Deliverable D2.1
SMARTMUSEUM Report of User Profile
Formal Representation and Metadata
Keyword Extension
Work package WP2 – Self Adaptive User Profile Management

T2.1 developing metadata keyword set extension for user profiling and data structuring
T2.2 developing resource efficient user profile representation format

Abstract

SMARTMUSEUM (Cultural Heritage Knowledge Exchange Platform) is a Research and Development project sponsored under the Europeans Commission’s 7th Framework. The overall objective of the project is to develop a platform for innovative services enhancing on-site personalized access to digital cultural heritage through adaptive and privacy preserving user profiling. Using on-site knowledge databases, global digital libraries and visitors’ experiential knowledge, the platform makes possible the creation of innovative multilingual services for increasing interaction between visitors and cultural heritage objects in a future smart museum environment, taking full benefit of digitized cultural information.

The main objective of this deliverable is to deliver formalization for user profile format as well as giving an extension of keywords used to describe the human side of access to cultural heritage.

Authors
SMARTMUSEUM Consortium (www.smartmuseum.eu)

Responsible Editors: KTH
Executive Summary

The purpose of this deliverable D2.1 is describing user profile formal representation and metadata keyword extension set. The report will contain the description of static user profile presentation and metadata keyword set to be used for user profiling.
1 Introduction .................................................................................................................................... 5
  1.1 SMARTMUSEUM Project ....................................................................................................... 5
  1.2 Deliverable purpose, scope and context .................................................................................. 5
  1.3 Audience .................................................................................................................................. 5
  1.4 Introduction to Personalization and Profiling .......................................................................... 6
    1.4.1 Adaptive, Adaptable and Personalization ........................................................................ 6
    1.4.2 User Profiling and User Modeling .................................................................................... 6
    1.4.3 Adaptive Systems ............................................................................................................ 7
  1.5 Overview of State-of-art User Profiling and Modeling Methodologies ..................................... 9
    1.5.1 The Multidimensional User Perspective .......................................................................... 9
      1.5.1.1 Context-driven user model .......................................................................................... 9
      1.5.1.2 Relationship-based models ......................................................................................... 10
      1.5.1.3 Mentalistic paradigms ................................................................................................. 10
      1.5.1.4 Task models ............................................................................................................... 10
      1.5.1.5 Environmental models ............................................................................................... 10
      1.5.1.6 Generic user models .................................................................................................. 10
      1.5.1.6.1 General User Model (and Context) Ontology ........................................................ 11
    1.5.2 Cross System Personalization ............................................................................................ 12
  1.6 Introduction to Metadata Standards and Initiatives .................................................................. 12
    1.6.1 Definition and properties of metadata ............................................................................. 12
  1.7 Overview of Metadata Standards and Initiatives ...................................................................... 12
    1.7.1 Extensible Mark-up Language (XML) .............................................................................. 12
      1.7.1.1 Dublin Core Metadata Element Set (DCMES) .......................................................... 13
      1.7.1.1.1 Semantic Web technologies ..................................................................................... 13
      1.7.2 Formal languages: RDF, DAML+OIL, OWL ................................................................. 14
      1.7.2.1 Ontologies ................................................................................................................ 14
      1.7.2.2 Topic Maps .............................................................................................................. 14
      1.7.3 Web Services ................................................................................................................. 15
      1.7.4 Metadata Harvesting - Open Archives Initiative (OAI) ................................................ 15
      1.7.5 Metadata for Personalization and Customization ............................................................. 15
  2 User Profile Formalization ............................................................................................................. 16
    2.1 User Profile Definition ........................................................................................................... 16
    2.2 User Profile Structure ........................................................................................................... 17
      2.2.1 Profile records ............................................................................................................... 17
        2.2.1.1 Context segment ....................................................................................................... 17
        2.2.1.2 Content segment .................................................................................................... 17
        2.2.1.3 Weight segment ...................................................................................................... 18
      2.2.2 Core profile format ......................................................................................................... 19
        2.2.2.1 Core format data types: ......................................................................................... 19
      2.3 User Profile Format ............................................................................................................. 20
    2.3.1 Lower level presentation formats ...................................................................................... 21
      2.3.1.1 Profile Content Mediation and Exchange .................................................................. 23
        2.3.1.1.1 Mediating RDF/XML content with SOAP messages ......................................... 23
        2.3.1.1.2 Semantic Annotations for WSDL and XML Schema (SAWSDL) ....................... 24
      2.3.2 Higher-level presentation format .................................................................................... 25
      2.3.3 Presenting context in user profiles .................................................................................. 29
      2.3.4 Mapping between Profile Levels .................................................................................... 30

Grant Agreement Number: FP7-216923
Acronym: SMARTMUSEUM
Project title: Cultural Heritage Knowledge Exchange
Platform
2.3.4.1.1 Mapping Intermediary nodes
2.3.4.1.2 Mapping Leaf nodes
2.3.4.1.3 Mapping Weight of nodes

3 User Profile Attributes
  3.1 User Perspectives and Attributes
  3.1.1 User Perspectives
  3.1.1.1 Basic User Perspective
    3.1.1.1.1 Demographics
    3.1.1.1.2 Preferences
      3.1.1.1.2.1 System preferences
      3.1.1.1.2.2 Transportation preferences
      3.1.1.1.2.3 Privacy/trust preferences
    3.1.1.1.3 Abilities
  3.1.2 Advanced User Perspective
    3.1.2.1 Relationship
      3.1.2.1.1 Role
  3.1.3 Context Perspective
    3.1.3.1 Context Dimensions
      3.1.3.1.1 Location
      3.1.3.1.2 Physical environment
      3.1.3.1.3 Visit
    3.1.3.2 Group dimensions
      3.1.3.2.1 Tours

4 User Profile Samples
  4.1 Presenting user information
    4.1.1 Presenting user’s basic information
      4.1.1.1 Age
    4.1.2 Presenting user cognitive patterns
      4.1.2.1 Interest
      4.1.2.2 Knowledge
      4.1.2.3 Abilities
    4.2 Presenting user’s history of museum experience and visits
      4.2.1 Incorporating context into visit history
  4.3 Presenting user target groups
  4.4 Presenting social relations of users

5 Bibliography

6 Appendix A: Complete List of Extended User Attributes
  6.1 Table1: Basic User Perspective
  6.2 Table2: Advanced User Perspective
  6.3 Table3: Context Perspective
  6.4 Table4: Group Perspective

7 Appendix B: Short List of Extended User Attributes
1 Introduction

The purpose of this section is to introduce the:

- SMARTMUSEUM Project
- Purpose, scope and context of this deliverable
- Intended audience for the deliverable

1.1 SMARTMUSEUM Project

SMARTMUSEUM (Cultural Heritage Knowledge Exchange Platform) is a Research and Development project sponsored under the Europeans Commission’s 7th Framework. The overall objective of the project is to develop a platform for innovative services enhancing on-site personalized access to digital cultural heritage through adaptive and privacy preserving user profiling. Using on-site knowledge databases, global digital libraries and visitors’ experiential knowledge, the platform makes possible the creation of innovative multilingual services for increasing interaction between visitors and cultural heritage objects in a future smart museum environment, taking full benefit of digitized cultural information.

The SMARTMUSEUM project supports achieving the following general goals:
- Lowering costs of on-site access to digital cultural heritage content,
- Improving structured, user behavior and preference dependent on-site access to the vast repository of cultural heritage,
- Improving the individual and shared experiences people receive from cultural and scientific resources,
- Bringing personalized cultural experience closer to non-expert communities,
- Making real reuse of personal experiences related to cultural heritage access for a variety of interest groups.

1.2 Deliverable purpose, scope and context

The purpose of this deliverable D2.1 is to develop 1) metadata keyword set extension for user profiling and data structuring and 2) Developing resource efficient user profile representation format.

1.3 Audience

The intended audience includes:

- Primarily SMARTMUSEUM Partners involved in developing the user application software for mobile device(s) including graphical user front end programming.
- Project partners involved in SMARTMUSEUM WP2
1.4 Introduction to Personalization and Profiling

1.4.1 Adaptive, Adaptable and Personalization

*Adaptive* is a property, which defines the ability to change, to suit to different conditions. In other words, something is adaptive, if it is able to change itself or something else, to fit to several circumstances, according to (Fröschl, November 2005).

According to (Oppermann, et al., 1997), there are two different types of adjustment:

First, we have *adaptable systems*, which refer to the property of changing system parameters. The user is able to change the behavior of the system. In other words, the user is able to modify the system in specified ways to fit the user needs.

Second, the term *adaptive* means the automatic tailoring of the system to the user. The needs of the user are assumed by the system itself. The user is not asked to change system parameters to its own needs, rather the necessities of the user are supposed by the system. The system changes its behavior according to these necessities. Weibelzahl (Weibelzahl, 2003) demands also another feature. He states that adaptive systems must obtain information about the user from observing the user.

*Personalization* can be provided by tailoring the content or the visualization of the system to the user’s preferences. According to (Weibelzahl, 2003), the term personalization represents the terms adaptivity and adaptability as synonyms. Thus, both types of systems, adaptive and adaptable systems, can be referred to as *personalized systems*.

According to Kim (Kim, 2002), there are at least two distinct origins for the term personalization. Firstly, dealing with the huge amount of information available today, what is referred to as information overload, it is necessary to gather and deliver only the information that is relevant to an individual or a group of individuals in the format and layout specified and in time intervals specified by the user. The second application of the term personalization originates from e-commerce field and is the concept of one-to-one marketing in which a business does marketing tailored to a group of individual users. This kind of personalization is motivated by a rise of the revenue of the business. The customer benefits through receiving useful and timely recommendations for purchasing goods or services in the most favorable terms.

1.4.2 User Profiling and User Modeling

User modeling and profiling has its roots in human computer interactions studies. Koch (Koch., May 8th, 2005) describes a user profile as a simple user model. A *user profile* is a collection of personal information. The information is stored without adding further description or interpreting this information. Koch states that user profiles represent *cognitive skills, intellectual abilities, and intentions, learning styles, preferences and interactions with the system*. These properties are stored after assigning values to them. These values may be
final or may change over time. Depending on the content and the amount of information about
the user, which is stored in the user profile, a user can be modeled. Thus, the user profile is
used to retrieve the needed information to build up a model of the user.
Koch also describes a user model as the representation of the system’s beliefs about the user.
The “real world” user is perceived by the system through the human computer interface.
According to Wahlster (Wahlster, et al., 1989), information about the user is usually collected
in a so-called user model and administrated by a user modeling system. They define (in the
context of a dialog system) the following two fundamental concepts:

A user model is a knowledge source in a system which contains explicit assumptions
on all aspects of the user that may be relevant to the behavior of the system. These
assumptions must be separable by the system from the rest of the system's knowledge.

A user modeling component is that part of a system whose function is to
incrementally construct a user model; to store, update and delete entries; to maintain
the consistency of the model; and to supply other components of the system with
assumptions about the user.

According to Middleton (MIDDLETON, et al., 2004), User profiling is either knowledge-
based or behavior-based.
Knowledge-based approaches engineer static models of users and dynamically match users to
the closest model. Questionnaires and interviews are often employed to obtain this user
knowledge. Behavior-based approaches use the user’s behavior as a model, commonly using
machine-learning techniques to discover useful patterns in the behavior. Behavioral logging is
employed to obtain the data necessary from which to extract patterns, according to Kobsa
(Kobsa, 1993).

Fröschl (Fröschl, November 2005), states that difference between user profiling and user
modeling relies in the different levels of sophistication. In general, the profiles contain “raw
material” gathered and acquired from a user while, when such data is processed it will be used
to build up a model of user, creating a sophisticated perception of user.

1.4.3 Adaptive Systems

All systems that perform an adaptation to the individual user are defined in (Jameson, 2003)
as adaptive systems.

According to Brusilovsky (Brusilovsky, et al., 2002), an adaptive system adapts itself or
another system to various circumstances. The process of adaptation is based on user’s goals
and preferences. These properties of the user are stored in a user model.
The user model is hold by the system and provides information about the user like for
example, knowledge, goals, etc. A user model gives the possibility to distinguish between
users and provides the system with the ability to tailor its reaction depending on the model of
the user.
In Figure 1, the structure of an adaptive system, according to (Brusilovsky, et al., 2002) is shown. The system intervenes at three stages during the process of adaptation. It controls the process of collecting data about the user, the process of building up the user model (user modeling) and during the adaptation process.

Figure 1: The Structure of an Adaptive System, Taken from (Brusilovsky, et al., 2002)

The behavior of an adaptive system varies according to the data from the user model and the user profile. Koch (Koch., May 8th, 2005) describes the application of user models as follows:

“Users are different: they have different background, different knowledge about a subject, different preferences, goals and interests. To individualize, personalize or customize actions a user model is needed that allows for selection of individualized responses to the user.”

Therefore, everywhere where an individualized response of the system is expected, a user model should be applied. Different types of applications can benefit from user models. Applications of user modeling are for example, search engines, recommender systems or help systems.
1.5 Overview of State-of-art User Profiling and Modeling Methodologies

Modern approaches to model-based user personalization fall into two main categories: *Multidimensional user perspectives* and *Cross context personalization*.

New approaches to model-based user personalization fall into two main categories: Multidimensional user perspectives and Cross context personalization. Multidimensional user approaches fall into following categories:

- Context-driven user representation
- Relationship-based models
- Mentalistic approaches
- Task Models
- Environment Models
- Generic approaches

Cross Context approaches fall into following categories:

- Unified User Context Model (UUCM)
- Context Passport
- W3C CC/PP

1.5.1 The Multidimensional User Perspective

Traditional approaches to personalization in digital libraries aim to narrow the gap between the vast amount of available content and individual, task-specific information needs by pre-selecting, structuring, and enriching content with metadata and further tailoring content to its users. According to Cingil (Cingil, et al., August 2000), one of the limitations of existing user-modeling techniques employed is that, they typically can manage only some limited number of dimensions (e.g., knowledge, or interest) of a user. But this limited view of user cannot answer the modern perspective to users of systems. According to Niederée (Niederée, et al., May 25, 2004), *Users are multi-dimensional, but models are not*. Users are diverse and complex: users have differing cognitive patterns, are embedded in a community, engage in multiple tasks or goals and have competing simultaneous roles that are: interactive and related to other entities in a given domain.

1.5.1.1 Context-driven user model

*Context* represents user’s intent in interaction with system, according to (Sieg, et al., 2007) Context can be thought of as the “extra”, often *implicit*, information (i.e. associations, facts, assumptions), which makes it possible to fully understand an interaction, communication or knowledge representation (Benerecetti, et al., July 2000).
An important issue of context modeling is the question on *which parameters or dimensions are to be taken into account for modeling context.*

1.5.1.2 Relationship-based models

**Relation-based models** of a user are information and community models that take into account the salient interrelationships of individuals in a cooperation or community context (McDonald, 2003). Such systems use graph-based or complex network structures to model interactions between human beings.

1.5.1.3 Mentalistic paradigms

Mentalist paradigms have emerged from machine-learning and artificial intelligence community (Pohl, 1997) (McTear, 1993).

**Mentalistic paradigms** are based on characteristics of the user which we refer to as **cognitive patterns**. These patterns represent user-specific aspects and include for example: interests, knowledge, preferences, misconceptions, or abilities (Wahlster, et al., 1989). There have been numerous systems which incorporated models of user’s cognitive patterns, (Kobsa, 2001) (Fink, 15. July 2003). Such systems have been widely used to selectively filter information on behalf of users from a large, possibly dynamic information source (Baudisch, 2001). A common example of an interest based model is a **collaborative filter** which infers a user’s interest and preferences from the ratings a user applies to an information item and from similarities between user’s interests (Konstan, et al., 1997) (Pazzani, 1999).

1.5.1.4 Task models

**Task models** of user are considered important (Kaplan, et al., 1993). Task models of the user, originated from goal-driven theories, state that goals of the users could define and influence their information needs. When these needs are known in advance, a system can better adapt to its users (Tyler, et al., 1989) (Vassileva, 14-19 August 1994).

1.5.1.5 Environmental models

**Environmental models** describe the surrounding facts or assumptions. Such additional and extra information provides a meaningful interpretation (context) to a users’ computer usage when their physical environment changes (Schmidt, et al., 1998).

1.5.1.6 Generic user models

Generic user models are, in theory, systems which have, among other aspects, two major goals (Kobsa, 2001):

1. **Generality:** which would allow a model of the user to be usable in a variety of application content domains;
2. **Expressiveness:** model is able to express a wide variety of assumptions about the user
At present, there is no unified theory which systematically integrates all dimensions. On the other hand, current systems which typically model a user along a single dimension suffer from a limited view of users and a significant amount of potentially useful information about the user may be lost; thereby demanding a need for more robust models (Sharma, 2001) (Callan, et al., June 2003).

1.5.1.6.1 General User Model (and Context) Ontology

Heckman (Heckmann, et al., 2005) introduce a user model ontology, which is generic enough to cover all possible dimensions of a ubiquitous user. This user model ontology includes the knowledge and inference mechanism for utilization of this ontology within the respective adaptive-system and context.

GUMO (general user model ontology) is a user model ontology, which models many static as well as dynamic dimensions of a user. GUMO tries to cover all classes, predicates and instances concerning the situational states and models of the users, as well as systems/devices and the environments. GUMO uses OWL for representation of user model terms and their interrelationships. Structure of GUMO, which is influenced by UserML and SitutationalStatements (Heckmann, 2003) is as following:

Subject {Auxiliary, Predicate, Range} Object

Each predicate corresponds to a user model dimension, in which each dimension expands into predicate, auxiliary and range. For instance, If we want to express knowledge about symphonies, we could divide this into the auxiliary=hasKnowledge, the predicate=symphonies and the range=poor-average-good-excellent. 1000 groups of auxiliaries, predicates and ranges have so far been identified and inserted into the ontology. Auxiliaries include (hasKnowledge, hasInterest, hasBelieve, hasPlan, hasProperty, hasGoal, hasRegularity, hasLocation), Predicates include Basic User Dimensions, Domain Dependant Dimensions include, Sensor Dimensions, Rating Dimensions include.

The following listing presents the concept Physiological-State defined as owl:Class (taken from (Heckmann, et al., 2005) ). It is defined as a subclass of BasicUserDimensions.

```xml
<owl:Class rdf:ID="PhysiologicalState.700016">
  <rdfs:label>Physiological State</rdfs:label>
  <rdfs:subClassOf rdf:resource="#BasicUserDimensions.700002"/>
  <gumo:identifier>700016</gumo:identifier>
  <gumo:lexicon>state of body or bodily functions</gumo:lexicon>
  <gumo:privacy>high.640033</gumo:privacy>
  <gumo:website rdf:resource="&GUMO;concept=700016"/>
</owl:Class>
```

The idea of having only “basic” user dimensions is allowing incorporation of other Ontologies such as SUMO (second upper merged ontology) (Pease, et al., 2002) into the respective ontology.
1.5.2 Cross System Personalization

Most of the present personalization experiences within the context of digital libraries gather information about users and their existing contexts within a profile which is restricted to a single system, although user might need sources of information which span beyond boundaries of a single system, like searching in different library collections. **Cross-systems personalization** is a way to meet the challenge of delivering personalization to users whose interaction with an information system is part of a larger task that covers several interactions with different systems, according to Stewart (Stewart, et al., 2004).

1.6 Introduction to Metadata Standards and Initiatives

1.6.1 Definition and properties of metadata

Metadata is defined as data describing data. Metadata is the value-added information which documents the administrative, descriptive, preservation, technical and usage history and characteristics associated with resources. It provides the underlying foundation upon which digital asset management systems rely to provide fast, precise access to relevant resources across networks and between organizations. (Hunter, 2003)

Sometimes a complexity level is defined for metadata. For instance, metadata required to describe the highly heterogeneous, mixed-media objects on the Internet is infinitely more complex than simple metadata for resource discovery of textual documents through a library database.

There are certain downsides and advantages to metadata, such as cost, unreliability, subjectivity, lack of authentication and lack of interoperability, with respect to syntax, semantics, vocabulary languages and underlying models.

1.7 Overview of Metadata Standards and Initiatives

In this section, an overview into existing major metadata standards and initiatives, with respect to utilization in cultural heritage, will be given.

1.7.1 Extensible Mark-up Language (XML)

XML and its associated technologies (such as XML Namespaces, XML Query languages, XML Databases) are enabling implementers to develop metadata application profiles (XML Schemas) which combine metadata terms from different namespaces to satisfy the needs of a particular community or application. (W3C, 2003)
Because XML makes it possible to exchange data in a standard format, independent of storage, it has become the de-facto standard for representing metadata descriptions of resources on the Internet

1.7.1.1 XML Schema

XML Schema Language (W3C XML Schema, 2003) provides a means for defining the structure, content and semantics of XML documents. Thus, the XML Schema language can be used to define, describe and catalogue XML vocabularies for classes of XML documents, such as metadata descriptions of web resources or digital objects. XML Schemas have been used to define metadata schemas for a number of specific domains or applications.

For instance, a particular community may want to combine elements of Dublin Core, MPEG-7 and IMS to enable the resource discovery of audiovisual learning objects.

1.7.1.1.1 Dublin Core Metadata Element Set (DCMES)

The Dublin Core metadata (DCMI, 2003) element set is a standard for cross-domain information resource description. Provides a simple and standardized set of conventions for describing online resources Dublin Core is widely used to describe Digital materials such as video, sound, image, text, and composite media like web pages. Implementations of Dublin Core typically make use of XML and RDF.

Dublin Core (DC) Metadata Element Set lists 15 standardized elements, such as dc:title, dc:creator, and dc:subject, with additional elements and element refinements.

DC is used as a basis in more detailed cultural metadata schemas, such as the Visual Resource Association’s Core Categories (VRA, 2007). VRA’s element set provides a categorical organization for Description of works of visual culture Images that document corresponding work. Most VRA elements are defined as sub-properties of corresponding DC elements.

1.7.2 Semantic Web technologies

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation." (Berners-Lee, et al., 2001)

There are two main building blocks for the semantic web: Formal Languages such as RDF, OWL and Ontologies.
1.7.2.1 Formal languages: RDF, DAML+OIL, OWL

As XML documents and schemas are ideal for defining the structural, formatting and encoding constraints for a particular domain's metadata scheme, a different type of language is required for defining meaning or semantics.

The **Resource Description Framework (RDF)** (W3C RDF, 1999) uses triples to make assertions that particular things (people, Web pages or whatever) have properties (such as "is a sister of," "is the author of") with certain values (another person, another web page).

The triples of RDF form webs of information about related things. Because RDF uses URIs to encode this information in a document, the URIs ensure that concepts are not just words in a document but are tied to a unique definition that everyone can find on the Web.

The W3C Web Ontology Working Group (WebOnt, 2003) is building upon the RDF Core work to develop a language for defining structured web based ontologies which will provide richer integration and interoperability of data among descriptive communities. This is the **Web Ontology Language (OWL)** (W3C OWL, 2003) which in turn is building upon the DAML+OIL (DAML+OIL, 2001) specification developed by DARPA.

1.7.2.2 Ontologies

Ontology consists of a set of concepts, axioms, and relationships that describes a domain of interest. An ontology is similar to a dictionary or glossary, but with greater detail and structure and expressed in a formal language (e.g., OWL) that enables computers to process its content.

Upper ontologies provide a structure and a set of general concepts upon which domain-specific ontologies (e.g. medical, financial, etc.) could be constructed. Existing research and standards groups are working on the development of common conceptual models (or upper ontologies) to facilitate interoperability between metadata vocabularies and the integration of information from different domains.

For instance, The CIDOC CRM (CIDOC-CRM, 2003) has been developed to facilitate information exchange in the cultural heritage and museum community.

1.7.2.3 Topic Maps

Topic Maps (TopicMaps, 2000) are an ISO standard for a system describing knowledge structures and associating them with information resources.

The difference between Topic Maps and RDF is that Topic Maps are centred on topics while RDF is centred on resources. RDF annotates the resources directly whilst topic maps create a "virtual map" above the resources, leaving them unchanged.
1.7.3 Web Services

web services (W3C Web Services Activity, 2003) will enable the building of software applications without having to know whom the user is, where they are, or anything else about them.

- **Web Services Description Language (WSDL)** (WSDL, 2003) which enables a common description of Web Services.
- **Universal Description, Discovery & Integration (UDDI)** (UDDI, 2003) registries which expose information about a business or other entity and its technical interfaces
- **Simple Object Access Protocol (SOAP)/XML Protocol** (SOAP, 2003) which enables structured message exchange between computers programmes

1.7.4 Metadata Harvesting - Open Archives Initiative (OAI)

Open Archives Initiative is a community that has defined an interoperability framework, the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (OAI-MHP, 2003), to facilitate the sharing of metadata. Open Archives Initiative (OAI) (OAI, 2003) provides a protocol for data providers to make their metadata and content accessible - enabling value-added search and retrieval services to be built on top of harvested metadata. To facilitate interoperability, data providers are required to supply metadata which complies to a common schema, the unqualified Dublin Core Metadata Element Set.

As an example OAI data providers now include multimedia collections such as University of Illinois historical images (UIL, 2002) and American Memory collection (Congress, 2002).

1.7.5 Metadata for Personalization and Customization

The individualization of information, based on users’ needs, abilities, prior learning, interests, context etc. is a major metadata-related research issue (Lynch, 2001). The ability to push relevant dynamically-generated information to the user based on user preferences may be implemented:

- Either by explicit user input of their preferences (obtrusive user monitoring);
- or
- learned by the system by tracking usage patterns and preferences and adapting the system and interfaces accordingly (unobtrusive user monitoring).

The idea is that the user can get what they want without having to ask. Systems which implement this type of information access are referred to as Recommender Systems. The technologies involved in recommender systems are: information filtering, collaborated filtering, user profiling, machine learning, case based retrieval, data mining, and similarity-based retrieval.
2 User Profile Formalization

2.1 User Profile Definition

As stated previously, existing approaches to modeling and profiling try to distinguish between usage model and usage profile. Within the context of our research, for the sake of uniformity, we consider the notion of model and profile of a user as equivalents and synonyms. We consider the profile to be capable of gathering as well as maintaining the usage experiences which presents and documents both knowledge and behavior perspectives of the user.

Definition – A profile is a structured collection of personal information about the user that has certain perspectives which covers different aspects of the personal attributes of the user. Profile content, documents the personal information about the user as well as history and evidence of the visiting experience of the user who is being profiled. Profile has depth (hierarchy) as well as length (flat structure), allowing us to create a level measure for details incorporated into the profile.

We consider a multidimensional view to user profile from the perspective of attribute types that we use for modeling the user. It would be important that we do not miss any kind of details that document the user’s personal data. We categorize the attributes that document the user and its behavior simply into perspectives (or dimensions) which allows us to define a hierarchy of concepts that describe the model of the user. Examples from the high-level perspectives include: interests, context and knowledge.

As it was stated previously, profile is not a repository of “raw” material as opposed to definition used within the adaptive systems context. We consider different levels for presentation of a user profile. A profile is presented in both high-level and low-level formats. From a high level perspective, profile is a hierarchy of related elements and concepts that describe user personal information such as interests. From a low level perspective we consider a profile as a record set of user information.

A profile builds up a partial interactive view of the users of the smartmuseum and allows the system to provide personalized experience for the users of the museum. The experiences of the user are observed and deduced interactively and gradually as user experiences and explores. The further the user interacts with the system the more the system gathers experiences of the user. The more system perceives the user and the better the user is modeled.
2.2 User Profile Structure

In this section, we define a structure that could be used for saving and retrieving different types of information that document both behavior and knowledge aspects of the user.

2.2.1 Profile records

A user profile contains user profiled records. In addition to user profiled information details we need a structure which could also record the history of the visit of the user as well as ranking weight of the information described. We would like to also incorporate security information describing the privacy of the profiled information as well as trusted arguments pertaining to profiled information. We divide the structure into different sections. We call each section of the profile records a segment. Each record is made up of three segments: Context, Content and Weight.

2.2.1.1 Context segment

Context is defined as extra information. We can consider all surrounding information as contextual information. We emphasize this perspective in the museum experience by dedicating the first segment of profiled materials to contextual information. Here, we assume that all surrounding facts could be considered context and information contained within them could be considered contextual information. In order to make such format more generic, we avoid inserting values directly to context segment and instead we can give a reference (in the form of URI) to existing contextual information. For instance we can define a context ontology that documents contextual concepts and it becomes instantiated when contextual information is available in the context database.

2.2.1.2 Content segment

This segment will contain the actual material which is being profiled. Most of existing approaches to user profiling, represent user profiles as record (vector) of attribute and value pairs. One of the most popular formats is vector space model (Cetintemel, et al., 2000) (Salton, 1989). In VSM, text-based documents are represented as vectors in a high-dimensional vector space where the values of dimensions are based on the words occurring in the documents. Documents describing similar topics are likely to be close to each other, as they possibly include common words. A profile (content) can also be represented as a single vector (record), or a collection of vectors (records), which can be derived from the previously judged document vectors. In general, a profile vector should have a position close to those of relevant document vectors (in the vector space).

If there are n distinct terms in a document d, then d will be represented by a vector

\[ V = ((t1; w1); (t2; w2); \ldots; (tn; wn)) \]

---

1 Uniform resource identifier
Ontologies, at the core of semantic web technologies, are playing a crucial role in profiling and modeling of usage-driven personalized software systems. Ontologies are defined as specification of concepts and relationships between the concepts within a certain domain. Ontologies have been extensively used for enriching user profiles with semantics of the concepts of the domain. Ontologies have been majorly used to deal with the so called “cold-start problem” (MIDDLETON, et al., 2004). Cold-start problem is the problem of lack of enough knowledge about user preferences to start provision of personalized information retrieval. Ontologies have been used extensively in personalizing the user experience on the web. (Sieg, et al., 2007)

By utilizing semantic web technologies, we use RDF described triplets of predicate, subject and objects to describe information contained in the profile.

For instance, “interest of user in science of Evolution” can be described as (hasInterest, user, “Evolution”) tuple. Here subject is used to describe the attribute and object is used to describe the value and the predicate describes the semantics of the relations of concepts. Predicate allows us to give further meaning to information which is being profiled. Using RDF-triplets cases later extraction and mapping of RDF data to lower level metadata formats such as XML or higher level formal presentational formats such as OWL.

2.2.1.3 Weight segment

In the VSM (Cetintemel, et al., 2000), each document is represented as a vector of term and weight pairs. So each single term representing the content of the user profile carries a weight. We have also considered assigning weights onto each piece of profiled material as well. VSM uses a one-dimensional weight for changing the interpretation of profiled material while we use a multidimensional and multimodal approach to assigning weight to profiled material. Each profile record contains a weight tuple which can specify the weight of information being profiled. Semantics of the weight segment could be different depending on the profiled data.

For instance if the profiled information is about user’s cognitive patterns such as interests, then the weight would describe the intensity of interest of user in the object atom specified in RDF triplets, meaning for instance, how much user is interested in a certain artifact.

In addition to the profiled information and its designated weight we incorporate a segment for security credentials. In general, trust describes the trust, belief and confidence of the user in the piece of information profiled while privacy describes the privacy of the piece of information recorded.

Privacy could have values between range [-1, 1] where positive values could describe the positive consent of the user towards sharing and disclose the information disclosed in profiled record, while negative value could describe the negativity of the user towards disclosure of the content recorded to outside world. Using such weighted information user can specify which atomic piece of information he/she would like to disclose or not to disclose to outside world.

Trust can be interpreted differently depending on the case scenario. For instance, trust in an artwork could document the originality of work being experienced by the user, while trust in profiled information documents the trust of user in the information documented and profiled or just trust of the user in system. Such atomic representation of trust could also be used for describing the trust in individuals while describing relationships between users.
2.2.2 Core profile format

By combining the segments of the information described, the generic structure will be constructed. Resulting generic structure is formalized as below:

A profile is a collection of records:

\[
\text{Profile: } = \langle \text{Record 1, Record 2, Record 3 ... } \rangle
\]

Each record is made up of three segments of context, content and weight:

\[
\text{Record: } = \langle \text{Context, Content, Weight } \rangle
\]

\[
\text{Context: } = \text{Contextual reference,}
\]

\[
\text{Content: } = \text{Predicate, Subject, Object,}
\]

\[
\text{Weight: } = \text{Rank, Trust, Privacy}
\]

As a matter of fact, resulting record structure will be a septuplets group as following:

\[
\text{Profile Record: } = \langle \text{Contextual reference, Predicate, Subject, Object, Rank, Trust, Privacy } \rangle
\]

2.2.2.1 Core format data types:

Now that the profile structure is defined we can define preferred datatypes for pieces of information which are being profiled.

<table>
<thead>
<tr>
<th>Value</th>
<th>Preferred Data Type</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>String</td>
<td>Reference, for instance</td>
</tr>
<tr>
<td>Reference</td>
<td>String</td>
<td>Uniform Resource Identifier (URI)</td>
</tr>
<tr>
<td>Predicate</td>
<td>String</td>
<td>Basic RDF element</td>
</tr>
<tr>
<td>Subject</td>
<td>String</td>
<td>Basic RDF element</td>
</tr>
<tr>
<td>Object</td>
<td>String</td>
<td>Basic RDF element</td>
</tr>
<tr>
<td>Rank</td>
<td>Integer/Float</td>
<td>Integer [0,1] or Floating range [-1,1]</td>
</tr>
<tr>
<td>Privacy</td>
<td>Float</td>
<td>Floating range [-1,1]: Negative values are important</td>
</tr>
<tr>
<td>Trust</td>
<td>Float</td>
<td>Floating range [-1,1]: Negative values are important</td>
</tr>
</tbody>
</table>
### 2.3 User Profile Format

In order to make the profiled user data interoperable in between systems and devices used in the architecture, we need to create different presentation formats and map profiled information between them. This also facilitates implementation and management of content profiling operations and services.

We present the profile structure in three levels: Located at the top, we have the high-level presentation in a hierarchical formation. We can extract the semantic profile content and map it to higher level semantics and present the knowledge documented the profiles in high-level presentation. At the middle-level, the core structure is depicted. At the bottom, we have low-level presentation, which allows us to create a flat structure that can be described using metadata schemas for better interoperability in between systems.

This abstraction is mainly based on the semantics of the content presentation. As depicted information in core profile structure is described in RDF, with basic semantics, while on higher level the knowledge can be further extended, using OWL, for instance. At the same time by flattening the profiled records, we can represent them in XML format in lower level.

Figure below depicts the three aforementioned levels and their corresponding profile languages.

![User profile presentation in different levels](image)

**Figure 2**- User profile presentation in different levels
2.3.1 Lower level presentation formats

We can consider using lower level semantics for profile presentation format. For instance, if we consider solely RDF presentation for the content of the profile, then we can create an RDF schema that describes the semantics of the resources (profiled materials) being documented (in the form of subjects and objects) as well as using predicates that document the relation in between resources.

In order to better understand how RDF/XML syntax works, and RDF content get stored in a file let us take a look at an example:

```
<http://www.smartmuseum.eu/MusuemOnt/rdf/people/Sandro> -> name -> Sandro
```

There are three pieces of information that will need to be stored, the resource subject, the predicate, and the target(object). We didn't put the 'name' predicate in a namespace for simplicity. However, we will need to add a namespace, and for the purposes of this example, we will just define one. Here is the information that will need to be stored:

| Subject: | <http://www.smartmuseum.eu/MusuemOnt/rdf/people/Sandro> |
| Predicate: | <http://www.smartmuseum.eu/MusuemOnt/rdf/people/name> |
| Object: | Sandro |

This is a triplet describing a resource to literal relationship. This type of relationship can be described using a single tag with two attributes:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
xmlns:people="http://www.smartmuseum.eu/MusuemOnt/rdf/people/">
  <rdf:Description rdf:about="http://www.smartmuseum.eu/MusuemOnt/rdf/people/Sandro"
    people:name="Sandro"/>
</rdf:RDF>
```

In order to better understand the usage of RDF schema, let us have an example. An RDF Schema for the classification of the resources presented in our sample, which is presented later on in Figure 6, with respect to their hierarchy and relationship can be described below.

We have defined a target namespace, namely SmartMuse. While for presenting the basic RDF elements (subject, object and predicate) as well as extended RDF material, we have used rdf and rdfs namespaces. The schema is presented in XML format. For defining the concepts in the schema, we have used rdfs:Class, while sub concepts are defined using rdfs:subClassOf. The relationships between concepts(edges on the networks) are defined using rdf:Property. For instance, hasPainted is defined as an rdf:Property on the domain of SmartMuse:Artist (source node) and on the range of any rdf:resource (as target node).

```
<?xml version='1.0' encoding='UTF-8'?>
--
<rdf:RDF xmlns:SmartMuse="&SmartMuse;" xmlns:rdf="&rdf;" xmlns:rdfs="&rdfs;">
```
Grant Agreement Number: FP7-216923
Acronym: SMARTMUSEUM
Project title: Cultural Heritage Knowledge Exchange Platform

Figure 3 – Sample RDF schema for user profile

If we populate and fill the schema with values, the network of the sample profile can be generated as following. This instance set of RDF triplets represents the nodes and their relationships. For instance, Sandro is instance of Artist resource who has painted Venere.
2.3.1.1 Profile Content Mediation and Exchange

As stated previously, one of the main motivations for having information available in different formats is maximizing interoperability in between systems involved in SMARTMUSEUM. XML is recognized to be one of the main standards for interoperable data exchange on heterogeneous environments such as internet. As a matter of fact, we can consider using XML as an enabler for inter-system data exchange in SMARTMUSEUM intranet or for online web services information exchange with other museums or perhaps other related online service providers.

In the context of SMARTMUSEUM, profiling content services can interact with other internal services or with external services offered and exposed by external systems. Based on web service enabled interfaces, we can describe the existing services and their interfaces using WSDL and services can exchange XML encoded contents in the payloads of SOAP messages.

In order to mediate the RDF profile content between systems and services, we can either encode the RDF/XML content into SOAP messages or we can use SAWSDL approach to mediating and exchanging RDF data.

2.3.1.1.1 Mediating RDF/XML content with SOAP messages

There are generally two options for encoding RDF/XML documents into SOAP messages:

1. Sending raw RDF
2. Sending RDF/XML data using SOAP encoding

There seems to be no official RDF encoding for SOAP messages at the moment. SOAP encoding of data in RDF/XML format is described at:

A set of of large examples can be found at:
2.3.1.1.2 Semantic Annotations for WSDL and XML Schema (SAWSDL)

SAWSDL (Kopecký, et al., 2007) is a set of extensions for WSDL, which provides a standard description format for Web services. WSDL uses XML as a common flexible data-exchange format and applies XML Schema for data typing.

The SAWSDL extensions take two forms: *model references* that point to semantic concepts and *schema mappings* that specify data transformations between messages’ XML data structure and the associated semantic model.

A *model reference* is an extension attribute, sawsd1:modelReference, that we can apply to any WSDL or XML Schema element in order to point to one or more semantic concepts. The value is a set of URIs, each one identifying some piece of semantics. Model references generically refer to semantic concepts, thus serve as placeholders for adding semantics. Model references can be used to describe the meaning of data or to specify the function of a Web service operation.


Lifting mappings transform XML data from a Web service message into a semantic model (for instance, into RDF data that follows some specific ontology), whereas lowering mappings transform data from a semantic model into an XML message.

Lifting and lowering transformations are useful for communicating with a Web service from a semantic client — for example, the client software will lower some of its semantic data into a request message and send it to the Web service; when the client software receives the response message, it can lift the data contained in the message for semantic processing. We can also use lifting and lowering annotations for XML data mediation through a shared ontology.

![Figure 5 - Lifting and lowering semantic data](image_url)

We can use lifting and lowering transformations for (a) Web service communication and (b) XML data mediation. Taken from (Kopecký, et al., 2007).
An automated mediator can lift the data in one XML format to data in the shared ontology and then lower it to another XML format using the lifting annotation from the first format’s schema and the lowering one from the second schema.

### 2.3.2 Higher-level presentation format

As stated previously, we assume a high level presentation format for user profiles. By making such assumption, we can create the possibility of adding further semantics into existing profiled material and we can describe and present the resulting knowledge in a higher level formal semantics language such as Web Ontology Language (OWL).

Similar as previous case, gathered usage profiles create a hierarchical structure for high-level profile representation where concepts documenting attributes of the users are presented as nodes in a hierarchical structure in the form of a graph and semantics of their relationships are documented and represented, as well.

In order to understand the structure and language of the profiles in each level, let’s take a look at sample partial profile of a user. Figure 6 depicts a sample high-level profile.

![Figure 6 - A high-level user profile depiction](image)

As depicted each concept forms a node in the hierarchy and carries a weight associated to it. Using RDF atoms we can describe the concept hierarchy between each parent concept.
(Subject) and child concept (Object) and using inter concept relations (predicate) we can describe the semantics of their interrelationships.

Let us consider an interpretation of the semantics of this profile: in this example, the history of the user experience corresponding to the artwork of renaissance have been recorded and presented. Visited artworks form instances and are located in the leaves of the graph. As depicted in hierarchy, the user has visited works corresponding to high and low renaissance of Italy, and from both artists two artworks have been visited and ranked respectively. As it is shown User has been fairly interested in works of both artists, which is described by the value recorded by rank tuple. It is well-described that the user is more interested in a certain artist’s personality while his artwork is not of high interest of the user (i.e. Michelangelo), while it is the opposite case with other artist (i.e. Botticelli) as the user is interested more in artists’ artworks rather than his personality.

In addition to interest and ranking weight, security credential has allowed the user to express the trust and privacy in recorded information as well. In the case of privacy, user has shown that it is not interested at all that his profiled visit be disclosed to public (negative values for privacy) while user is positive with sharing his/her high-level preferences with outside world (Italian renaissance has positive value for privacy). Trust in this context describes the user’s belief in originality of the art works visited, in the case of leaf nodes and in the case of art workers and artists, it could describe the trust in the factual information described and presented regarding the artists. For instance, the user has no trust in the integrity of the information presented by the system on Italian renaissance.

Now that the profile content and respective meanings of the semantic user profile have been presented, let us see how such information can be described and presented using the corresponding presentation format.

By defining an ontological classification of concepts of usage domain, we can extend the semantics of core format and add further semantics to it. As an example, a partial schematic classification for the concepts of our sample domain can be defined. Figure 7 shows a sample schema of the concepts corresponding to domain, with respect to their hierarchical ordering.

![Figure 7 - Partial schema for sample profile](image-url)
We can present the corresponding ontological classification using a formal language. In this instance using OWL, the classification can be presented as following. Concepts such as Artist has been defined using owl:Class. For defining parent and child relationship, we have used rdfs:subClassOf on the rdf:resource, acting as parent. Relations in between the concepts are defined using owl:ObjectProperty. For instance, hasPainted is defined using owl:ObjectProperty on the domain of Artist using rdfs:domain, which creates an outgoing edge from Artist (source node) to destination node with label of hasPainted.

```xml
<rdf:RDF>
  ...
  <owl:Class rdf:ID="Artist">
    <rdfs:subClassOf rdf:resource="#Person"/>
  </owl:Class>
  <owl:Class rdf:ID="ArtworkTimeLine">
    <rdfs:subClassOf rdf:resource="#Art"/>
  </owl:Class>
  <owl:Class rdf:ID="ArtworkType">
    <rdfs:subClassOf rdf:resource="#Art"/>
  </owl:Class>
  <owl:Class rdf:ID="Era">
    <rdfs:subClassOf rdf:resource="#ArtworkTimeLine"/>
  </owl:Class>
  <owl:ObjectProperty rdf:ID="hasArtist">
    <rdfs:domain rdf:resource="#Renaissance"/>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="hasPainted">
    <rdfs:domain rdf:resource="#Artist"/>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="hasSculpted">
    <rdfs:domain rdf:resource="#Artist"/>
  </owl:ObjectProperty>
  <owl:Class rdf:ID="ItalianRenaissance">
    <rdfs:subClassOf rdf:resource="#Renaissance"/>
  </owl:Class>
  <owl:Class rdf:ID="Person"/>
  <owl:Class rdf:ID="Renaissance">
    <rdfs:subClassOf rdf:resource="#Era"/>
  </owl:Class>
  <owl:Class rdf:ID="Sculpture">
    <rdfs:subClassOf rdf:resource="#ArtworkType"/>
  </owl:Class>
  ...
</rdf:RDF>
```

**Figure 8** – A sample OWL/RDF presentation of profile schema

When described ontology becomes instantiated, resulting instances are in RDF format. For instance figure below is a sample RDF instance set, generated as a result of instantiation of previously mentioned OWL schema.

In order to fully map profile records onto the high-level presentation format, we can consider modeling equivalents for Weights of the profiled records. We can define Trust, Privacy, Rank and Context as rdf:Datatype. We can also consider just mapping RDF data from profile records onto high-level presentation format.
Figure 9 – RDF/XML presentation of user profile. Values are realistically presented to reflect the precision of the application and sensory hardware.

Having the relationship between the concepts of the domain of usage, the instantiated RDF content of user profiles creates a network. This instance set of RDF triplets represents the nodes and their respective weights. Each node is represented as a tag with the name of the concept and id of its instance. For instance, Venere, as an instance of Painting, is presented as with a tag, while rdf:ID represents the name label of this instance.

The RDF network of the sample visit is visualized in Figure 10. In this figure the purple triangles represent the instances of the domain ontology for user profiles, while directed edges in between them represent the properties (or predicates). The green boxes represent the weights that each term carries and outgoing edges to them represent the type of the edges.

Figure 10 – Networked visualization of a user profile: Purple triangles represent the instances, green labels represent the weights (trust, privacy and rank) assigned to items. (Visualized using RDF Gravity²)

Since the whole visit is being documented in the same context, same literal value for all of the profiled information is recorded, and all the edges with label context are pointing to the same literal value, at the center of the figure.

### 2.3.3 Presenting context in user profiles

As described the role of the context in the museum experience could give an exceptional opportunity to create much more personalized experience for visitors and users. By considering a hierarchy for context we can incorporate more than one single dimension for context. As stated context could have multiple dimensions such as visit, physical environment and location.

In order to point out the effect of having context in the presentation format, we can consider a sub graph of our sample profile. This structure describes the profile of a user in different contexts. The first one describes the profile in the context of *visit with family* (with role father or parent perhaps) while the second one is the same profile, but in the context of visit as a *scholar*. These presentations are depicted in Figure 11.

As context changes the semantics of the visit, we can model the effect of such change by mathematically alternating the weight of profiled materials. For instance, in visit with family, perhaps user does not want to pay attention to details of the artworks, while as scholar the user has paid more attention onto details of the artworks and has eventually become more interested in the artist as well. As a matter of fact, since user didn’t pay much attention to artistic details, the trust levels were low, while under the context of visit as a scholar trust values seems quite higher. In order to present this fact in core format, we need to create a record set per each visit of the user. In order to further emphasize this difference, we present the effect of context in high-level profile format. As the user might visit the same artwork more than once then, we can create multiple instances of the same concepts of knowledge, with different weight, each representing visits of the user under different contexts. So if user visits painting of primavera for instance, we have two instances for “primavera”, like primavera and primavera2.

---

**Grant Agreement Number:** FP7-216923  
**Acronym:** SMARTMUSEUM  
**Project title:** Cultural Heritage Knowledge Exchange Platform
Figure below depicts networks of the semantic profiles, in different contexts. As mentioned previously, for each concept that is revisited, a new instance has been created and new weights are generated and assigned. As it is depicted, the user shows more confidence and interest in revisited artwork in the context of visit as a professional rather than a group member.

Figure 12 – Visualization of profiles under different contexts: Left) profile with context “visit with family”, right) profile with context “visit as scholar”. Weight segments take different values depending on the context.

As an effect of having different contexts for profiled information, as opposed to having a single context, RDF networks generated are separated and independent.

2.3.4 Mapping between Profile Levels

On a theoretic level we can describe the mapping between segments of profiled records in different levels of presentation.

Figure below depicts a sub graph of high-level presentation of profiled data and it’s mapping to a core profile record. In order to point out the relationship between a profiled record and its high-level presentation, we have to make certain assumptions. As described by considering an ontological classification of a domain, a profile can be seen as an instance of this ontological classification with literals describing weights and references to contexts. In order to better understand the correlations, we scope the correlation between the levels onto clusters of instances which form a profile record.

2.3.4.1.1 Mapping Intermediary nodes

We map each profile record (a single profile unit) to a record cluster on the network. A record cluster always contains a set of nodes that form subject and object and an edge between them.
that takes the *predicate* label defining their relationship. So each two nodes with an edge in between can form a cluster which is documented in a single profile record.

![Figure 13 - Mapping Intermediary nodes](image)

In this example, we have a cluster which contains two nodes where the one with outgoing edge forms the subject and the one with incoming edge forms the object, while the edge between them has the label which forms the predicate. The correlation between high-level profile data and low-level profile data is depicted in the figure. This record documents information about an artifact, a museum item, which belongs to the artist described in related record. The content of this record describes the artist has painted this painting along with weighted segment values, describing interest, trust and privacy.

Using our core format the record is presented as below:

```
<Reference to context, has Painted, Michelangelo, Holy Family, 0.9, 0.8,-1>
```

We can present the profile record using RDF/XML metadata notation as following:

```xml
<Artist rdf:ID="Michelangelo">
  <hasPainted rdf:resource="#Holy_Family"/>
</Artist>
<Painting rdf:ID="Holy_Family"/>
```
2.3.4.1.2 Mapping Leaf nodes

Leaf nodes are nodes which has no outgoing edges on our network. This means that we have no value to document as their predicate. This also means we have no value to document as object. For instance in last figure, the Holy family has this situation.

We deal with this problem by creating an inverse edge; from the leaf node to one of the consecutive parent nodes we create an edge to allow us to document the leaf nodes properties. For instance in this case we can define BelongsTo predicate which allows us to create the following record:

\[
\text{< Reference to context, BelongsTo, Holy Family, Michelangelo, 0.67, 0.4,-1>}
\]

Or:

\[
\text{< Reference to context, hasPainter, Holy Family, Michelangelo, 0.67, 0.4,-1>}
\]

\[
\text{< Painting rdf:ID="Holy_Family">}
\text{<hasPainter rdf:resource="# Michelangelo "/>}
\text{</Painting>}
\]

2.3.4.1.3 Mapping Weight of nodes

So far we have made an assumption that profiled material is in RDF format of triplets of subject, object and predicate. In higher level presentation formats such as RDF or OWL we can also consider mapping the rest of the recorded material; context reference and weight values of rank, trust and privacy to corresponding and saving them in rdf/xml format.

For instance in that case of example in Figure 13, the weight documented in the record is the weight of the subject node. For instance, in this example the weight in the record corresponds to weight of node with instance of artist as Michelangelo.

For instance this weight information can be presented in rdf/xml as below. All the values are defined using rdf:datatype of floating point.

In the case of Context, since contextual reference is presented using URI, then we can also map and save the context reference using rdf:datatype of string.

\[
\text{<Privacy_value rdf:datatype="&xsd;float">-1.0</Privacy_value>}
\text{<Rank rdf:datatype="&xsd;float">0.9</Rank>}
\text{<Trust_value rdf:datatype="&xsd;float">0.8</Trust_value>}
\]
3 User Profile Attributes

In order to be able to extend existing keywords used for describing cultural heritage to include usage and user-driven attributes, first we need to understand the attributes of usage.

In Figure 14, attributes of usage are depicted as the extension of existing metadata descriptions of museum content. Usage attributes pertain to attributes describing user as an individual or a group of users. These attribute sets are further categorized into perspectives and depending on the importance they can be even further categorized until they reach atomic attributes. Along with attributes of usage, context is also considered as an extended keyword set, which has its own perspectives and categorization in return.

In order to create the connection between the user model and extended keyword set, we have used references.

---

**Figure 14 - Extending Metadata Descriptions to include Usage Attributes**
3.1 User Perspectives and Attributes

We have defined an ontological structure for usage modeling in SMARTMUSEUM.

The high level user model presentation is depicted in Figure 15.

Beginning with user model at the root, on the second level we have the perspectives (dimensions) of usage. By defining perspectives of the user we try to document the personal attributes that the user will have and these attributes could be used to describe the user and different aspects of it such as personal attributes, cognitive patterns (interests), intellectual abilities and possibly disabilities, preferences as well as interactions with the system.

We have considered 4 main attribute categories for modeling the usage:

*User (basic and advanced) attributes, context attributes, and user group attributes.*

![Diagram showing user perspectives and attributes](image)

**Figure 15** – Main perspectives of extension set with their respective attributes.
On the third level we have *attribute bags (or composite attributes)*, where each of them divides into different attributes respectively. We have only shown high-level organization of usage attributes. Depending on the importance and usage of respective attributes, we might have deeper levels until we reach an atomic attribute slot (where a value will be assigned).

For creating this structure, we have partially used ontological structure and organization of GUMO (general user model ontology), (Heckmann, et al., 2005) and UserML (Heckmann, 2003) to describe the attributes and perspectives which pertains to users. As GUMO uses an ontological metadata hierarchy, we have followed the same organization in order to categorize and represent selected user attributes.

### 3.1.1 User Perspectives

This is the fundamental perspective of the model. This perspective allows us to document attributes pertaining to users. This perspective is divided into two sub-perspectives: *Basic* and *Advanced*. Basic perspective allows documentation of basic attributes of users, such as name and address, while advanced attributes allows documentation of advanced attributes of the user such as interest and relationship.

#### 3.1.1.1 Basic User Perspective

Basic perspective of the user documents attributes that shape a basic profile of the user. For instance, User intellectual abilities or disabilities could be used as user basic attributes, such as typing abilities and reading abilities.

Demographics of the user could describe demographical information of the user such as culture or mother tongue.

As basic user perspective is the fundamental section of usage model, it is important that certain attributes in this category need to be specified as obligatory to make sure system has enough credential about user. Perhaps a minimal set can be specified.

Figure 16 depicts the basic user perspective hierarchy.
Figure 16 – Basic user perspective with subcategories

3.1.1.1.1 Demographics

Demographics of the user could describe demographical information of the user such as mother tongue and other languages user might speak, gender, education level.

Among the attributes in this category perhaps preferred language could be set as mandatory.

A sample RDF presentation of this attribute set can be presented as following:

```
<Demographics rdf:ID="Demographics_21">
  <HighestEducationLevel rdf:datatype="&xsd;string">High School</HighestEducationLevel>
  <Age rdf:datatype="&xsd;string">25</Age>
  <Education_Level rdf:datatype="&xsd;string">High School</Education_Level>
  <Preferred_Language rdf:datatype="&xsd;string">Spanish</Preferred_Language>
  <Gender rdf:datatype="&xsd;string">Male</Gender>
</Demographics>
```
3.1.1.1.2 Preferences

This category allows us to document preferences of the user. Attributes under this category allows further personalization of the experience of the users both online and onsite by documenting how user customizes the experience, depending on the choices he or she prefers.

Attributes in this category fall in three main subcategories: system preferences, personal preferences and privacy/trust preferences. These categories and their successive attributes are presented in figure below.

![User preferences diagram](image)

**Figure 17** – User preferences

3.1.1.1.2.1 System preferences

This category of attributes allows documentation of user preferences with regards to system. For instance, user can specify if he/she wants the commercials to be shown on the device while visiting the museum or not.

Friend Finding capability specifies if the user wants to be browsed by others or wants to find others in the museum perhaps with same interests, using Bluetooth connectivity or wireless access.

Recommendation capability allows user to state if he/she prefers receiving recommendations or suggestions from the system.

3.1.1.1.2.2 Transportation preferences

This category specifies certain attributes pertaining to personal attributes of the user.
For instance, using transportation type user can specify which type of transport he/she likes to have when leaving museum. Using this information system can announce the availability of a specific type of transportation to user at the end of visit.

3.1.1.2.3 Privacy/trust preferences

This category could be used to allow users to manually specify the trust and privacy values. Two scales are specified to allow users to show much they want their profile to be private and how much they trust the system.

A sample RDF presentation of this attribute set can be presented as following:

```
<Preferences rdf:ID="Preferences_29">
  <Commercials_Availability rdf:datatype="&xsd;boolean">true</Commercials_Availability>
  <Privacy_Scale rdf:datatype="&xsd;int">0</Privacy_Scale>
  <Trust_Scale rdf:datatype="&xsd;int">5</Trust_Scale>
  <Friend_Finding_Capability rdf:datatype="&xsd;int">5</Friend_Finding_Capability>
  <Recommendation_Capability rdf:datatype="&xsd;boolean">true</Recommendation_Capability>
</Preferences>
```

3.1.1.3 Abilities

This category allows documentation of user intellectual capabilities. Abilities in general fall under two main categories: Cognitive (mental) capabilities as well as physical capabilities. We have considered mostly physical abilities here.

A sample RDF presentation of this attribute set can be presented as following:

```
<Abilities rdf:ID="Abilities_27">
  <Ability_to_UseStairs rdf:datatype="&xsd;boolean">true</Ability_to_UseStairs>
  <Ability_to_Walk rdf:datatype="&xsd;boolean">true</Ability_to_Walk>
  <Reading_Skills rdf:datatype="&xsd;string">Advanced</Reading_Skills>
  <Ability_to_Talk rdf:datatype="&xsd;boolean">true</Ability_to_Talk>
  <Typing_Skills rdf:datatype="&xsd;string">Novice</Typing_Skills>
  <Ability_to_Hear rdf:datatype="&xsd;boolean">true</Ability_to_Hear>
  <Ability_to_See rdf:datatype="&xsd;boolean">true</Ability_to_See>
  <Ability_to_Touch rdf:datatype="&xsd;boolean">true</Ability_to_Touch>
</Abilities>
```

3.1.2 Advanced User Perspective

As stated advanced perspective of the user could document advanced attributes of the user. Attributes in this perspective fall into two major categories: relationship and Interest.
3.1.2.1 Relationship

Relationship perspective allows us to model connectivity of individuals with other users. Relationship can have different sub-dimensions. We have considered role as the most important one.

3.1.2.1.1 Role

One of the most important subdimensions of relationship is the role of the user. Role of the user in experiencing the museum could have different interpretations. For instance, if the user is visiting as a businessman perhaps he won’t pay much attention to artistic details of an art gallery due to time restriction, but might want to read all details of craftworks if the user is experiencing in the role of tourist.

We have considered the following roles as most important ones:

Inhabitant, teacher, tourist and parent.

A sample RDF presentation of this attribute set can be presented as following:

```xml
<Role rdf:ID="Role_1">
  <Role_Type rdf:datatype="&xsd;string">Parent</Role_Type>
  <Role_Type rdf:datatype="&xsd;string">Tourist</Role_Type>
</Role>
```

3.1.3 Context Perspective

As described previously, context means extra information. Contextual information is mostly implicit and they are processed along with existing information to change the meaning of interpretation, when they are available. Context can be seen as an important perspective in modeling the human side as it could improve the personalization of the museum experience.

As different sources of information can be considered as context, then we assume different dimensions for context.

Figure below presents the context along with dimensions we have considered for it:
3.1.3.1 Context Dimensions

3.1.3.1.1 Location

Location could be seen as the most important attributeset for the observing user behavior. Location sensors provide contextual information regarding the physical location of the user during the onsite experience. The attributes defined for location context, depend heavily on sensors at hand and their precision.

Location dimension is divided into following subcategories:

- Coordinates
- Physical Location
- Virtual Location
- Spatial Location
3.1.3.1.2 Physical environment

Physical environment could be seen as other important aspect of context. For instance, temperature (for indoor experiences) or weather (for outdoor experiences) could be seen as the other important aspects of physical surrounding facts.

Physical Environment is divided into following subcategories:

- Humidity
- Level of wind
- Light Level
- Noise Level
- Opening hours
- Temperature
- Weather

3.1.3.1.3 Visit

As museum visit can be seen as a proper context dimension. Visit can have different subcategories, namely visit companion such as with family or alone, visit motivation such as art and culture or wellness and relaxing, visit duration such as 1 hour, 2 hour or longer, travel season and etc. It could be of use if we can understand this user is a first time visitor. We have also taken into account goals and plans visitor might have for visiting the exhibit. From a more generic view, we can have season or year considered as time context sub-dimension. We can also state what the attributes for typology of the visitors (Sparacino, 2002) are.

Figure 19 – Visit as Context
3.1.3.2 Group dimensions

We can group a set of users with a shared set of attributes as target groups (or stereotypes).

Grouping of users can be done on many bases. This grouping can be done based on age, culture, and knowledge and group size. One of the most useful contexts to study the user groups is museum tour.

![Figure 20 – Group Dimensions](image)

3.1.3.2.1 Tours

In order to better understand the tours and their corresponding attributes, consider the following tour description:

**Art and Junk**

*This workshop offers children, teenagers and families to conduct an investigation with historical sources in form of writing, pictures and objects.*

*A tour around the permanent exhibition’s upper floor will enable the participants to recognize the different forms and characteristics of historical sources on the basis of original examples. The aim of the following teamwork period inside the exhibition is the critical reflection, analysis and interpretation of various historical sources.*

<table>
<thead>
<tr>
<th>Target Group:</th>
<th>Children from 14 to 18 years and families</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration:</td>
<td>120 minutes</td>
</tr>
<tr>
<td>Fee:</td>
<td>€25 per group plus €5 admission; for children and teenagers under the age of 18 admission free</td>
</tr>
<tr>
<td>Group Size:</td>
<td>Up to 10 individuals</td>
</tr>
</tbody>
</table>

3 Taken from [http://www.dhm.de/ausstellungen/museumspaedagogik/staendige-ausstellung/english/kinder.html](http://www.dhm.de/ausstellungen/museumspaedagogik/staendige-ausstellung/english/kinder.html)
Based on given example, we can consider following attributes for tours:

- Tour Types
- Tour Availability
- Tour Times
- Tour Age Group
- Tour Size
- Tour Fee
- Tour Duration

**Figure 21 - Tour Attributes**

```xml
<Tour rdf:ID="Tour_9">
  <Tour_Name xml:lang="en">An insight to Galileo’s Telescopes</Tour_Name>
  <TourAgeGroup rdf:datatype="&xsd;int">4</TourAgeGroup>
  <TourSize rdf:datatype="&xsd;string">Less Than 4</TourSize>
  <TourFee rdf:datatype="&xsd;int">10</TourFee>
  <TourType rdf:datatype="&xsd;string">GuidedTour</TourType>
</Tour>
```
4 User Profile Samples

In this section examples of profiled information will be presented as a guide to the structure of the information in profile format proposed.

4.1 Presenting user information

In this section we describe profile records defined for recording and presenting user attributes.

4.1.1 Presenting user’s basic information

4.1.1.1 Age

In the case of contact details we can record and present the information using our template in the following manner with following meanings:

“User’s age is 17 and this information might not be true (showing low trust level, since user might have entered wrong information) and information is somehow private”.

We can describe it using:

   < Reference to context, is, Age, ”18”, 0.3, 1,-0.3>

A sample RDF syntax for this slice of profile can be presented as following:

<SmartMuse:Age rdf:about=”&SmartMuse;User8012314699”
   SmartMuse:is= ”18”/>

4.1.2 Presenting user cognitive patterns

4.1.2.1 Interest

for instance, if we want to state that;

   “User’s interest in art is high and information is trusted and is totally private”.

We can describe it using:

   < Reference to context, hasInterest, user, ”art”, 0.8, 1,-1>

A sample RDF syntax for this profile slice can be presented as following:

<SmartMuse:Art rdf:about=”&SmartMuse;Art1”/>
<SmartMuse:User rdf:about=”&SmartMuse;User8012314699”>
   <SmartMuse:hasInterest rdf:resource=”&SmartMuse;Art1”/>
</SmartMuse:User>

The important thing to consider here is that we have described the Art as an instance of art concept on our museum ontology. So the predicate hasInterest is defined on the range of instance of art, Art1, and the weight recorded in the profile corresponds to the weight of the Art1 instance.
4.1.2.2 Knowledge

If we want to state that;

“User has knowledge in science of astronomy and is very interested in it and information is rather trusted and can be shared”.

We can describe it using:

<Reference to context, hasKnowledge, user, "astronomy", 0.5, 0.7, 1>

A sample RDF syntax for this profile slice can be presented as following:

<SmartMuse:Astronomy rdf:about="&SmartMuse: Astronomy1"/>
<SmartMuse:User rdf:about="&SmartMuse: User8012314699">
  <SmartMuse:hasKnowledge rdf:resource="&SmartMuse:Astronomy1"/>
</SmartMuse:User>

The same as previous example, we can consider here is that we have described the Astronomy1 as an instance of Astronomy concept on our museum ontology. The same as previous example, the weight recorded in the profile corresponds to the weight of the Astronomy1 instance.

4.1.2.3 Abilities

We can describe intellectual abilities of the user in the following manner:

“User has sensory impairment (blindness) and the level of blindness is 70% and information is more or less trusted and is totally private”.

We can either describe it as following:

<Reference to context, hasDisability, user, "blindness", 0.7, 0.5, -1>

Or we can describe it as following:
We consider Reading ability as a subconcept of ability to read and we create and instance of it and relate it to user with hasDisability predicate.

<SmartMuse:AbilitytoRead rdf:about="&SmartMuse: User_Disability1"/>
<SmartMuse:User rdf:about="&SmartMuse: User8012314699">
  <SmartMuse:hasDisability rdf:resource="&SmartMuse: User_Disability1"/>
</SmartMuse:User>

4.2 Presenting user’s history of museum experience and visits

In order to document the history of visit to the museum, the following template can be used:

<Reference to Context, visited, "artifact name", atTime time value, Interest rate, trust, privacy>
For example the following record states that:

“With reference to a certain time (for instance day or night), user visited Venere artwork and liked it very much and user trusts the originality of the work visited and has average consensus on the disclosure of the information to the public.”

<Ref. to context, visited, “Venere”, atTime Day*, 0.8, 0.6, 0.5>

*Time could be expressed in different formats depending on the usage. For instance time could be also the season of the year or could be hours or minutes which visits took place. Contextual information and how to incorporate them into the profiled records are discussed in the following subsection.

If we consider Venere as a concept described using museum ontology or vocabulary then we can instantiate it for each of the visits by the visitor and record it using the described template.

For instance, we have the visit of Venere1 instance, at Day time with mentioned weight values, described using RDF/XML syntax below:

<SmartMuse:Venere rdf:about="&SmartMuse;Venere1"
SmartMuse:visited="atTime Day"/>

4.2.1 Incorporating context into visit history

As previously stated, we consider modeling context as a multidimensional entity, and we consider different dimensions for context. For instance, we consider time (hours, minutes, seconds or day, month, season), travel, physical environment and location (of user or artifacts) as certain dimensions of context. One of the most important things to consider is, on availability of any contextual information afore mentioned; such contextual information can have different weights.

In order to cope with availability more than one contexts and variety of weights of contextual information, we can define an ontological description context and on availability of contextual information we can provide reference to the contextual taxonomy.

For instance, we can describe a history of visit of the user, with availability of outdoor sensory information (physical environment dimension) in the following manner:

“User has visited Mnajdra (Maltese temple) and didn’t like it [Context: the humidity was 70%-probably a reason affecting the experience of user] and information is more or less trusted and is shared”.

We can describe it using:

<Ref. to humidity in physical context, visited, "Mnajdra", at Time Day, 0.1, 0.5, 1>

For instance, we have the visit of Mnajdra instance, at Day time with mentioned weight values, described using RDF/XML syntax below:
4.3 Presenting user target groups

Community models of users allow modeling a group of individuals with attributes which are shared between most of the members of community. We refer to such group models as user target groups.

A target audience, or **target group** is the primary group of people that something, usually an advertising campaign, is aimed at appealing to. A target audience can be people of a certain age group, gender, marital status, etc. (ex: teenagers, females, single people, etc.) A certain combination, like men from twenty to thirty is often a target audience. Other groups, although not the main focus, may also be interested.

For instance, a “teenage” target group is a group of users with the age attribute range between 12-18 years old, so all the users with value between this ranges, is considered a “teenager”. In museum context, target groups are considered as individuals who share common traits such as culture, ethnic or social affiliation, educational level, and leisure preferences. In order to model target groups using our generic template, we take into account the following example:

“User belongs to target group “**Scientists**” and has high interest in the group and trusts the information received (recommended) from the group and prefers the information to be private”.

We can describe it using:

```xml
<Ref. to context, belongstoGroup, user, "Scientists", 0.8, 0.7, -1>
```

Again, If we consider Scientists as a concept describing a subcategory of user groups based on knowledge, then we can instantiate it to describe a target grouping. For instance, we can describe the profile slice mentioned using RDF/XML syntax as following:

```xml
<SmartMuse:Scientists rdf:about="&SmartMuse;Scientists1"/>
<SmartMuse:User rdf:about="&SmartMuse;User8012314699">
  <SmartMuse:belongstoGroup rdf:resource="&SmartMuse;Scientists1"/>
</SmartMuse:User>
```

4.4 Presenting social relations of users

Social relations of users can be described and presented using our generic template. As we are using resource description format in order to describe and document the information content, we can reuse an existing vocabulary for describing social relations between users. For instance we can use FOAF (friend of a friend). FOAF uses RDF for describing user’s personal information as well as social connections.
For instance, we can describe social connections of users in the following manner:

“User Alice knows user Bob. She doesn’t like him but trusts him very much. Alice prefers that this relationship remains private”.

We can describe it using:

< Ref. to trust (FOAF) ontology, foaf_knows, foaf_person:"Alice", foaf_person:"Bob", 0.2, 0.85, - 1>

In order to map to RDF/XML, perhaps we can describe the example record as following:

<SmartMuse:User rdf:about="&SmartMuse;User_Alice"
    SmartMuse:Context="http://xmlns.com/foaf/0.1/"
    SmartMuse:Privacy="0.85"
    SmartMuse:Rank="0.2"
    SmartMuse:Trust="-1.0">
    <SmartMuse:foaf_knows rdf:resource="&SmartMuse;User_Bob"/>
</SmartMuse:User>
5 Bibliography

Grant Agreement Number: FP7-216923
Acronym: SMARTMUSEUM
Project title: Cultural Heritage Knowledge Exchange Platform

6 Appendix A: Complete List of Extended User Attributes

6.1 Table1: Basic User Perspective

<table>
<thead>
<tr>
<th>Basic User Attributes</th>
<th>Abilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AbilityTo Hear</td>
</tr>
<tr>
<td></td>
<td>AbilityTo See</td>
</tr>
<tr>
<td></td>
<td>AbilityTo Talk</td>
</tr>
<tr>
<td></td>
<td>AbilityTo Touch</td>
</tr>
<tr>
<td></td>
<td>AbilityTo Use Stairs</td>
</tr>
<tr>
<td></td>
<td>AbilityTo Walk</td>
</tr>
<tr>
<td></td>
<td>Reading Skills</td>
</tr>
<tr>
<td></td>
<td>Typing Skills</td>
</tr>
<tr>
<td>Demographics</td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td>First Language</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
</tr>
<tr>
<td></td>
<td>Highest Education Level</td>
</tr>
<tr>
<td></td>
<td>Preferred Language</td>
</tr>
<tr>
<td>Preferences</td>
<td>Personal Preferences</td>
</tr>
<tr>
<td></td>
<td>Transportation Preference</td>
</tr>
<tr>
<td></td>
<td>Commercials availability</td>
</tr>
<tr>
<td>Device Preferences</td>
<td>Friend Finder Capability</td>
</tr>
<tr>
<td></td>
<td>Recommendation Capability</td>
</tr>
<tr>
<td>Privacy/Trust Preferences</td>
<td>Privacy Preference Scale</td>
</tr>
<tr>
<td></td>
<td>Trust Preference Scale</td>
</tr>
</tbody>
</table>

6.2 Table2: Advanced User Perspective

<table>
<thead>
<tr>
<th>Advanced User Attributes</th>
<th>Relationship</th>
<th>Role/RoleType</th>
<th>Inhabitant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Parent</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teacher</td>
<td></td>
</tr>
</tbody>
</table>
### 6.3 Table 3: Context Perspective

<table>
<thead>
<tr>
<th>Location</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SpatialLocation</td>
</tr>
<tr>
<td></td>
<td>VirtualLocation</td>
</tr>
<tr>
<td></td>
<td>Timestamp</td>
</tr>
<tr>
<td><strong>Physical Environment</strong></td>
<td></td>
</tr>
<tr>
<td>Humidity</td>
<td></td>
</tr>
<tr>
<td>LevelOfWind</td>
<td></td>
</tr>
<tr>
<td>LightLevel</td>
<td></td>
</tr>
<tr>
<td>NoiseLevel</td>
<td></td>
</tr>
<tr>
<td>OpeningHours</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td></td>
</tr>
<tr>
<td>Weather</td>
<td></td>
</tr>
<tr>
<td>Transportation Availability</td>
<td></td>
</tr>
<tr>
<td><strong>Visit</strong></td>
<td>Plan / Goal</td>
</tr>
<tr>
<td>First Time Visitor</td>
<td></td>
</tr>
<tr>
<td>VisitorTypology</td>
<td>Busy</td>
</tr>
<tr>
<td></td>
<td>Selective</td>
</tr>
<tr>
<td></td>
<td>Greedy</td>
</tr>
<tr>
<td>Companion</td>
<td>Alone</td>
</tr>
<tr>
<td></td>
<td>WithFamily</td>
</tr>
<tr>
<td></td>
<td>WithFriends</td>
</tr>
<tr>
<td></td>
<td>WithTouristGroup</td>
</tr>
<tr>
<td></td>
<td>SchoolGroup</td>
</tr>
<tr>
<td>Motivation</td>
<td>Adventure</td>
</tr>
<tr>
<td></td>
<td>ArtAndCulture</td>
</tr>
<tr>
<td></td>
<td>WellnessAndRelaxing</td>
</tr>
<tr>
<td></td>
<td>Educational</td>
</tr>
<tr>
<td>Duration</td>
<td>Hourly-based</td>
</tr>
<tr>
<td>Season</td>
<td>Autumn</td>
</tr>
</tbody>
</table>

Grant Agreement Number: FP7-216923
Acronym: SMARTMUSEUM
Project title: Cultural Heritage Knowledge Exchange Platform
<table>
<thead>
<tr>
<th>Age</th>
<th>AgeGroup</th>
<th>Child</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Teenage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elderly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GroupSize</th>
<th>Knowledge</th>
<th>Culture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Tours</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour Types</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour Availability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour Times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour AgeGroup</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour Size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour Fee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tour Duration</td>
</tr>
</tbody>
</table>

6.4 Table 4: Group Perspective

- Spring
- Summer
- Winter
- Guided Tours
### Appendix B: Short List of Extended User Attributes

<table>
<thead>
<tr>
<th>Attribute Name</th>
<th>Attribute Type</th>
<th>Attribute Value</th>
<th>Default Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>EducationLevel</td>
<td>String</td>
<td>All, School, HighSchol, College, University</td>
<td>School</td>
<td>Highest education level of the user. The extension is aimed at understandability of the artifacts, the more complex the artifact is the higher the level of the education of the user needs to be.</td>
</tr>
<tr>
<td>ArtifactLanguage</td>
<td>String</td>
<td>All, English, Italian, French, German, ...</td>
<td></td>
<td>The language of the artifact.</td>
</tr>
<tr>
<td>ImportanceLevel</td>
<td>String</td>
<td>High, Normal, Low</td>
<td>Normal</td>
<td>Importance of the artifact (with regards to artist's importance or just attractiveness). For instance, most attractive artifacts could be recommended to users.</td>
</tr>
<tr>
<td>LightLevel</td>
<td>String</td>
<td>High, Normal, Low</td>
<td>Normal</td>
<td>Measured level of light used over the artifact or needed for artifact visit. (This attribute is aimed at personalizing the user's with visual impairment.)</td>
</tr>
<tr>
<td>NoiseLevel</td>
<td>String</td>
<td>High, Normal, Low</td>
<td>Normal</td>
<td>Measured level of noise around the artifact. (This attribute is aimed at personalizing the user's with audial impairment.)</td>
</tr>
<tr>
<td>OpeningHours</td>
<td>String</td>
<td>All, Mornings, Evenings, Afternoons</td>
<td>All</td>
<td>The availability duration for a certain artifact.</td>
</tr>
<tr>
<td>Duration</td>
<td>Int</td>
<td>1, 2, 3, 4, 5</td>
<td>1</td>
<td>Minutes needed to visit a certain artifact.</td>
</tr>
<tr>
<td>Season</td>
<td>String</td>
<td>Summer, Fall, Winter, Spring</td>
<td>Summer</td>
<td>Season a certain artifact can be visited.</td>
</tr>
<tr>
<td>TourName</td>
<td>String</td>
<td>Galileo's Telescopes, Knights of Malta, ...</td>
<td></td>
<td>Name of the physical (or virtual, or online) tour to the online museum.</td>
</tr>
<tr>
<td><strong>AgeGroup</strong></td>
<td><strong>String</strong></td>
<td><strong>All</strong></td>
<td><strong>Target grouping the users of the museum based on their age group.</strong></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>KnowledgeGroup</strong></td>
<td><strong>String</strong></td>
<td><strong>School, HighSchool, University, Researcher, All</strong></td>
<td><strong>All</strong></td>
<td><strong>Knowledge group the user belongs to with respect to his/her academic knowledge.</strong></td>
</tr>
<tr>
<td><strong>CultureGroup</strong></td>
<td><strong>String</strong></td>
<td><strong>West, Asia, East, Africa</strong></td>
<td><strong>All</strong></td>
<td><strong>General geographical categorization of the cultures the work could be recommended to.</strong></td>
</tr>
<tr>
<td><strong>VisitorGroup</strong></td>
<td><strong>String</strong></td>
<td><strong>Casual, Usual, Special</strong></td>
<td><strong>Usual</strong></td>
<td><strong>Grouping of the users based on their attention to exhibition. Categorization is roughly equivalent to visitor typology (Busy, Greedy, Selective).</strong></td>
</tr>
</tbody>
</table>