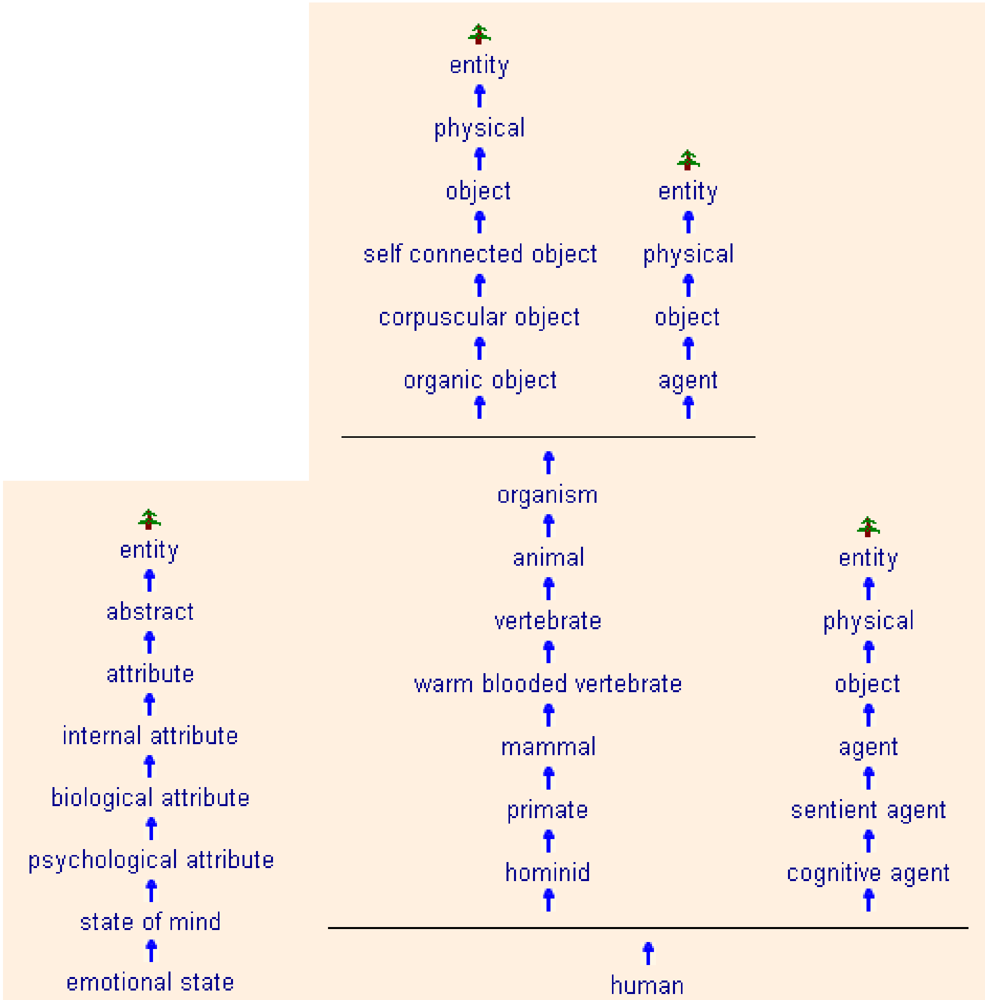

Figure 3: (a) Superclasses of Emotional State as shown in the SUMO browser and (b) multiple inherited superclasses of the Human concept.

Multiple inheritance is realized by tree copying. SUMO is intended as a domain-independent knowledge base for designing domain ontologies and is therefore very interesting for the design, integration and extension of GRAPPLE ontologies. The combination of ontological concepts and natural language words has partially been established by connecting SUMO to the *WordNet* lexicon.

4.1.2 Gumo = General User Model Ontology

GUMO[7] is a huge mid level ontology for user models and user context dimensions. This ontology is represented in the semantic web language OWL and thus via internet available for all user-adaptive systems. The initial idea is the simplification for exchanging user model data.
GUMO presents a collection of user's dimensions that are modelled within user-adaptive systems like the *user's heart beat*, the *user's age*, the *user's current position*, the *user's birthplace* or the *user's ability to swim*. Furthermore, the modelling of the user's interests and preferences like *reading poems*, *playing adventure games* or *drinking certain French Bordeaux wines* is enabled. The complete ontology can be inspected with a foldable tree browser at http://www.ubisworld.org.

[7] GUMO website: http://www.gumo.org

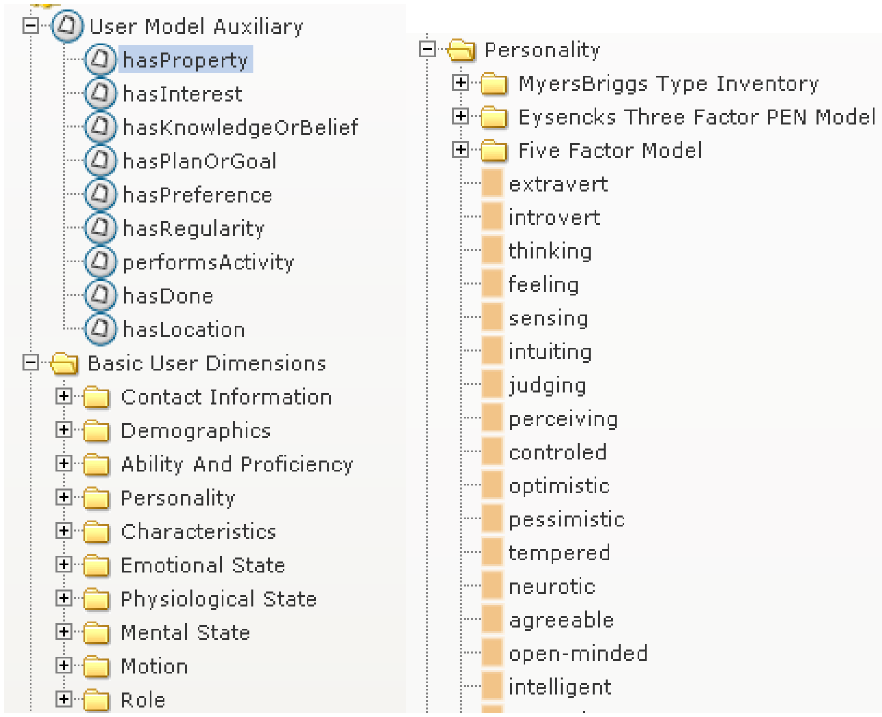

Figure 4: GUMO - user model auxiliaries, some basic user dimensions and personality traits.

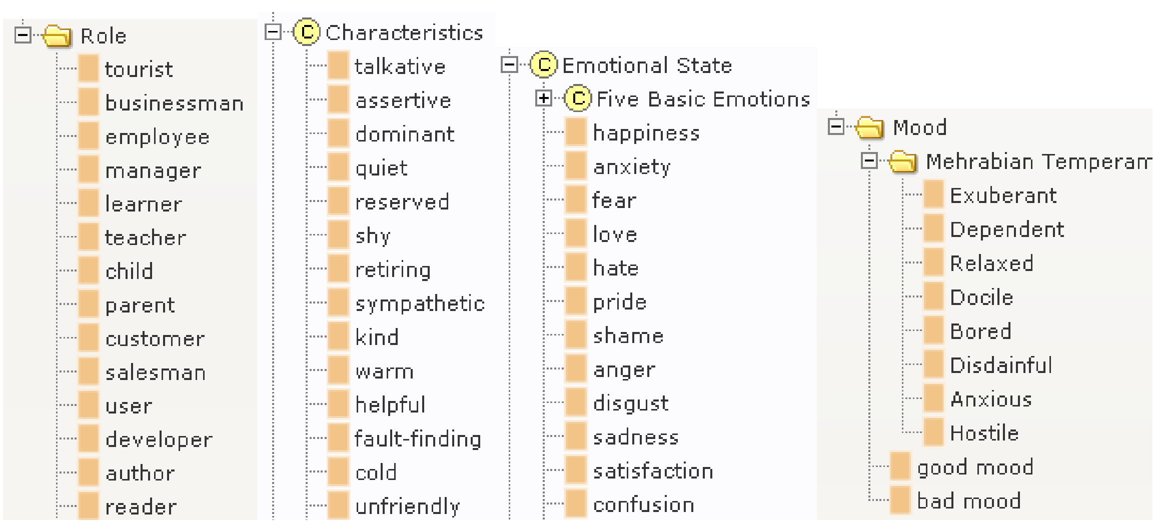

Figure 5: GUMO - user role, user characteristics, emotional state and mood.

4.1.3 WordNet

WordNet[8] is a semantic lexicon for the English language. It groups words into sets of synonyms called synsets, provides short, general definitions, and records the various semantic relations between these synonym sets. The purpose is twofold: to produce a combination of dictionary and thesaurus that is more intuitively usable, and to support automatic text analysis and artificial intelligence applications.

[8] WordNet homepage: http://wordnet.princeton.edu/, description from http://en.wikipedia.org/wiki/Wordnet

That is the point where it might become very interesting for GRAPPLE's interoperability task: if different systems use synonym words to define the semantics of the described GrappleStatements, the defined synonym relations might be used for automatic alignment.

4.1.4 DBpedia.org

DBpedia[9] is a community effort to extract structured information from Wikipedia and to make this information available on the Web. DBpedia allows users to ask expressive queries against Wikipedia and to interlink other datasets on the Web with DBpedia data.

Interesting for GRAPPLE is the enormous coverage of available concepts on the Web: as of September 2008, the DBpedia dataset describes almost 2.5 million "things", including at least 108,000 persons, 392,000 places, 57,000 music albums, and 36,000 films. Thus in certain field this ontology could serve as intermediate domain ontology.

Also very important for GRAPPLE is the fact that the DBpedia project already uses the Resource Description Framework as a flexible data model for representing extracted information. Which means that the entities can be used directly through their URIs without any further pre-processing.

4.1.5 Ontologies for technology enhanced learning

It could be interesting or necessary for GRAPPLE to transform existing approaches for interoperability in the technology enhanced learning environment into the semantic web ontology languages. Candidates are for example:

- LOM = Learning Objects Metadata[10]
- SCORM = Sharable Content Object Reference Model [11]
- IMS TI = Global Learning Consortium Tools Interoperability Guidelines

The IEEE LOM standard[12] conceptual data schema lists all the meta-data elements in tabular format. To complement this table, Figure 6 presents a graphical illustration of the elements in the data schema, which shows how the elements are divided into nine top level categories: General, Life Cycle, Meta-Metadata, Technical, Educational, Rights, Relation, Annotation, and Classification.

---

[9] DBpedia website: http://dbpedia.org, description from: http://en.wikipedia.org/wiki/DBpedia

[10] LOM: http://en.wikipedia.org/wiki/Learning_object_metadata

[11] SCORM: http://en.wikipedia.org/wiki/SCORM

[12] IMS Global: http://www.imsglobal.org/metadata/

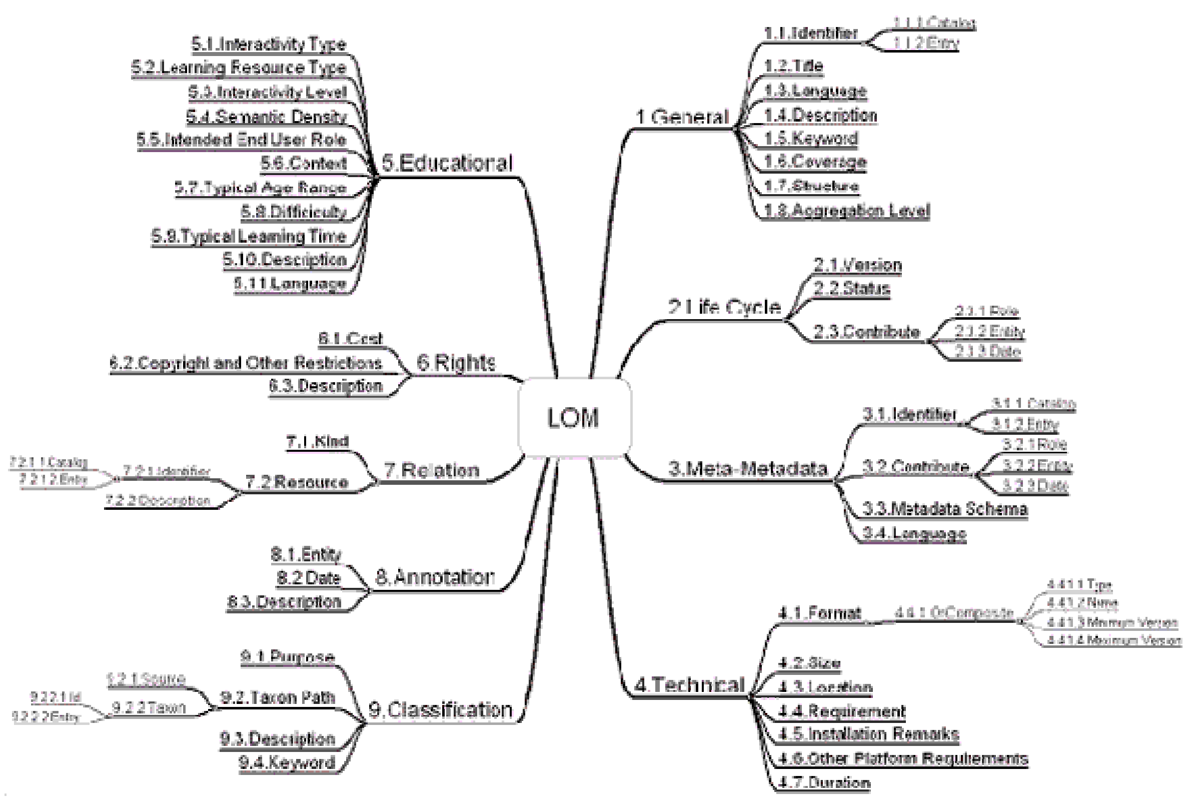

Figure 6: Standard for Learning Object Metadata Copyright © IMS Global Learning Consortium 2006. All Rights Reserved.

4.2 A new GrappleOntology?

The so far discussed existing ontologies could be partially used for GRAPPLE's semantic. However, since none of them perfectly fits to the task of supporting interoperability between ALE and LMSs, the design and development of a new GrappleOntology could become interesting.

A "GrappleOntology" could focus on collecting and developing all aspects of LMS and ALE related aspects that are not part of any other well structured available ontology. Since this issue and idea is new and has not been included in the initial description of work, further discussions and decisions have to be made in the near future.

As nucleus of the new GrappleOntology we could follow the approach in Task 5.2 "Conversion Models between GRAPPLE and LMSs" in work package 5. They try to identify a set of minimal elements necessary to characterize the GRAPPLE ALE data model that needs to interact with the existing LMSs (Moodle, Sakai, Claroline, learn eXact, and IMC Clix) data models.

4.2.1 GRAPPLE data model

The minimal GRAPPLE data model as suggested in D5.2 - Conversion models between GRAPPLE and LMSs has been achieved by analyzing the data that the existing LMSs can pass to the GRAPPLE system:

Attended Courses

The learner information that consists of the education/training, work and service (military, community, voluntary, etc.) record and products (excluding formal awards). This information may include the descriptions of the courses undertaken and the records of the corresponding evaluation.

- *UserX hasAttended CourseXY*

- *UserX isEnrolled CourseXY*

Roles

Roles describe a group of interest to the Learning Management environment. There are many types of groups that may be shared between systems.

- *Teacher*
- *Tutor*
- *Learner*
- *System Administrator*

Competences/Skills

This learner information consists of the descriptions of the skills the learner has acquired. These skills may be associated with some formal or informal training or work history and formal awards. The corresponding level of competency may also be defined.

- *UserX has CompetenceY with LevelZ*

Personal User data

The learner information that contains all of the data for a specific individual or organisation. This includes data such as: names, addresses, contact information, demographics and agent.

- *Username*
- *Password*
- *Last Name*
- *First Name*
- *Email*
- *Language*
- *Gender*
- *Date of birth*
- *Street*
- *Town*
- *Postal code*
- *Region*
- *Country*

Accessibility Info.

The learner information that consists of the cognitive, technical and physical preferences for the learner, their language capabilities, disability and eligibilities.

The attended courses will be part of the domain model, however the “Roles” and the “Personal User data” could be the start for the new GRAPPLE naming convention and the start of the GrappleOntology.

5 Conclusion

This first version of the data models and related documentation introduces how the new approach of the semantic web layers can be applied to the GRAPPLE data modelling task. The three important parts identification, representation and semantics are collected into the three challenges:

- GrappleNames
- GrappleStatements
- GrappleOntologies

The *GrappleNames* face the problem of the non existing unique name assumption in GRAPPLE: everything, every concept, every user can have several names and identifiers by different subsystems. Since the necessary mapping in order to share information can technically not be done by the system itself, we need to form a group of consensus for naming resources under the GRAPPLE project partners. The ideas behind the GrappleNames are discussed in section 2.

The *GrappleStatements* face the problem of the knowledge representation task in a distributed heterogeneous environment of independent systems where stability and formality is needed. Formality allows for sharing of information, however, the wide sharing scope increases the costs of negotiation. GrappleStatements form a compromise of uniformity and flexibility, thus several versions of GrappleStatements are defined in section 3.

The *GrappleOntologies* are an attempt to define the semantics of the concepts within GrappleStatements. Since the GRAPPLE projects tries to solve real world problems rather than develop a light weight test-tube demonstrator, we are confronted with a difficult “openness” for knowledge representation and data modelling: agents and services may enter or leave the scene, we may be confronted with conflicting statements, and we have to apply the “Open World Assumption” and not the “Closed World Assumption”. This means that missing statements are just unknown rather than false. All these issues lead to the need of reusing existing ontologies and gradually building new ontologies that are presented in section 4.

The great challenge of data modelling for sharing knowledge between LMS and ALE completely under the influence of the semantic web is new in its completeness and not straight forward at all, since the tools for the semantic web are still under development. We need to do pioneer work which however promises to come up with a better solution for integration and interoperability. The task of this work package is to carefully select and collect data models for GRAPPLE from parallel work packages with the focus on harmonization. This has to be done in several iterative rounds.

References

Berkovsky, S., Cheverst, K., Dolog, P., Heckmann, D., Kuflik, T., Picault, J., et al. (2007). Ubiquitous and decentralized user modeling. In *Workshop UBIDEUM 2007, 10th International Conference on User Modelling, .* Corfu: Springer.

Berners-Lee, T. (2005). *RFC 3986 / STD 66, the current generic URI syntax specification.* Abgerufen am 30. 10 2008 von http://tools.ietf.org/html/rfc3986

Carmagnola, F., & Dimitrova, V. (2008). An Evidence-Based Approach to Handle Semantic Heterogeneity in Interoperable Distributed User Models. *Adaptive Hypermedia, AH2008* (S. 73-82). Hannover, Germany: Springer.

Fensel, D. (2001). *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce.* Berlin: Springer.

Heckmann, D. (2003). Introducing "Situational Statements" as an integrating Data Structure for User Modeling, Context-Awareness and Resource-Adaptive Computing. *Abis'03 Workshop*, (S. 283-286). Karlsruhe, Germany.

Heckmann, D. (2006). *Ubiquitous User Modeling.* Berlin, Germany: Akademische Verlagsgesellschaft Aka GmbH.

Heckmann, D., Schwartz, T., Brandherm, B., Schmitz, M., & Wilamowitz-Moellendorff, M. (2005). Gumo - the general user model ontology. *International Conference on User Modeling (UM 2005)* (S. 428–432). Edinburgh, Scotland: Springer.

Internet2. (2008). *Shibboleth.* Retrieved 10 22, 2008, from http://shibboleth.internet2.edu/

McGuinness, D., & van Harmelen, F. (2003). *OWL web ontology language overview.* Von W3C Recommendation: http://www.w3.org/TR/owl-features/ abgerufen