FIDIS

Future of Identity in the Information Society

Title:	"D2.3: Models"
Author:	WP2
Editors:	Thierry Nabeth (INSEAD, France) and others (see the table contributors)
Reviewers:	Emmanuel Benoist, Bernhard Anrig (VIP, Switzerland) Martin Meints, Marit Hansen (ICPP, Germany) Mark Gasson, Kevin Warwick (Reading University, UK)
Identifier:	D2.3
Type:	[Deliverable]
Version:	2.0
Date:	Thursday, 06 October 2005
Status:	[Final]
Class:	[Public]
File:	fidis-wp2-del2.3.models.doc

Summary

The objective of this document is to present, in a synthetic way, different **models of representation of a person** ("person schema") that can be used in different application domains.

In particular it presents:

- Different **perspectives for modelling of the person** in Information Systems.
- Different **categories of attributes** (or data schema) that can be used to define a person, and the different domains in which they are used;

This document is aimed at an audience of non-experts and experts who are interested in a broad overview of existing models of representation of the person in the Information Society.

Information Society

Copyright Notice:

Future of Identity in the Information Society (No. 507512)

Members of the FIDIS consortium

1.	*Goethe University Frankfurt*	Germany
2.	*Joint Research Centre (JRC)*	Spain
3.	*Vrije Universiteit Brussel*	Belgium
4.	*Unabhängiges Landeszentrum für Datenschutz*	Germany
5.	*Institut Europeen D'Administration Des Affaires (INSEAD)*	France
6.	*University of Reading*	United Kingdom
7.	*Katholieke Universiteit Leuven*	Belgium
8.	*Tilburg University*	Netherlands
9.	*Karlstads University*	Sweden
10.	*Technische Universität Berlin*	Germany
11.	*Technische Universität Dresden*	Germany
12.	*Albert-Ludwig-University Freiburg*	Germany
13.	*Masarykova universita v Brne*	Czech Republic
14.	*VaF Bratislava*	Slovakia
15.	*London School of Economics and Political Science*	United Kingdom
16.	*Budapest University of Technology and Economics (ISTRI)*	Hungary
17.	*IBM Research GmbH*	Switzerland
18.	*Institut de recherche criminelle de la Gendarmerie Nationale*	France
19.	*Netherlands Forensic Institute*	Netherlands
20.	*Virtual Identity and Privacy Research Center*	Switzerland
21.	*Europäisches Microsoft Innovations Center GmbH*	Germany
22.	*Institute of Communication and Computer Systems (ICCS)*	Greece
23.	*AXSionics AG*	Switzerland
24.	*SIRRIX AG Security Technologies*	Germany

Versions

Version	Date	Description (Editor)
0.1	January 2005	INSEAD: Thierry Nabeth. Initial release
0.2	24 February 2005	ICPP: Martin Meints. (first version) Types of IMS with respect to social Systems
	7 April 2005	JRC: Sabine Delaitre. Ubiquitous Services, CC/PP and UAProf
	7 April 2005	TUD: Sandra Steinbrecher. Feedback + reference Interoperability
	11 April 2005	Univ. Reading: Mark Gasson. Biometrics perspective.
	12 April 2005	TUD: Richard Cissée. vCards, LDAP
	15 April 2005	VIP: Bernhard Anrig, Emmanuel Benoist. Identification in government
	9 May 2005	VIP: Emmanuel Benoist. Review of the deliverable
	9 May 2005	ICPP: Martin Meints & Marit Hansen. Review of the deliverable
	23 May 2005	Univ. Reading: Mark Gasson. Proof read of the deliverable.
1.0	23 May 2005	INSEAD: Thierry Nabeth. Final version
2.0	30 September 2005	INSEAD: Thierry Nabeth. Glossary, global picture, PKCS#9, etc.

Foreword

FIDIS partners from various disciplines have contributed as authors to this document. The following list names the main contributors for the chapters of this document:

Chapter	*Contributor(s)*
all	Thierry Nabeth (INSEAD)
	Mark Gasson (Reading)
Categorising the IMS; a non technical perspective	Martin Meints (ICPP)
Directory systems and standards (vCard, LDAP, etc.)	Richard Cissée (TUD)
Interoperability	Sandra Steinbrecher (TUD)
Identity schemas in ubiquitous (and mobile computing)	Sabine Delaitre (JRC)
Biometrics & AmI	Mark Gasson (Reading)
Identity & government	Bernhard Anrig and Emmanuel Benoist (VIP)

Table of Contents

Executive Summary

The objective of this document is to provide an overview of different models of person representation that can be used in different application domains.

In particular it presents:

- Different **perspectives for modelling of the person** in Information Systems.
- Different **categories of attributes** (or data schema) that can be used to define a person, and the different domains in which they are used;

It also points out

- Different **processes** that can be used to manage *person's information*;
- Different approaches used to **acquire** *person's information*

Note: these two later aspects (processes and acquisition) are not considered as central in this document, and are only present to facilitate the comprehension of the modelling of the person. These two aspects are more precisely covered respectively in WP3 (IMS Technologies) and WP7 (Profiling).

This document is aimed at an audience of non-experts and experts who are interested in a broad overview of the existing models. A more in-depth presentation of models and approaches will be presented in the more specialised workpackages of FIDIS[1].

This document does not aim to be exhaustive, rather it aims to help the reader better understand the level (and detail) to which the identity of a person can be represented in different domains and what the processes used for dealing and manipulating it are.

This document is organised into sections providing:

- An introduction presenting the objective and identifying the issues of this document
- An overall presentation of the management of identity categorisation, of the identity representation, as well as the different approaches used to acquire this identity information (*person's information*)
- An overview on different categories of attributes used to represent the persons, and how the attributes are used in different application domains
- An overview of the representation of the person in different application domains, and of the different standards that can be used "identity-informed" Information Systems.
- A concluding section
- An appendix listing standards for the representation of the person

---

[1] For instance WP3 provides in D3.1 a more technical overview of identity management systems, WP4 is concerned with interoperability issues, and WP7 with profiling

1   Introduction

1.1 Scope

This document, created in the context of the workpackage 2 of the **FIDIS Network of Excellence**, proposes to give an overview of the different **models (schemas)** that are used to represent people's information in **different application domains**.

This document also points out some peripheral aspects to this representation, such as the management of identity, or the acquisition of this identity. However these aspects –which are more thoroughly covered in WP3 (Technologies) and WP7 (Profiling) of FIDIS- are only present as a means for facilitating the comprehension of this document.

Besides, this document aims at making a (partial) inventory of different standards and schema that are used for the representation of the person in Information Systems. This inventory which is presented in the Annex includes a variety of standards of representation of the person (and includes "schemas" that are used in many application domains such as mobility, justice or human resources). It does not however cover the different protocols standards that are used for the manipulation of this "Identity" (for instance SIP, SAML or PKIX are not covered and only PKCS#9 of PKCS is presented).

The audience of this document comprises both non-experts (citizens, employees, civil servants, …), who need to have an understanding of the way that identity is managed in different application domains (and in particular what the categories of people's attributes that are dealt with or manipulated are); and experts who want to get a more global (multidisciplinary) understanding of the different models that are used to manage identity (in particular in the application domains that they are less familiar with).

This document tries to be comprehensible (but not simplistic) even to the non-specialist, and favours easy comparison of the different approaches for dealing with identity. The emphasis has been put on comprehensibility of the different models, rather than completeness (more detailed descriptions are to be found in the more specialised workpackages).

This document is part of a series of deliverables aiming at facilitating the comprehension of the identity concept including "D2.1 Inventory of identity terms" aiming at identifying the terminology used for dealing with identity and "D2.2 Set of Use Cases and Scenarios", which provides examples, cases, stories and scenarios which deal with identity issues.

1.2 Objective

The objective of this document is twofold: firstly, to contribute to the awareness and the understanding of the models of representation of personal information that are currently used in Information Systems, and in particular in IMS (Identity Management Systems). Secondly, more as a consequence, to facilitate cross fertilisation of the approaches dealing with identity representation in various application domains.

For instance, this document may help to answer (or provide some guide) to a set of identity-related questions that different audiences may ask, such as:

- From the end user: What can an Information System really "know" about me?
- From the IMS designers: How do different application domains deal with identity (in particular one that I do not know)? Are there some standards that I could use or that could help me structuring my work?
- From the researcher: What is done in other disciplines? Can I get inspired, and find new ideas?
- From the defenders of privacy: What categories of information are stored in Information Systems?
- From government officials: What information on people could I potentially exploit to provide better public services? (On the condition that the usage of this information is in conformance with the law).
- For companies exploiting customer information: What information can I get about my clients (existing or potential)? (On the condition that the usage of this information is in conformance with the law).

This document relies in particular on a review of different standards used for representing identity (such as LDAP for directory management, HR-XML used in human resource, CIQ - Customer Information Quality - used in direct-marketing, or Global JXDM used in law enforcement) to understand the level of representation of identity of the person and the usage in different application domains.

Note: The description of the standards, which is out of the scope of this document (and is something that may subsequently be addressed in WP4 'interoperability'), is only provided in the Annex to illustrate more concretely some of the attributes used to represent *person's information* in digital infrastructures.

1.3 The content and structure of this document

The first section of this document is this introduction that sets the scope, the objective and the structure of this document.

The second section provides an overall presentation of the management of identity categorisation and of the representation of this identity (or how the person is represented in Information Systems).

The third section presents in more detail the different categories of attributes used to model the person, and how they are used in different application domains.

The next section makes an overview of the representation of the person in different application domains, and of the different standards that can be used "identity-informed" Information Systems.

The last section concludes the document and provides some directions for future work.

This document finishes with an annex presenting some of the standards (LDAP, vCard, HR-XML, IMS LIP) that are used to represent *person's information*.

2 The Management and the Representation of Identity in Information Systems: an Overview

This section provides a categorised overview of how identity is managed in Information Systems (for instance how this identity can be controlled by the user, by a trusted third party, by a non-trusted third party), and how this identity can be represented (for instance, the different categories of identity attributes that can be distinguished). For a better understanding two perspectives are taken: the social perspective and the perspective of a person or individual. This section also indicates what different approaches can be used to acquire identity related information to describe a person.

Parts of the chapter 2.1 are summarised for a better understanding of "Del 5.2 Thematic Workshop on ID-theft, privacy and security" and "Del 3.1 Overview on IMS".

2.1 Categories of IMS, and their application in social systems

2.1.1 Types of IMS

Taking a look at the market of existing IMS (Identity Management Systems), on prototypes, concepts and IM-related tools, we determine several approaches towards IMS which differ, e.g. in

  a. Procedure of management (by whom? which operations on data offered?)

  b. Type of managed data (possession of the data: personal or organisational? comprehensive profiles or selection of roles or partial identities? anonymity or identifiability?)

With respect to these properties, we observe three main types of IMS (which are explained and further investigated in Deliverable 3.1):

Type 1: IMS for account management

Type 1 IMS are used for implementing authentication, authorisation, and accounting[2]:

  a. Procedure of management mainly centralised; within groups of organisations which trust each other administration can be somewhat decentralised to a group of (trusted) administrators (federated identity management systems, see D3.1)

  b. Type of managed data mainly organisational, roles determined by the provider of the central system, data in most cases linkable to the individual person

---

[2] http://infosecuritymag.techtarget.com/2002/apr/cover_casestudy.shtml,
http://www.oracle.com/technology/products/id_mgmt/index.html
http://www3.ca.com/Solutions/ProductFamily.asp?ID=4839

Type 2: IMS for profiling of user data by an organisation

Type 2 IMS are e.g. for logging or data warehouse tools which support personal profiling e.g., personalised services, or group profiling such as the analysis of customer behaviour[3]:

a. Procedure of management mainly centralised

b. Type of managed data personal as well as organisational, mainly comprehensive profiling, linkability to an individual person is often the target of these approaches

Type 3: IMS for user-controlled context-dependent role and pseudonym management

a. Procedure of management mainly decentralised

b. Type of managed data is mainly personal; tool-supported selection of roles and relating partial identities, privacy protection and in some cases anonymity is the target of these approaches

2.1.2 Social systems

From a traditional sociological point of view, the specific identity as "person" (Luhmann 1991) is a construction through a specific situation which is mainly formed by a specific social system. Sociologists know at least three types of social systems (Luhmann 1997):

- **Interactional systems**
  - forms of community in which participants are not subject to documented rules, but nevertheless schemes apply; examples are neighbourhood, friendship, spontaneous encounters (Kesterling 2000)
- **Organisational systems**
  - characteristics are membership and effective production of decisions; examples are public bodies, institutes and companies (Baecker 1999, Luhmann 2000)
- **Functional systems**
  - economy, law, politics and science as main "self-conducted" communication systems

Functional systems are characterised by communication that is specialised in functionality. Organisations have to be connectable to all four main functional systems, but normally have a main emphasis on one of them:

- **Economics**: payment / non-payment;
  - programme: price; generic person: e.g. "client" and "employee"
- **Law**: legal / non-legal;
  - programme: laws; generic person: e.g. "citizen"

---

[3] http://www.lumeria.com/what.shtml
http://www.epic.ca/TechnologyDay/October05_2004/MoreInformation/Presentations/RandallBartsch%20-%20Identity%20Mgmt.pdf

- **Politics**: power / non-power;
  - programme: political programmes; generic person: e.g. "responsible citizen" in the meaning of the French term "citoyen"
- **Science**: true / false;
  - programme: theories and methods; generic person: e.g. "the human being"

Sociologists understand social systems as pool of schemes, events and communicational components which are used by persons. In this context persons could be understood as personal inventory of social systems; the thinking of persons taking part in communication is focused by the mentioned components within the appropriate social systems. The different types of social systems operate on different addressing modes to link these communicational components.

The social subsystems reproduce particular patterns of communication that have particular social functions (cp. e.g. the above-mentioned generic persons, which also correspond to typical roles within these systems). These functions, in turn, generate so called "pointed sense horizons" (Luhmann 1997) for organisations, which create particular sets of expectations (role conformity as "client", "citizen", "responsible citizen", "human being") for the persons acting in them.

Communication thus personally involves persons in socially typical processes, e.g. as a citizen of a certain country, as a company's client, as a patient in hospital or as teacher in a school, but also as "my friend", "my mother" or "my neighbour".

2.1.3 Application of IMS within social systems

Looking at the described types of IMS and the different social systems we observe specific connections between them.

Type 1 IMS: organisational function

Type 1 IMS are mainly established within organisational systems where we find e.g. enterprises (economic subsystem) or public bodies (political functional system) using them. The data managed are mainly centrally administrated account data for employees and customers / citizen. The type of management in most cases is strictly organised, role assignments regulated by processes which are well documented. The person behind the managed account in most cases has small influence on the roles she or he is allowed to take within this IMS; it matches usually quite well with the roles taken or assigned within the organisational system. In this case we can speak of an "assigned identity".

Type 2 IMS: profiling function

Type 2 IMS, much like Type 1, are mainly used by organisational systems; we find here for example enterprises doing profiling on customer data they receive off-line (e.g. via questionnaires) or on-line (web, cash box using customer account cards etc.). Another

example is the profiles generated by public bodies e.g. in the case of criminal investigations. The employees doing the review are mandated by the organisation and specially educated to do profiling. The techniques and parameters used for the profiling are in most cases (public or trade) secrets – the profiled persons get, in most cases, no feedback as to how the profiling was done and what resulted from the profiles. Looking at the results of profiling based on authentic personal information, but in central aspects not being influenced by the profiled person itself, we can speak of a "derived identity". Profiling is done on both members and clients of the organisations (for more details please refer to deliverable 7.2).

Type 3 IMS: individual function

Type 3 IMS are used in both interactional as well as in organisational systems although organisational aspects are dominating. We suggest examining internet-typical communications like chat, blogging and avatars for their interactional character. In this context the new description of borderlines between interactional and organisational systems might be necessary. To use those services with interactional aspects we find requirements such as written communication, membership and the use of services of telecommunication providers which are typical for organisational systems.

Another example of the use of type 3 IMS within organisational systems is the use of different web services with different accounts by a customer (client role outside of the organisation). In this example type 3 IMS are the decentralised, user driven view on account management systems (type 1 IMS) for public web based services.

Due to the existence of alternate services and the user control of which one to use with which account, we can speak of "chosen identity".

2.2 Modelling the Person in Information Systems

IMS, or systems that integrate an IMS component, use a variety of attributes to represent (model) a person and to later manage this *person's information*. For instance attributes can be used to represent the identifiers of a person (such as name or pseudonym), her biological characteristics (gender, hair colour), her location (permanent address or geo-location at a given time), competences (diploma, skills), social characteristics (affiliation to groups, friends), and even behaviours (personality or mood).

In some cases, standards and specifications have even been elaborated to facilitate the design and the interoperability of such systems. For instance LDAP schemas have been defined to specify how to represent *person's information* in directories. In the human resources domain, the HR-XML specification has been elaborated to standardise the way information about employees are represented in the management software (see the Annex for an overview of different standards and specification for people representation).

Actually, an important strand of research has been conducted for many years in user modelling, aiming at enhancing the interaction between users and systems via the design of adaptive systems (Fischer, 2001; Brusilovsky, 2001; Stephanidis, 2001; Kay, 2000; Andre et al., 2000; Fink and Kobsa, 2000, etc.). The goal of research on personalisation is to improve the efficiency and effectiveness of user interaction by taking into account the specificity of the

end-user (such as his cognitive style, or his competence) as well as the context of activity of this user (for instance the current tasks in which he is engaged or the organisational context (Nabeth, Angehrn, and Balakrishnan, 2004)). Practically, adaptive systems are able to support the user better by filtering the irrelevant information (reducing cognitive load), by delivering this information at the right time, by choosing a form of delivery that maximises its impact on this user, or by proposing very contextualised help. Research on adaptive systems has been conducted for applications in a number of domains such as e-learning Diogene (2002), e-commerce (Kobsa et al., 2000) or knowledge management (Razmerita, 2004).

In this document, we will not enter into the details of these theories or standards (which would be out of the scope of this document) but just make a tentative attempt to find some way to categorise these attributes.

In particular, we are going to present categorisations according to a

- temporal perspective
- functional perspective
- domain perspective

2.2.1 Temporal categorisation

The different attributes can be first categorised by the level of permanence of the information they represent:

- permanent – given
- permanent – acquired
- persistent situations
- temporal state

Permanent – given: Some attributes are used to represent some permanent (given) characteristics that were given to a person and on which he usually has no influence. Examples include for instance the biological characteristics (gender, eye colour, fingerprint, etc.), some socio-cultural-economical characteristics (parents, country of birth, etc.), basic personality traits (for some psychologists such as Hans J. Eysenck, personality has an important genetic basis), etc. Some exceptions such as gender changing have to be made regarding the person's non-influence.

Permanent – acquired: Some other attributes are used to represent permanent (acquired) characteristics that the person was able to acquire because of some circumstances or because of a deliberate action. Examples include qualification (either because of a deliberate action like graduating at a University or because of circumstances like learning a new foreign language during the stay in a country), behavioural characteristics.

Persistent situations (or states): Other attributes are used to represent a situation that is not permanent, but that has some persistence (for instance several years). Examples include the address of a person, a job position (title, employer, etc.), marital status, social status, or a network of friends.

Temporal states: Finally, other attributes are used to represent very temporary situations that are attached to a particular context. Examples in this case include for instance the geographical position of a person at a given time or the mood of the person.

2.2.2 Functional categorisation

The attributes can also be categorised according to some functional characteristics.

Examples of such categories of attributes include:

- identification (such as a name, the social security number, password, …)
- location (geographical location, addresses, ...)
- biological characteristics (biometrics, age, …)
- personal - psychological (personality, psychological state, preferences, …)
- group - sociological (affiliations, social group, social networks…)
- …

This categorisation will be detailed in the following chapter.

2.2.3 Categorisation by domain / spheres

The attributes can also be grouped according to their application domain / activities in which these attributes are used such as:

- work (employer, title, roles, expertise, acquaintances, work context / tasks, ...)
- education (university, degrees, …)
- leisure (pseudo used in chat spaces, friends, sexual preferences, …)
- government (registration information, tax services, …)
- justice and police (criminal files, …)
- health (social security number (ssn), medical information, …)
- …

A subsequent chapter of this document will present more in detail how the categories of attributes are managed in different domains of application.

2.3 Acquisition of the Person's Information (Profiling)

Profiling of a given user is the process of obtaining the values associated with the different attributes that constitute the user model (note: refer to the work conducted in WP7, especially D7.2, for a more elaborated study of profiling).

The different means that can be employed to get the information associated with the different attributes include:

- The direct entering of the personal information by the end-user

- The extraction of this information from querying existing data sources (such as databases) or/and captured during different processes (such as the recording of a transaction)
- The calculation from existing attributes (simple algorithms or expert systems)
- The extraction via the mining of information

Of course, one should be aware that the quality of the information differs, and depends largely on the means that have been employed to get this information. For instance, one can easily imagine that certain information originating from a governmental database is much more reliable than the same information found in a personal web page which the user has entered by her- or himself. In a similar way, we can assume that information that has been calculated or inferred is more prone to error than information that has been entered directly (although the phenomenon of obsolescence of information can mitigate this assertion).

Similarly, the control of this information is strongly correlated to the means that have been employed to collect the information. For instance, information that is present in web pages is totally controlled by the users themselves, whereas information that is present in a governmental database is principally controlled by a third party (the government). In the later case, legislation and the possibility of the end user bringing some correction can help to share part of this control. In the case where the information has been extracted by some data mining procedures, it is totally controlled by some third parties. There is almost no possibility of intervention by the end-users (who may simply be totally unaware that their personal information is being exploited).

2.3.1 Direct entering of the personal data

Direct entering of personal data consists of mechanisms in which the end users are able to enter explicitly their personal information. For instance, a typical example of such systems is an online system in which the users have to describe themselves by specifying their name, addresses, preferences and other characteristics.

The Type 3 IMS (individual function), presented previously, represent a typical category of systems that employ this method.

The personal data directly entered is mostly under the total control of the user who is able to modify this information whenever he likes.

This mode of collection of the personal data appears to better preserve the privacy of the user than the centralised solutions, although it is not without a certain number of limitations. Firstly, this information may not be very reliable nor up-to-date, since it relies on the willingness of the users to enter this information and to be honest. Equally, the users may even involuntary introduce some errors that originate from an incorrect perception of reality (such as rationalisation). Secondly, the entering of this information and its update can be considered too time consuming for the user who may not be ready to spend the effort.

The typical attributes that can be captured in this way include name, addresses, pseudonyms, short descriptions (such as picture) and preferences (basic).

2.3.2 The extraction from data sources and from processes

In this case, the values associated to the attributes originate from two different sources: (1) databases; and (2) processes.

In the first case, the databases may be governmental (such as police or tax), human resource databases (enterprise resource planning and knowledge management systems such as payrolls, or training information) or health file databases (managed by hospitals or by social security units).

In the second case, the data can originate from a series of processes that can be used to capture the data (and that will be stored in databases). Examples of such processes include e-commerce systems (such as Amazon) and fidelity programs that can capture the history of different transactions associated with each of the customers, or virtual community systems that can capture the history of activities of the different members (such as age in the community, and number of posting).

The type 1 IMS (organisational function), presented previously, represents a typical category of systems that employs this method, although it can also be used in the type 3 IMS (individual function).

The personal data that is present in databases or captured via a set of processes is mostly outside the user's control (the possibilities of correction by the end user are often limited). These data are also often very regulated by some legislation specifying the type of data that can be represented, the possible usage of this data, including combining databases.

Even if this mode of collection of personal data appears to be more intrusive to people's privacy, it is not without some advantages, even for the people themselves. First, the data captured via this means can be considered much more reliable, since it directly reflects the activities of people, and not only the perception of these activities. Second, because this data collection is automatic, it can be considered less demanding for the end-users.

The values of many attributes that can be recorded in this way include characteristics that have a certain level of permanence, while other categories of *person's information* can include all the transactions (commercial or not) in which the people have been engaged.

2.3.3 Data calculated and inferred from other attributes

In this case, unknown values associated to particular attributes originate from the calculation of other attributes (typically the ones that have been extracted from the previous two methods). This category is relatively similar to the category previous described, however, it differs in the level of sophistication of the systems that make use of it. Notably, these are more frequently used in Type 3 IMS (individual function) applications that use it to provide some level of adaptability (for instance in e-learning systems or e-commerce systems).

The reliability of these calculated attributes is generally less accurate than for non-calculated attributes. For instance in Amazon the assertion "a customer that has bought a book about children is interested by children and is likely to buy other books about children" is only correct in average, since they may only have bought this book once in order to offer a present to somebody else.

The level of control on these calculated attributes is often limited by the simplicity of the algorithm used, and the way it was configured for the calculation. Thus, people that read the value of these attributes usually have, at best, only a vague idea about the underlying principles that have been used. For instance, a calculated attribute could be a level of risk that a bank could calculate on a particular client, which results from a combination of values of attributes such as the gross salary of the person, the assets such as real-estates that the person may own, his family status, or the postal code of his place of living or even his ethnic origin. Another application is certain e-commerce websites, where the preferences of a customer are determined automatically.

2.3.4 Data extracted via mining the information

The extraction of values via data mining techniques could appear similar to the previous calculated methods. They differ however in that the algorithms are being applied globally to the data of (very large) groups of people, and not on the data set that is associated with a single person. The algorithms used are also of a more statistical and probability based nature, and often rely on the use of Heuristics. Finally, these algorithms may also be used to help the creation process of the user model itself, and in particular help to determine the set of attributes required to "summarise" the problem (for instance, in a banking application, an algorithm may determine that the knowledge of the age and of the postal code information represent sufficient information to discriminate a reliable customer from an unreliable one, with a limited risk of error).

Type 2 IMS (profiling function), presented previously, represent a typical category of systems that employs this method.

The types of attributes that are extracted via mining typically include people related categories such as social categories or life styles. These attributes can be considered to be more abstract and less directly associated to the individuals.

At a more micro-level, these attributes can represent some user characteristics and behaviours that can be automatically extracted from the use of some Information Systems. For instance such attributes, in the context of an e-commerce system, can reflect reliability characteristics (likeliness of fraud), and, in the context of a virtual community, can reflect the level of participation (such as the activity of the people in SourceForge.net).

3 The Categories of Attributes of the Person

The previous chapter has given us an overview of how the *person's information* is managed in Identity Aware Information Systems and in particular in IMS.

In this chapter, we are going to look more in detail what are the individual facet of the identity of the person, and in particular what are the different categories of attributes can be used to represent this person.

Note: The next chapter will provide a more integration perspective, and will look at how these different categories of attributes can be assembled in order to represent the identity of the person in an application context.

3.1 The attributes of the person: a categorisation

As described in the previous section, a variety of attributes can be used to represent the person. Each attribute is used to represent a particular characteristic or property of a person such as the identification of this person (how we refer to this person), some biological characteristic (the physical characteristic of this person), some location information (where the person lives, where the person is now), personal and behavioural characteristics (how the person behaves, the preferences of this person, etc.) or social characteristics (who this person knows, the affiliations of this person, etc.).

This chapter, shall analyse the different categories of attributes, and for each of them describe what they are, in which context they are used, how they are acquired and if there are any standards that are used to represent them.

3.2 Identification

3.2.1 Description

The identification attributes represent a category that is used <u>primarily</u> as a "referent" to a person. Their main function is to provide support for linkability.

The name

One of the most well known identifiers is **the name**. A name is a "label for a thing, person, place, product (brand name), and even an idea or concept, normally used to distinguish one from another" (Wikipedia). A person's name is principally used to refer to a particular person. For instance, the name of a person can be used in an article to indicate who the author of a particular writing is, i.e. to create a link (or association) between a document and its author. It is also used in people's communication (such as in a discussion) to make reference to a person that is not present. A name can also be used in Information Systems to reference a person.

Different names can be associated with a person such as a surname (inherited from the parents), given name (chosen by the parents), married name (acquired from the husband), nick

name (chosen by close acquaintances), pseudonym (chosen by the person, and used to reduce linkability) or stage name (a particular pseudonym chosen by an actor to help separate his private life from his work life).

Pure identifiers

Other categories of identifier exist that are often used for non-human processing. A typical person identifier is for instance the ssn (social security number). The ssn is an identifier that is issued to people by governmental agencies mainly for managing people's identification in their relationships with the government (and not only social security).

Other attributes also used as identifiers

It is important to indicate that other attributes that can be associated to the person can also be used as identifiers, such as biometric information.

More detail on identifiers can be found in FIDIS WP3, with deliverables related to identification and authentication.

3.2.2 Examples of attributes

- Name
    - o First name or given name
    - o Last name
    - o Married name
    - o Maiden name
    - o Nick name
    - o Pseudonym
        - ▪ Stage name
- Identifier
    - o ssn (social security number)
    - o Login
    - o used as identifiers (identifier is only a secondary function)
        - ▪ Biometric attributes
        - ▪ Etc.

3.2.3 Application domains

The identification attributes appear in every application domain (for instance the name of a person is present in almost every user model). These attributes have however a very central role in applications that focus on authentication and identification, such as applications which function to manage access control (such as a passport or other ID cards) or applications concerned with the management of contacts (business cards).

3.2.4 An example of the use of an attribute: the ssn (social security number)

Governments in general tend to make their habitants identifiable for different reasons. Several techniques for this purpose exist, as well as different legislations in various countries which imply different types. For the purpose of giving concrete examples, we will focus here on ssn implementation in three different countries, namely Switzerland, the US and France, and specifically what information is coded in the ssn.

The ssn in Switzerland

In Switzerland there is in fact no nationwide identification number available at the moment. This is the case for several reasons: first the federal structure of the country, second no one is obliged to possess either a passport or identity card, but also because many citizens are opposed to such a unique number in general. Notably however, there are plans for a unique number to be used for health care. At the moment, the only thing close to a unique identifier for people is the social security number (called AHV). This is a unique number attached only to persons who receive income and is delivered to the person usually when he/she starts employment. Hence only part of the population has and is therefore identifiable by such a number. Further, depending on different reasons, this number changes with time (marriage, change of name, sex, nationality, etc.). The number is used for different purposes, mainly for contacts between people and governmental agencies, and not only social security. Rarely is it used in the private sector.

Technically, the number consists of 11 digits, from which information about the holder can be deduced[4]: Consider for example the number 123'45'678'113. The first three digits (123 in the example) depend on the name of the holder, beginning with 100 for names between "A" and "Abi", with 101 for names between "Abi" and "Abl", and so on. The next two digits (45 in the example) denote the year of birth, the following three (678 in the example) the day and month of birth as well as the sex of the person, where roughly the first one denotes the quarter of the birth (1-4 for male, 5-8 for females) and the other two the day in the respective quarter. For example 678 means that the holder of the number is female and born June 16th (i.e. the 78th day of the 2nd quarter). The next two digits (11 in the example) are a running number delivered by a central office in Switzerland in order to differentiate between persons having identical first 8 digits. The second of these numbers tells if the holder is a Swiss citizen (numbers between 1 and 4) or foreign citizen (numbers from 5 to 8). The last digit is a weighted check sum used to ensure the integrity of the number when used and depends on the other 10 digits.

The SSN in the US

A concept very close to the Swiss one exists in the US, the SSN (Social Security Number[5]). Each person over 18 who receives income must have an SSN, although it is possible for anyone, including children to have one. In the US, the SSN is used in a variety of sectors besides social security, for example for opening a saving account; hence a large part of the

---

[4] http://www.ahv.ch/Home-D/allgemeines/31810612d.pdf

[5] http://www.ssa.gov/

population has a SSN. The SSN does not change in the person's lifetime (besides some very rare cases), only the physical card carrying the number is replaced, e.g. when changing name because of marriage.

Technically, the SSN consists of nine digits which decompose into area numbers (3 digits), group numbers (2 digits) and serial numbers (4 digits). The area number relates to the state the application for the number comes from. The group numbers are used to "break" the area numbers into different pieces following a specific scheme. The serial numbers run through 0001 to 9999 without further information contained therein. Hence there is almost no information about the holder contained in this number. Some ideas about the age might be gathered from the group number.

The SSN in France

The French Social Security Number[6] contains information, that is clearly readable, about the user. The first digit stands for the gender (1 for male and 2 for female), then comes two digits for the year of the birth and two other digits for its month. The next two digits represent the department of birth (one of the 95 administrative subdivisions of the country). The three next digits represent the city of birth and the last three digits are for the rank within the register of births in this city. It is therefore possible to know precisely the sex and age of the SSN holder, and from which region he or she comes from.

The differences

What is the difference between coding lots of information in the AHV-number (like in Switzerland) opposed to coding almost no information in the SSN (like in the US)? There is on one hand the problem of data protection: everyone in possession of the AHV number can deduce all things described above, e.g. age, sex of the holder. This information is not protected. On the other hand, everyone can – at least to some degree – check if some number really belongs to a person claiming to be its owner. Typically this is done by cross referencing data contained in some other credential like a passport (containing name and date of birth) with the info contained in the AHV number. On the other hand, the SSN allows the holder to not disclose any information about himself. Clearly this raises the problem of SSN falsification, fraud detection, etc.

3.2.5 Relevant standards and specifications

Most standards specify some attributes which have identification as a principal role. For instance the name of a person, if present, is the major representation specification in LDAP, vCard, HR-XML, IMS-LIP, JXDM, etc.[7]

Some specifications are however addressing more specifically the identification dimension, and in particular provide more sophisticated "identification attributes".

---

[6] http://www.dads.cnav.fr/TDS/stru0103.htm

[7] See the Annex for more detailed information about these specifications.

Such specifications include: LDAP (directory services); vCard (the digital business card); Liberty Alliance (and Microsoft Passport); JXDM (global Justice mark-up language).

More specifically, LDAP schema includes the "identification attributes" password and user certificate, and JXDM (used in the US) includes an attribute that is used to specify many (14) assigned ids of a person (SSNID, TaxID, DriverLicenseID, FBIID, StateID, AFISID, OtherID, RegisteredOffenderIndicator, FirearmSalesDisqualifiedIndicator, LicenseID, GeneralLedgerID, PersonHumanResourcesID, PersonVendorID, PersonNationalID).

Finally, it is important to mention biometrics attributes (presented in the next section), whose function is essentially to contribute towards identification.

3.3 Biological characteristics (biometrics, medical)

3.3.1 Description

The biological attributes represent the category that is used to represent the biological (or physiological) characteristics of a person. The representation of the biological characteristics can be done for several reasons such as identification, verification (access control), criminal investigation or healthcare.

Biometrical information

The first category of attributes that have already been mentioned is related to the **identification of the person** and includes all the **biometrical** information. The underlying premise is that some of the biological characteristics are permanent, intimately associated to the person, difficult to forge and unique enough so that they can be used for identification purpose. For instance they can be used to link a person to a passport or, in the context of a criminal investigation, to link the presence of a person to the scene of a crime.

The biometrical characteristics can vary considerably, and includes elements that are highly visible for the human (such as a Facial Features) or need some sophisticated mechanisms to be analysed (such as the DNA). These characteristics can either be physiological (passive), such as iris or face recognition or behavioural (active), such as lip movement, gait or keystroke dynamics. Within the physiological biometric methods we can distinguish between morphological methods (such as facial features, iris, fingerprint or palm geometry) and those being related to the senses (including voice, thermal patterns, body odour etc.). Biometric methods and their use for identification and verification are investigated further in FIDIS D3.2.

Physiological & medical information (patient data)

Another category of biological information is related to **healthcare** and includes the physiological characteristics that can be recorded in a **medical record**. Examples of biological information that can be recorder include: blood characteristics (pressure, level of albumin, cholesterol, etc.), known disease, etc.

It is important to mention that the use of these physiological characteristics can also be relevant outside of the medical domain, such as ability to practice a sport or to perform a job, insurance, etc., though in some case it raises a series of questions related to privacy protection .

3.3.2 Examples of attributes

- Biometric
  - Physiological (or passive)
    - Morphology
      - Facial features
      - Fingerprint
      - Palm geometry
    - Senses
      - Voice
      - Body odour
      - Thermal patterns
    - Other
      - DNA
  - Behavioural (or active)
    - Gait
    - Lip movement
    - Keystroke dynamics
- Physiological and medical
  - Physiology
    - Sex
    - Weight
    - Length
    - Strength
    - Biological clock (morning / evening)
  - State
    - Awake / asleep
  - Health characteristics
    - Known diseases
    - Vaccinations
  - Health instant state
    - Blood pressure
    - Temperature of the body

3.3.3 Application domains

The main application domains of biological characteristics are security (authentication / verification), and healthcare.

3.3.4 Relevant standards and specifications

Biometrics

Certain standards exist that can be used to specify biometrical characteristics, especially concerning the way processed biometric data, so called 'templates', are stored. The template formats depend highly on the algorithms used to process the 'raw' biometric data (such as a digital photo, in a defined digital format, of a face or iris). In some cases the algorithms are protected by patents. These aspects are further described in FIDIS D3.2.

For transporting of raw biometric data and templates, standards such as XCBF (OASIS XML Common Biometric Format) can be used. The International Committee for Information Technology Standards, a predominantly US organisation, is attempting to derive standards for interoperability (data exchange formats) for biometrics with M1[8], but is only one of several parallel organisations and initiatives (BioAPI, ISO SC37[9], BioFoundry x9.84 and others ..) deriving 'standards' for biometric information..

Finally, it is worth mentioning the representation of biometric information in JXDM (global Justice mark-up language), which can be used in a crime investigation context.

3.4 Location

3.4.1 Description

The location refers to the geographical position of the person. This property can apply to the instant position of a person (typically the place in which a person is present at a particular moment and that can be given by e.g. Geo-location (GPS) coordinates**,** or even the geographical mapping of an IP number), or a more permanent position of this person (typically the living or working addresses of a person).

3.4.2 Examples of attributes

- Address
    - Home
    - Work
    - Leisure
- Geographical location (instantaneous)
    - GPS coordinate
    - IP location mapping

3.4.3 Application domains

The location of a person is used in a variety of application domains. For instance the instant coordinate of a person represents a major element in **mobility applications**. Indeed, this information can be exploited to offer services that are geographically close to the current position of a person. GPS can also be used to reveal some of the behavioural characteristics of a person (Mountain and Raper, 2000).

Security and personalisation, in particular in **e-commerce applications**, can also make use of the geographical location of the person determined via IP address (albeit with not a very good level of reliability, because the table of IP addresses is not updated in real-time, and a range of IP address may be reallocated to users from a totally different region). In the first case, this

---

[8] InterNational Committee for Information Technology Standards: M1 http://www.ncits.org/tc_home/m1.htm

[9] ISO SC 37 "Biometrics" data interchanged format.

http://isotc.iso.org/livelink/livelink/fetch/2000/2122/327993/2262372/customview.html?func=ll&objId=2262372

information can help to identify some instances of fraud[10], and as a consequence, a payment with a credit card issued in a different country may not be accepted. In the second case, the knowledge of the IP address can help to determine the nationality of the person, and therefore result in an interaction that is aware of geographical characteristics (such as the language used).

The permanent locations of a person (the addresses) represent an attribute that is present in many identity management applications since it represents one of the most important means to reach the person in the off-line world (via the mail). For instance, the address of the person will be used in an e-commerce application to deliver the physical goods that have been purchased by the person.

3.4.4 Relevant standards

Addresses

We will not try to identify all the different specifications that define the address of a person since these are numerous. However we can mention the services and specification for which this information can be considered more important such as LDAP (directory services) or vCard (the digital business card). In the domain of commerce, CIQ (OASIS Customer Information Quality) together with xNAL (extensible Name and Address Language) provide a relatively detailed way to represent a person's address.

Geo-location

The applications and specifications that represent the geo-location of the person are rarer, and often relate to mobility. However, surprisingly this information does not seem to be present in the mobility standards that are used for mobile devices such as CC/PP (Composite Capabilities / Preferences Profile) proposed by W3C and UAProf (User Agent Profile) proposed by the WAP (Wireless Access Protocol) forum (See the Annexe for a description of CC/PP and UAProf). These latest standards are mainly used to manage user preference information that is typically used to customise the mobile device.

On the other hand, the Liberty Alliance developed a standard (ID-SIS-GL) that provides support for geo-location information representation (it defines a specialised schema to represent this information). This geo-location information includes the position of a principal, speed and direction related information, and information related to the quality of the position information.

---

[10] A brochure from the company Quova, Inc. indicates that, "in 2004, LexisNexis studied over 100,000 transactions executed by a major US online retailer, all with US credit card numbers. The study found that 75 percent of the identified fraudulent orders with US billing addresses had been placed from overseas. 97.9% of all transactions originating in Africa and 74.8% of all transactions originating in Asia (including Russia) were fraud". http://www.quova.com/technology/E-RET_FraudWP.pdf

3.5 Assets & transactions

3.5.1 Description

Assets & transactions relates to the information concerning owning a belonging and acquisition of "goods". These goods include a vast variety of "things" such as real-estates (a house), a digital item (such as a sound), a sum of money, or even a piece of information. The acquisition refers to the different processes used to transfer this information to a person, for instance via some purchasing action, or indeed any other means (for instance using a peer-to-peer network tool).

The management of the "assets & transactions" information is related to the association of items to a person (typically the items that belong to this person) as well as the historical data concerning this association. The role of the transaction attributes is to encapsulate this historical information.

3.5.2 Examples of attributes

- Income
  - Salaries
  - Other revenues
- Financial
  - Bank accounts
  - Loans
  - Stocks
- Physical properties
  - Real estates properties
  - Car
- Digital properties (digital items)
  - Sounds
  - Movies
- Socio-economical category

3.5.3 Application domains

Many application domains deal with asset & transaction information. The first obvious sets of applications are e-commerce, since they have a direct interest in making a commercial transaction with a person. For this purpose, companies sometime use data-mining techniques to profile their existing or potential clients to improve the effectiveness of their commercial campaigns. These companies are also interested in a better knowledge of their clients (what they own, what their financial capabilities are, what their socio-economical categories are) in order to adapt the interaction, and guide them (or orient them) to products that they are likely to buy. Of course, the access of this personal information (or set of attributes) may not be acceptable to the person.

The "market" of digital goods (for instance sound tracks) is slightly more complex since, with the advent of peer-to-peer networks, the exchanges of digital items may occur outside of a commercial context. In this case, a participant of a peer-to-peer network will display to other participants a profile describing the digital items that he "owns". This information may be

used by peer-to-peer nodes to identify the item to be exchanged, but can also be used by the major media companies to track violation of digital rights.

In the case of reputation based e-commerce systems (such as eBay) the transaction history information of the sellers and the buyers are made visible to all, providing transparency mechanisms in order to enforce trust.

E-banking represents another application domain that is likely to manage *person's information*, and in particular the "assets & transactions" attributes. The first motivation is, as in the previous case, to sell customers some financial products. The second motivation is to construct the risk profile of the person, and their capacity to reimburse a loan.

The government, and more particularly its tax department, is interested in accessing all this information, in order to calculate the financial contribution of the person to the society.

3.5.4 Relevant standards

Assets & transactions information, or attributes that represent information about people assets and transactions, are not often the subject of open specifications. For instance, government systems or e-commerce systems, may manage this information internally using proprietary standards. However, HR-XML, which is used to manage human resource information, does represent some employee financial information with attributes such as "Pay" (a brief description of the salary, rate of pay or bonus provided by the job/position) or the "employee StockPlan". Needless to say, the sensitivity of this information (and therefore the desire for confidentiality), in some of the European countries in particular, is considered high.

In the domain of electronic commerce, CIQ (OASIS Customer Information Quality) aims to deliver global, application-independent, open XML specifications for party/customer information and profile management.

Finally, it is interesting to note that RFID may in the future provide the possibility (under the condition that it will be authorised) to instantly determine the belongings on a person (for instance when this person enters a shop).

3.6 Citizenship, justice

3.6.1 Description

Citizenship and justice related attributes are for information that is related to the interaction of the person from the legal perspective. They are used to represent the *person's information* related to the rights and obligations of the person towards society, as well as the expression of this person in a democracy.

Of course this information is very sensitive, since the level of its exploitation is strongly related to the level of democracy within a society and the right of expression.

3.6.2 Examples of attributes

- Authority (representative of justice, journalists, etc.)
- Political affiliation

- Union affiliation
- Criminal records
- Actions (liabilities)

3.6.3 Application domains

The application domains include the government, police and justice.

3.6.4 Relevant standards

The Global Justice Extensible Markup Language Data Model (Global JXDM) and Dictionary (Global JXDM) are the result of an effort by the justice and public safety community to produce a set of common, well-defined data elements to be used for data transmissions.

For instance, it can be used for crime reporting, or to define an arrest warrant, and be exploited in the Information Systems used by policemen (for instance computer systems present in police cars).

3.7 Preferences

3.7.1 Description

Many applications including IMS support preferences as expressed by the user. These preferences are mainly related to the interface e.g. font type, colour of the display, use of toolbars etc. Other preferences could be privacy preferences, for example, which are supported by a number of web browsers[11].

3.7.2 Examples of attributes

- Level of detail of the interface (simple or advanced)
- Desired layout (colours, etc.)
- Level of disclosure of personal information

3.7.3 Application domains

Preferences are used, for instance, in web and mobile applications, but they are also available for other applications including IMS of various types. Privacy preferences are implemented particularly in web browsers, allowing the comparison of the local policy with the policy of websites supporting the same standard, such as P3P[12].

---

[11] See e.g. http://www.datenschutzzentrum.de/p3p/

[12] See http://www.w3.org/P3P/

3.7.4 An example of preference in ubiquitous and mobile services

The progressing evolution of technology allows the telecommunications and information services to provide omnipresent services and applications, which facilitate the users' everyday activities. Ubiquitous services must overcome the limitations of time and space: must be accessible anytime and from any location. It is envisaged that ubiquitous services will become one of the main pillars that will support future telecommunications services.

Ubiquitous services have the following characteristics:

- Services are provided, or often triggered, based on physical conditions
- Since services are invoked when such conditions are satisfied, services are aware of the real-world status and users situations
- Services are offered without an explicit request from the user, because they are triggered automatically by the system and not by the user's intention

One of the main challenges of ubiquitous services is the capture (discovery and/or acquisition) and communication of profile data, commonly known by Capability and Preference Information (CPI), in order to provide an adaptive response. Hence, the presentation of information must be adaptable to different users and user platforms, categorised by the following features:

- **Devices**
  (PC, PDA, WAP Phone, WebTV, etc.)

- **Device capabilities**
  (display size, memory size, network speed, etc.)

- **User preferences**
  (desired layout, navigation patterns, etc.)

- **User browsing history**

In this context, some standards exist such as CC/PP (Composite Capabilities / Preferences Profile) proposed by W3C[13] and UAProf (User Agent Profil) proposed by the WAP (Wireless Access Protocol) forum.

See the Annexe for a description of CC/PP UAProf.

3.7.5 Relevant standards

Relatively few specifications appear to exist related to the representation of users' preferences, although in the domain of Mobile applications, CC/PP (Composite Capabilities / Preferences Profile) is utilised.

---

[13] www.w3c.org

For the comparison of a local privacy policy defined within an appropriate web browser with the privacy policy of a visited web site, P3P[12] can be used. P3P is a standard provided by the W3C consortium.

3.8 Competences

3.8.1 Description

Competency relates to all the capabilities of the person. It includes both the official competences that have been certified by a certifying institution (e.g. school, university, governmental bodies, etc.) such as a diploma, the competences that are testified by a certain amount of experiences in an area (job, life, etc.), and competences that are more implicit (for instance deriving from the interest of the person to a particular topic).

3.8.2 Examples of attributes

- Certified competences
  - Driving licences
  - Diplomas
- Skills and know-how
- Experiences
  - Job
  - Life

3.8.3 Application domains

The two main application domains that deal with competency information are education and work. In the first case, educational systems (e-learning systems) need to track and record the evolution of competence of the students. In particular, LMS (Learning Management Systems) usually maintain a model of the student that includes the different abilities of the students, as well as the different grades that these students obtained in different disciplines. More intelligent e-learning systems have also tried to exploit this information in order to provide more customised interaction (note: the more advanced systems not only try to use competency or preference information, but also cognitive characteristics such as learning style).

In the second case, Enterprise management systems (and in particular Human Management Systems, or Knowledge Management Systems) are interested in managing people competency information for different companies' processes, such as recruitment (some competences may be desirable or even required to be certified by a diploma), career management, and the management of the intellectual capital of the organisation (used for instance in the cartography of competence of organisations).

Competence may also be managed in other application domains. For instance, a person may have to justify a driving license in order to drive a car, a bus or a truck.

3.8.4 Relevant standards

The domain of e-learning has defined with IMS/LIP (Learner Information Package) a series of specification for the representation of the learner information. Two different slots are used to represent competency, the 'Competency' slot, which is used to represent skills, knowledge, and abilities acquired in the cognitive, affective, and/or psychomotor domains and the 'QCL' slot (Certifications and Licenses) which is used to represent qualifications, certifications and licenses granted by recognised authorities.

HR-XML defines a schema for specifying the competency (measurable characteristics) of the employee. The competencies schema allows the capture of information regarding the evidence used to substantiate a competency and ratings and weights that can be used to rank, compare, and otherwise evaluate the sufficiency or desirability of a competency.

3.9 Functions, position and roles

3.9.1 Description

Function, position and role relates to the formal or informal "position and role" that the person is fulfilling as member within an organisation such as an enterprise or governmental service, and to the tasks in which this person is engaged.

3.9.2 Examples of attributes

- Job title (CEO, Director, …)
- Title (Dr., Prof., Her Majesty, …)
- Tasks
- Working activity state. (instantaneous information)
- Informal role (and authority), castes (see social characteristics)

3.9.3 Application domains

Work appears the most obvious domain in which functions, position and roles intervene. However, many other domains are also concerned by this facet such as Justice (role related to authority), commerce (defining responsibilities) and even leisure (for instance in a Massively Multiplayer Online Game).

3.9.4 Relevant standards

For formal roles, HR-XML defines for instance the slot JobAndPosition. The Job and Position Header specifications are high-level entities that may be used within a variety of human resource management (HRM) models and business processes. Both entities may be categorised into fragments such as Duties and Responsibilities, Work Policy, Requirements, and Work Schedule.

Global Justice Extensible Markup (JXDM) defines different types of persons such as: EnforcementOfficial, JudicialOfficial, CaseOfficial, Juror, MissingPerson,

RegisteredOffender, Subject, Witness. Each of them share some attributes, but also includes a set of specific attributes.

Finally, as described in the next section, the role can also be informal. For instance a person in an organisation or in a group can have an authority over other persons (such as leadership) that may be implicit and not appear in the formalised structure of the organisation.

3.10 Social characteristics (sociology)

3.10.1   Description

The "social characteristics" deals with all the information regarding the relationship that people have with one another individually, or as part of a group.

3.10.2   Examples of attributes

- Social networks (FOAF)
  - o   Personal network (mates, friends, ...)
  - o   Family network
  - o   Professional network
  - o   Political network
- Affiliation
  - o   Groups
  - o   Professional groups
  - o   Political party
- Social categories (extracted via data mining)
- Social identity (reputation, informal roles & authority, etc.)
- Social style (competitive / collaborative, introvert / extravert)
- Social role
  - o   Team role (Belbin model)
    Plant, Resource investigator, Co-ordinator, Shaper, Monitor evaluator, Teamworker, Implementer, Completer
  - o   Reputation

3.10.3   Application domains

Many application domains deal with social characteristics. For instance, on the Internet, online social networking systems (such as LinkedIn, Friendster, Orkut) directly represent the social networks in order to facilitate the creation and the exploitation of social structure. Similarly, the Blogosphere (the Internet space that consists of a set of blogs) is often considered as a giant social network linking the different owners of blogs which refer to one another. Affiliation is also an attribute that is present in many systems relying on informal communication such as virtual communities.

In the domain of electronic commerce, social characteristics can appear via the socio-economical group categories that can be extracted using data-mining techniques.

3.10.4 Relevant standards

Standards such as FOAF (Friend Of A Friend) or XFN represent specifications that have been created specifically to represent people relationships.

In the domain of eCommerce, CIQ (OASIS Customer Information Quality) also provides a way to specify relationships between a vendor and a customer, or a customer and a vendor organisation.

3.11 Psychological characteristics (psychology)

3.11.1 Description

The "psychological characteristics" deals with the behavioural aspects of the person. It includes both the relatively permanent behavioural characteristics of the person such as personality or cognitive style, to the most instant behavioural state of the person such as the current psychological state such as mood or attention.

3.11.2 Examples of attributes

- Psychological state
  - Mood
  - Attention
- Personality (5 factor model)
  - Extroversion (Sociable / Reserved)
  - Conscientiousness (Self-disciplined / Impulsive)
  - Emotional Stability (Self-Confident / Insecure)
  - Agreeableness (Sympathetic / Cold)
  - Openness to Experience (Curious / Unimaginative)
- Motivation (Steven Reiss 16 basic factors model)
  - Avoiding Pain & Anxiety
  - Citizenship
  - Curiosity
  - Family
  - …
- Cognitive style
  - Learning style
    - Transforming Learners, Performing Learners, Conforming Learners, Resistant Learners
  - Innovation style
    - Innovators, Early Adopters, Early Majority, Late Majority, Laggards
  - Social style
    - Connector, Maven, Salesman

3.11.3 Application domains

Behavioural characteristics are a very sensitive domain, intimately related to the person. Several domains however are concerned in one way or another with these behavioural characteristics such as HR (for instance in the hiring process with personality tests or the

Future of Identity in the Information Society (No. 507512)

career management of the employees), e-learning (several researches are trying to take into account the cognitive style of the learner), justice (criminal profiling) or e-commerce (categories of customers).

3.11.4   Relevant standards

No open standard is currently available on this subject.

3.12 Summary

The following list represents a (very limited) set of attributes indicating the large variety of facets that can be used to represent the person. Its aim is simply to indicate the variety of characteristics that can be used to represent the person.

Facet	Examples of Attributes	Some domains of use
Identification	Name ssn (social security number) Pseudonym	Security, commerce
Biological characteristics	Gender Eye colour Height DNA Finger print Retina Iris Face Gesture Medical information	biometrics, medical, forensic
Location	Home location Work location Instant location (GPS/mobile)	Mobility, e-commerce
Assets & transactions	Assets, (real estate, etc.) Liabilities, Financial information Tax information Income Transactions	E-Taxation, e-commerce, fighting crime
Citizenship, justice	Political opinion Union affiliation Criminal records attributes Actions (liabilities)	Law enforcement, police, democracy, society
Preferences	Interests	Organisation, education, commerce, mobility, leisure
Competences	Diploma Expertise	Organisation, education

Functions, position and roles	Organisation Title Role Tasks Responsibilities	Organisation, Law
Social characteristics	Personal network (mates, friends, ...) Affiliations Reputation	Society, personal life
Psychological characteristics	Psychological state Personality, cognitive style Motivation Sexual orientation Behaviours	Human resources, education, criminal profiling

4 The Representation of the Person in Different Application Contexts, and the Use of Standards

In the previous section, we looked at the different categories of attributes that can be used to represent the person.

In this section, we are going to see how these attributes are being assembled to represent the person for an application in different contexts. We will also have a look at some of the standards of representation of the "identity" of the person. These standards rely on well specified sets of attributes that are used to represent the "identity" of the person in "Identity informed" Information Systems.

The objective of this chapter is not to redefine again the identity concept (this has been largely covered in the deliverable del.2.1 of the FIDIS project), nor to make a complete inventory of the technical standards (which are tasks that belongs more to other workpackages such as "WP4 Interoperability" or "WP3 Technologies" of this project), but to provide a more concrete perspective of identity based on how the identity is represented and use in different application domains, and represented by standards.

4.1 Identity in Different Application Contexts

The objective of this section consists mainly in providing the application context perspective of the work that was done in chapter 3. More concretely, it consists in identifying application domains in which identity represents a key role, and for each of them in indicating the set of attributes that are relevant for the representation of the person. It also mentions the relevant standards that can be used (although these standards will be presented more specifically later).

This section provides us with a basic comparison of the approaches that are being used to represent the person in the different domains, and in particular what are the aspects in which they differ or on the contrary are very similar.

In the following table, we provide an overview of key aspects of identity in different application contexts, as well as some relevant standards. In some cases (for instance for Work and "Commerce") we have distinguished the organisational perspective from the individual perspective is order to reflect different way the identity is managed.

Future of Identity in the Information Society (No. 507512)

Domain	Some Key aspects of identity (attributes)	Indicative Standards
Mobile applications (mobile phone, ambiant intelligence)	Name, address, location, preferences,	CC/PP, UAProf,
Education (e-learning, learning networks)	Name, address, competence, Social networks	IMS/LIP; LDAP (Educause)
Human Resources	Name, address, competence, psychology, sociology, social networks, roles	HR-XML, FOAF
Work (Organisations)	Name, address, location, security, roles	LDAP, PKCS, etc.
Work (knowledge worker)	Name, pseudonym, social networks, competence,	FOAF
Justice & police (forensic, money laundering, law enforcement)	Name, SSN, address, location, biometrics,	JXDM
Government (taxation, democracy/voting)	Name, address, assets & liabilities, …	
Commerce (companies)	Customer information, reputation	CIQ
Commerce (customer)	Name, address, preferences, banking information,	
Banking, insurance	Name, address, banking information, Assets & liabilities,	
Personal sphere (Leisure, family)	Name, pseudonym, address, preferences, social networks, psychology	FOAF
Health	Name, SSN, address, Biological information	

This table represents a very preliminary work, that should need to be consolidated in the future (using a much more rigorous methodology of classification).

It indicates the important overlapping of the identity from a domain to another (see the Del 2.1 for a more theoretical perspective, as well as the possible implications for the privacy of the person), but also the specificity of each application.

In some cases, there exists some strong similarity in the way an aspect of the identity is managed. For instance the name of the person or her address is an element that will be found in many different application domains, and will be managed in a very similar way. Other domains may have a strong overlapping in the way they represent a facet of the identity, such as for instance the management of competence in Human Resource Management or Education.

In some other cases, certain aspects of identity can be considered as very specific to the application domain. Sets of ad hoc and much elaborated attributes are then used to represent and capture the complexity of a particular element of the identity. For instance, in the case of mobile application (mobile phone, ambient intelligence), the location of the person represents an important dimension, and can therefore be represented with a high level of precision (this will include for example the precise and instant location of the person). In the case of commerce application, the address of the person can be represented with a high level of precision in CIQ/xNAL (information which is important for direct marketing communications).

4.2 A Overview of Standards for the Representation of the Person

The details of the different standards that are referenced in this section are provided in "6 Annexes: An Overview of Standards for Representing *Person's Information*".

The objective of this overview is not to make a complete inventory of these standards (such task would rather belong to a workpackage such as "WP4 Interoperability" or "WP3 Technologies"), but to contribute to a better understand the Identity concept by analysing, thought different standards, contexts in which the identity is represented and used.

Indeed standards of representation of the persons consist in a well defined specification of a set of attributes that can be used to represent the identity of a person in a particular application domain or in general purpose Identity Management Systems.

Note: the standards for the exchange of identity information (.i.e. protocols) are considered to be outside the scope of this document since they are not considered to be central to the comprehension of the structure of the Identity. Processes and protocols will more specifically be addressed in "WP4 Interoperability" and in "WP3 Technologies" (Identity Management systems).

We can distinguish two main categories of standards for representing person's identity:

- General Identity representation standards.
- Application specific standards.

General Identity representation standards of representation of the person have been defined mostly independently from any application domain, in order to specify the representation of a particular identity aspect. Usually led by a technical standardisation group such as ISO, IETF, ANSI, Liberty Alliance or W3C, they represent some of the elements participating in the construction of general purpose identity management systems. Some of the most important elements of the identity that are represented in these standards are the identifiers that can be associated to the persons (such as named, Ids, signatures), and other directory information (addresses, title, etc.). Typical standards include directory schema (LDAP, Liberty Alliance), but we can observe an increasing importance in the scope of these standards (for instance Liberty Alliance ID-SIS-GL cover instant location, and general standards such as XCBF are being elaborated in order to represent biometric information).

Application specific standards for the representation of the person have been defined specifically to address the needs of a particular application domain (such as Justice, Human Resource Management, Education, Commerce, etc.). The standards of the representation of the person usually represent a subpart of the more general standards that are used to model any categories of information in this domain. Application specific standards are usually defined by organisations that are important stakeholder in the domain (such as IMS Global in the domain of e-Learning or the U.S. Department of Justice (DoJ) for JXDM), and sometime coordinated by business oriented standardisation organisations (and more particularly OASIS). Typical standards include IMS Global / LIP (learner information package) for the domain of e-learning, HR-XML used to design human resource management systems (an important facet of this standard is dedicated to the modelling of the employee) or JXDM (Justice Markup language) that is used by the Us Department of Justice to describe in particular the person information related Justice.

Future of Identity in the Information Society (No. 507512)

These two categories of standards can be considered as very complementary and connected, and the application specific standards are expected to increasingly rely on the general identity representation standards as those later ones are extending their scope of application.

The following table list many of these different standards (See "6 Annexes: An Overview of Standards for Representing *Person's Information*" for more detailed description of these standards).

Standards			
Standard		Chapter in Annex 6	details
CC/PP	Composite Capability/Preference Profiles	6.5.1	Used in mobile applications
CIQ	Customer Information Quality	6.6	Customer information data model
FOAF	Friend Of a Friend	6.7.1	Social network
FTML	Family Tree Markup Language	6.7.3	Genealogy Web Format
HR-XML	Human Resource XML	6.3	XML for human resource management systems.
HumanML	Human Markup language	6.8	
IMS/LIP	IMS Global / Learner Information Packaging	6.2	Student representation in e-learning systems
JXDM	Global Justice Extensible Markup	6.4	Justice
LDAP schemas	Lightweight Directory Access Protocol	6.1.1	directory
Liberty Alliance ID-SIS-CB, ID-SIS -GL	Liberty Alliance Identity Service Interface Specifications -CB (Contact book); -GL (Geo-Location)	6.5.3	Identity management
PKCS #9	Public-Key Cryptography Standards #9	6.1.4	authentication
UAProf	User Agent PROFile	6.5.2	Used in mobile applications
vCard	Virtual Card ?	6.1.2	vCard is the electronic business card
XCBF	XML Common Biometric Format	6.5.4	biometric (OASIS)
XFN	XHTML Friends Network	6.7.2	XFN is a simple way to represent human relationships using hyperlinks.
XRI and I-Name	Extensible Resource Identifier	6.1.3	identifiers
xNAL	Extended Name and Address Standard	6.6.1.1	Names and addresses. (part of CIQ)

5 Summary and Future Work

5.1 Summary

In the previous section, the different facets of representation of a person have been examined, and for each of them the application domains that are the most likely to utilise each of these facets has been identified. To date and to our knowledge, a similar work had not been conducted on a large scale, although some tentative attempts, specific to some domains (in the domain of e-learning, for instance, Diogene (2002) tries to identify other user models in different domains in order to provide guidance in the definition of the student model), are worthy of mentioning.

This document has revealed some notable findings. Firstly, it has been observed that several application domains (human resource, e-learning, justice) have already started to create quite complete and holistic models of the person that fulfil the needs of this domain (see the example of HR-XML, IMS/LIP, or JXDM in the annex). In some cases, domains have independently developed different representations (in some cases very elaborated) for representing the same objective (such as for the representation of the address of a person), raising the question of the duplication of work.

In other cases, attempts to reuse some of the existing specifications have been observed. For instance there is a clear similarity between the specifications developed for e-learning in IMS/LIP and in Human resources with HR-XML that indicates some clear interdependence between the two specifications. Also, the Liberty ID-SIS Service Specification explicitly mentions the use of another specification (vCard) as part of the specification.

Currently, the modelling of identities appears to be composed of dispersed initiatives, is still very much under construction (and incomplete), and very complex for the most advanced initiatives. Such initiatives would probably benefit from additional cross-disciplinary studies, such as this, even if the benefit of a holistic approach still needs to be demonstrated (because of its complexity). It has been noted in particular that there is relatively limited interaction between different "Universes" (such as identity management and intelligent adaptive systems) for which representation of persons is very important.

Finally, and not surprisingly, it can be noted that some aspects of the person modelling, such as the behavioural or social characteristics, are still largely underdeveloped, especially considering that they will be essential in the new technologies of tomorrow, such as intelligent mobile and ubiquitous applications.

5.2 Future work

In Chapter 4, we provided a first preliminary application perspective of identity representation. For instance, we have mentioned the work of different application domains such as human resources (HR-XML), Education (IMS/LIP - Learner information package), government, police and justice (JXDM), healthcare, commerce (CIQ - Customer Information

Quality), banking, business (business card), mobility or leisure. We have also found that some of the application domains such as human resources (with HR-XML) and justice (with JXDM) are already pretty advanced in the elaboration of relatively holistic models, although it remains to be seen what the impact is on the complexity of the systems that will have to implement them.

It is clear that a more in-dept and systematic work of identifying the attributes of identity in different application domains, and an analysis of the commonalities between the different domains would be useful.

Finally, other routes that deserve studies in their own right are related to:

1. The control and the distribution of the personal information.

People may not want in the future to see their personal information managed within a single centralised system. Federated identity management systems such as the Liberty Alliance may have to deal with *several standards* of representation of *person's information*.

2. The management of multiple identities.

The grouping of attributes may have to be considered more dynamic in order to address a larger diversity of contexts (an application may for instance have to deal with several identity representations varying over time)

3. The legal and ethical issues associated to the representation, the control and the use of person's information.

A distinction should indeed be formed between the characteristics of the person that can technically be represented and captured (this limit is continuously pushed forward with progress in the technology and in knowledge modelling), and the one's that should be represented and captured - society has to determine limits on information that can be represented and processed in Information Systems.

References

Andre, E., Klesen, M., Gebhard, O., Rist, T.,. (2000); "'Exploiting Models of Personality and Emotions to Control the Behavior of Animated Interactive Agents'"; Fourth International Conference on Autonomous Agents, pp. 3-7, Barcelona.

Baecker, D., (1999); "Organisation als System"; Suhrkamp, Frankfurt am Main 1999.

Brusilovsky, P., (2001); "Adaptive Hypermedia, User Modeling and User-Adapted Interaction"; Kluwer Academic Publishers, Printed in the Netherlands, pp. 87-110

Diogene (2002); "Survey on Methods and Standards for Student Modelling"; Diogene project, September 2002. http://www.diogene.org/archive.html

Fink, J., Kobsa, A., (2000); "A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web"; in User Modeling and User Adapted Interaction, Special Issue on Deployed User Modeling, 10, p.204-209 [Fischer, 2001]

Fischer G. (2001); "User Modeling in Human-Computer Interaction"; in User Modeling and User Adaptive Interaction, Kluwer Academic Publishers, Printed in Netherlands, 2001. p.69-85

Kay, J, (2000); "User modeling for adaptation"; in User Interfaces for All; Stephanidis (ed), C, Salvendy, G, (General Editor), Human Factors Series, Lawrence Erlbaum Associates, 271—294.

Kesterling, A., (2000); "Kommunikation unter Anwesenden – Studien über Interaktionssysteme"; Frankfurt am Main.

Kobsa, A., Koenemann, J., and Pohl, W., (2000); "Personalized hypermedia presentation techniques for improving online customer relationships"; The Knowledge Engineering Review 16, p111-155

Luhmann, N., (1991); "Die Form 'Person'"; in Soziale Welt; 42 (2), p. 167-175.

Luhmann, N., (1997); "Die Gesellschaft der Gesellschaft"; Suhrkamp, Frankfurt am Main.

Luhmann, N., (2000); „Organisation und Entscheidung"; 1st Edition; Westdeutscher Verlag, Opladen/Wiesbaden.

Mountain, D., and Raper, J., (2000); "Designing geolocation services for next generation mobile phone systems"; the AGI (association for geographic information) conference at GIS 2000.

Nabeth, T., Angehrn, A.A. and R. Balakrishnan (2004); "Integrating 'Context' in e-Learning Systems Design"; IEEE International Conference on Advanced Learning Technologies (ICALT 2004), Joensuu, Finland

Razmerita, L., 2004; "User modeling and personalization of the Knowledge Management Systems"; book chapter in Adaptable and Adaptive Hypermedia, edited by Sherry Chen and George Magoulas, published by Idea Group Publishing.

Stephanidis, C., (2001); "Adaptive Techniques for Universal Access"; User Modeling and User-Adapted Interaction 11: 159-179, Kluwer Academic Publishers

See also:

The Fidis Wiki (restricted access).

http://internal.fidis.net/178.0.html?tx_a1wiki_pi1[keyword]=t2.1%20definition

The FIDIS WP2 collection of identity resources at:

http://www.calt.insead.edu/fidis/information/news/

6 Annexes: An Overview of Standards for Representing *Person's Information*

This Annexe provides some brief descriptions of some of the identity schema standards for representing user information in different domains. They are aimed at giving the reader some global overview on the categories of attributes manipulated in the different standards, without this having reader to go into the full specifications. The full specifications, to which this document provides references, should be used in the case the reader would like to engage into elaborated work.

6.1 Directories and Business Cards (LDAP, vCards, Active Directory, Liberty Alliance schemas, PKCS, etc.)

Directory services provide the means to manage the identities and relationships that make up network environments. Directory services can for instance be used to manage the email addresses, but also to manage the permission and the access to resources. The objects that are stored in directories can include people, but also organisational structures and resources.

6.1.1 LDAP schema

6.1.1.1 Presentation

The Lightweight Directory Access Protocol (LDAP) is a protocol for accessing directory services, which can be seen as specialised databases, e.g. over the internet. It replaces its predecessor, the Directory Access Protocol (DAP), which has not become widely accepted due to its complexity. Version v3 of the LDAP is specified in RFC3377. A main feature of the LDAP is its focus on security aspects, it provides AAA (authentication, authorisation and accounting) functionality to secure the information within directory services by supporting encryption between the client and the directory server, and by providing access control lists (ACLs) defining who may access specific directory entries or parts thereof.

An LDAP directory contains entries in a hierarchical structure. A directory entry may refer to any kind of entity, object, or resource. Each entry is identified by a distinguished name (dn), and consists of one or more attributes, i.e. type-value pairs. A directory schema specifies a set of rules defining what may be stored in a directory, the valid attribute types etc.

Table 1 shows an example of a LDAP directory entry representing a person. The attribute types used in the example are commonly used and therefore given as short keyword names.


The Distinguished Name of the entry, containing in this case the person's Common Name (cn), Organisation (o) and Country (c)	`cn=Richard Cissee,`
	`o=DAI-Labor,`
	`c=Germany`

Email address attribute.	`mail=richard.cissee@dai-labor.de`
LocalityName attribute, typically the city.	`l=Berlin`

Table 1: An Exemplary LDAP Directory Entry

6.1.1.2 Presentation: application to higher education

The EDUCAUSE/Internet2 eduPerson task force has the mission of defining an LDAP object class that includes widely-used person attributes in higher education.

In particular EDUCAUSE has defined an LDAP schema for representing the person in an educational environment.

top
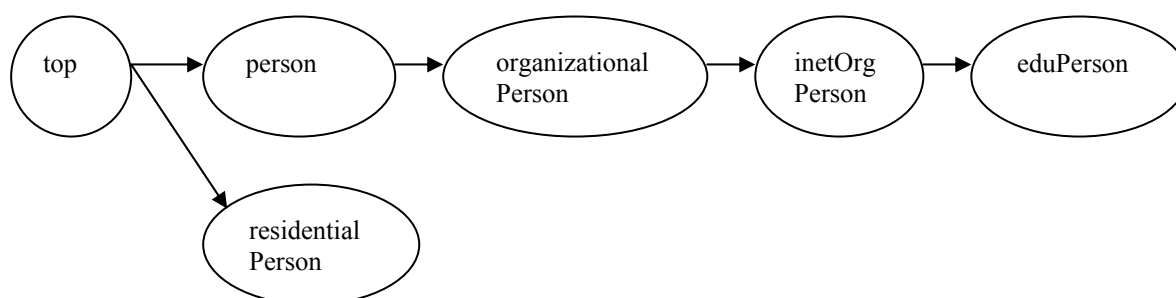
- **Person**
    - o   sn TDIRNAME_NAME (mandatory)
    - o   cn TDIRNAME_NAME (mandatory)
    - o   description (string)
    - o   seeAlso (string)
    - o   telephoneNumber TDIRTELE_TELEPHONE
    - o   userPassword (string)
    - o   **organizationalPerson**
        - ▪   title PRSN_WORKING_TITLE
        - ▪   ou TDIRORGN_ORGAN
        - ▪   preferredDeliveryMethod (string)
        - ▪   st ADDR_STATE_CD
        - ▪   telexNumber (string)
        - ▪   l unknown
        - ▪   telephoneNumber TDIRTELE_TELEPHONE
        - ▪   physicalDeliveryOfficeName (string)
        - ▪   postalCode ADDR_ZIP_CD
        - ▪   internationalISDNNumber (string)
        - ▪   x121Address (string)
        - ▪   registeredAddress TDIRADDR_ADDRESS fields from appropriate type record.
        - ▪   street TDIRADDR_ADDRESS
        - ▪   postalAddress TDIRADDR_ADDRESS
        - ▪   facsimileTelephoneNumber TDIRTELE_TELEPHONE
        - ▪   teletexTerminalIdentifier (string)
        - ▪   postOfficeBox ADDR_...
        - ▪   destinationIndicator (string)

- **inetOrgPerson**
  - userCertificate DGCR_DIGITAL_CR from appropriate type
  - uid TDIRIDEN_IDENTFIER
  - homePostalAddress TDIRADDR_ADDRESS
  - employeeType unknown
  - preferredLanguage PRSN_PREF_LANG
  - mail PRSN_UF_EMAIL_AD/ORGN_UF_EMAIL_AD
  - homePhone TDIRTELE_TELEPHONE
  - roomNumber (string)
  - x500UniqueIdentifier TDIRDGCR_DIGITALCR
  - employeeNumber PRSN_UFID
  - photo (string)
  - businessCategory (string)
  - pager TDIRTELE_TELEPHONE
  - o ORGN_DISPLAY_NM
  - jpegPhoto (string)
  - secretary unknown
  - audio (string)
  - userPKCS12 DGCR_DIGITAL_CR from appropriate type
  - displayName ORGN_DISPLAY_NM/PRSN_DISPLAY_NM
  - mobile TDIRTELE_TELEPHONE
  - labeledURI unknown
  - carLicense (string)
  - givenName TDIRNAME_NAME
  - manager (string)
  - userSMIMECertificate DGCR_DIGITAL_CR from appropriate type
  - initials TDIRNAME_NAME inetOrgPerson
  - departmentNumber ORGN_UFID
  - **eduPerson**
    - eduPersonOrgUnitDN relationships
    - eduPersonOrgDN relationships
    - eduPersonPrincipalName CMAC_USERID of GatorLink type, + @UFL.EDU
    - eduPersonNickname TDIRNAME_NAME
    - eduPersonAffiliation relationships
    - eduPersonPrimaryAffiliation PRSN_PRI_AFF_TYPE
- **residentialPerson**
  - l (string) (mandatory)
  - preferredDeliveryMethod NONE
  - st ADDR_STATE_CD
  - businessCategory (string)
  - telexNumber (string)
  - telephoneNumber TDIRTELE_TELEPHONE
  - physicalDeliveryOfficeName (string)
  - postalCode ADDR_ZIP_CD
  - internationalISDNNumber (string)
  - x121Address (string)
  - registeredAddress TDIRADDR_ADDRESS fields from appropriate type record.

- street TDIRADDR_ADDRESS
- postalAddress TDIRADDR_ADDRESS
- facsimileTelephoneNumber TDIRTELE_TELEPHONE
- teletexTerminalIdentifier (string)
- postOfficeBox ADDR_... residentialPerson
- destinationIndicator (string)

Example 1: An simple LDAP file with two entries

```
version: 1
dn: cn=Barbara Jensen, ou=Product Development, dc=airius, dc=com
objectclass: top
objectclass: person
objectclass: organizationalPerson
cn: Barbara Jensen
cn: Barbara J Jensen
cn: Babs Jensen
sn: Jensen
uid: bjensen
telephonenumber: +1 408 555 1212
description: A big sailing fan.


dn: cn=Bjorn Jensen, ou=Accounting, dc=airius, dc=com
objectclass: top
objectclass: person
objectclass: organizationalPerson
cn: Bjorn Jensen
sn: Jensen
telephonenumber: +1 408 555 1212
```

6.1.1.3   References

RFC 2849 - The LDAP Data Interchange Format (LDIF) - Technical Specification, June 2000

http://www.faqs.org/rfcs/rfc2849.html

6.1.2 vCards

The vCard standard defines a format for electronic business cards intended for Personal Data Interchange (PDI), i.e. exchanging personal information between individuals. The current version 3.0 of the vCard specification, which has been developed by the Internet Mail Consortium, is defined in RFCs 2425 and 2426. The vCard specification is operating system independent; all information is stored in a file (usually recognisable by its extension ".vcf")

that is mainly text-based. VCards are supported by and used in various applications and contexts, such as Web browsers, email clients and Personal Information Managers.

A vCard may contain the following types of information:

- Personal information, such as identification information (the name, nickname, and birthday of a person), addressing information (a person's home, business and email addresses, various telephone numbers), and organisational information (the organisation a person is associated with, his title and role within the organisation).

- Additional geographic information which may be useful when contacting the respective person, such as the time zone or geographical co-ordinates of the address where the person can usually be reached. This is usually static and does not indicate a person's current location.

- Additional multimedia information, such as a photography of the person or a sound file containing the correct pronunciation of the person's name.

Table 1 shows an example of a vCard containing only text-based elements.

	`BEGIN:VCARD`
The version of the VCard specification.	`VERSION:2.1`
The name of the person (family name; given name; additional names; pre- and suffixes). Required entry.	`N:Cissée;Richard`
The formatted name. Required entry.	`FN:Richard Cissée`
The organisational name and units associated with the person.	`ORG:DAI-Labor, TU Berlin`
The person's position or job title.	`TITLE:Research Assistant`
Various telephone numbers (possible types include voice, fax, mobile, business, home).	`TEL;WORK;VOICE:030 /31473612` `TEL;WORK;FAX:030 /31421799`
Address information (again, different types of addresses are possible here).	`ADR;WORK;ENCODING=QUOTED-PRINTABLE:;;Technische Universit=E4t Berlin / DAI-Labor=0D=0ASekretariat GOR1-1=0D=0AF=ranklinstrasse 28/29;Berlin;;10587;Deutschland`
The URL associated with the person.	`URL;WORK:http://www.dai-labor.de`
The person's email addresses.	`EMAIL;PREF;INTERNET:richard.cissee@dai-labor.de`
The revision of the information contained in the vCard (a timestamp indicating the last modification).	`REV:20050411T125131Z`
	`END:VCARD`

Tableau 1: An Exemplary VCard

6.1.3 Universal private address with XRI and I-Names

I-names are universal private addresses intended to provide persistent addressing. An I-name integrates a person's different electronic addresses, such as email, instant messaging,

Future of Identity in the Information Society (No. 507512)

SMS/MMS etc. An I-name assigned to a person is, similarly to a domain name, supposed to be relatively long-lived and does not change even when the person's contact data changes. Additionally, it protects the person from undesired communication attempts, such as spam, by operating on a permission-based communication infrastructure.

I-names are based on the XRI (Extensible Resource Identifier) and XDI (XRI Data Interchange) specifications under development from OASIS. XRIs are a kind of Uniform Resource Identifier (URI) providing a standard syntax and resolution protocol for abstract identifiers, i.e. identifiers that are independent of a specific location, domain, application, or protocol. I-names are human-friendly identifiers (for a person, they may simply have the form "=GivenName.FamilyName") that are usually combined with I-numbers, i.e. persistent machine-friendly identifiers assigned to resources (persons in this case).

An I-name may be used to contact the respective person via a specific channel, such as email, only if the person has given permission to do so. Otherwise, the I-name may only be used to request this permission. These requests can automatically be filtered by an I-broker using a personal contact page via which the person may specify rules for contacting.

6.1.3.1 References

I-Names Explained http://www.xdi.org/i-names-explained.html

6.1.4 The Public-Key Cryptography Standards #9 (PKCS #9)

In the domain of authentication, PKCS defines some specification providing some definition of the person.

The naturalPerson object class is a general-purpose auxiliary object class that is intended to hold attributes about human beings. It has been designed to be used within directory services based on the LDAP protocol and the X.500 family of protocols, where support for these attributes is considered useful.

NaturalPerson

- emailAddress
- unstructuredName
- unstructuredAddress
- dateOfBirth
- placeOfBirth
- gender
- countryOfCitizenship
- countryOfResidence
- pseudonym
- serialNumber
- ... -- For future extensions

References

RSA Laboratories (2000); "PKCS #9 v2.0: Selected Object Classes and Attribute Types" RSA Laboratories, February 25, 2000; http://www.rsasecurity.com/rsalabs/node.asp?id=2131

6.2 IMS/LIP (Learner Information Packaging)

E-learning represents a domain that has been relatively far in trying to create standards for addressing the interoperability of internet based facilitated e-learning systems (typically Learning Management Systems). Examples of standards that have been designed include: (1) SCORM (Sharable Courseware Object Reference Model) an e-learning specification that is used to represent learning content in a standardised way, and to define how this content interacts with learning management systems; and (2) IMS LIP (Learner Information Package), a standard for representing *person's information* in e-Learning systems.

IMS Learner Information Package is based on a data model that describes those characteristics of a learner needed for the general purposes of:

- Recording and managing learning-related history, goals, and accomplishments;
- Engaging a learner in a learning experience;
- Discovering learning opportunities for learners.

The specification is aimed at supporting the exchange of learner information among learning management systems, human resource systems, student Information Systems, enterprise e-learning systems, knowledge management systems, resume repositories, and other systems used in the learning process.

LIP is structured around the following eleven core data structures:

1. **Accessibility** – Data regarding the accessibility of learner's information as defined through:

    − Language: the definition of a learner's language proficiencies.
    − Preference: the definition of a learner's cognitive, physical and technological preferences.

2. **Activity** – The activity the learner is engaging in, comprising:

    − Learning activity reference: an external reference mechanism to the learning materials.
    − Definition: the definition of the materials studied.
    − Product: the materials developed by the learners themselves.
    − Testimonial: statements attesting to the capabilities of the learner.
    − Evaluation: the results of the evaluations undertaken.

3. **Affiliation** – The learner's professional affiliations and associated roles.

4. **Competency** – The competencies of the learner.

5. **Goal** – The learner's goals and sub-goals.

6. **Identification** – The learner identification data. They comprise:

    − Formatted Name: the learner's name, formatted.
    − Name: the learner's name.
    − Address: the learner's addresses.
    − Contact info: electronic-based contact information about the learner.

> − Demographics: demographics information about the learner.
>
> − Agent: the representatives permitted to act on behalf of the learner.

7. **Interest** – Hobbies and recreational interests of the learner.

8. **Qcl** – A description of the qualifications, certifications and various licenses of a learner.

9. **Relationship** – the set of relationships that are to be defined between the learner and their identification, accessibility, qualifications, competencies, goals, activities, interests, transcripts, security keys and affiliations.

10. **Security key** – the security-related information for the given learner.

11. **Transcript** – the transcripts that summarise the performance of the learner.

6.2.1 Competency

Skills, knowledge, and abilities acquired in the cognitive, affective, and/or psychomotor domains.

This learner information consists of the descriptions of the skills the learner has acquired. These skills may be associated with some formal or informal training or work history (described in the 'activity') and formal awards (described in the 'qcl'). The corresponding level of competency may also be defined;
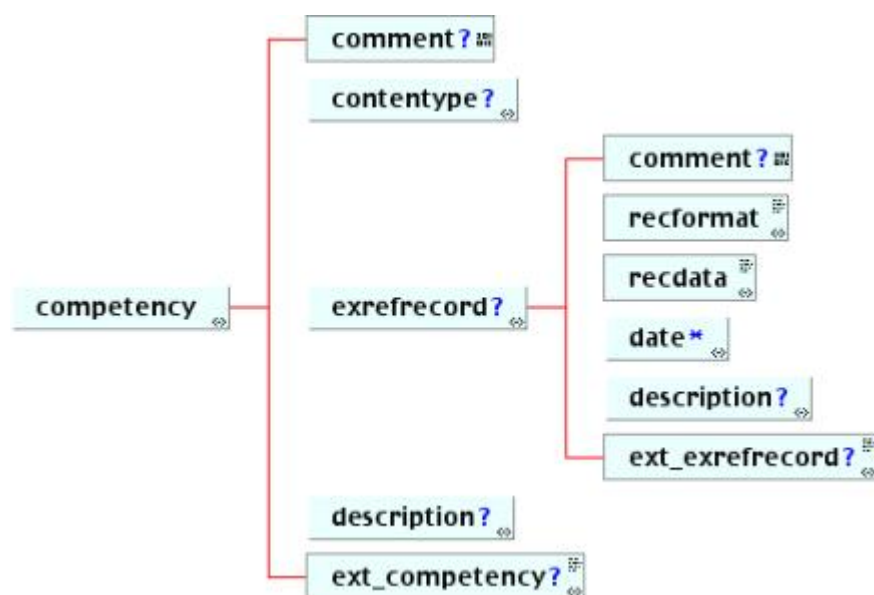
6.2.2 QCL

Qualifications, Certifications and Licenses (qcl): Qualifications, certifications and licenses granted by recognised authorities.

This learner information consists of the qualifications, certifications and licenses awarded to the learner i.e. the formally recognised products of their learning and work history. This includes information on the awarding body and may also include electronic copies of the actual documents.
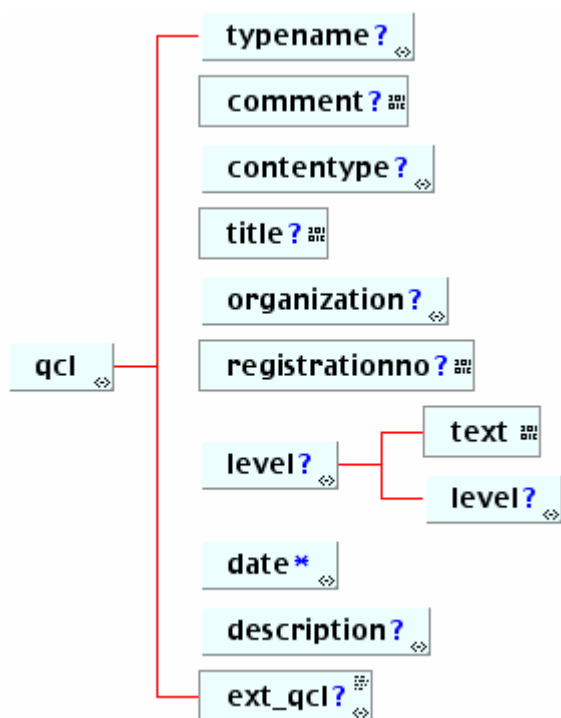
6.2.3 An example: Identification Address

This example creates the address for a learner. The example contains all of the IMS LIP information required to construct the instance record.

```
<learnerinformation>
        <comment>A basic example of an address in an Identification.</comment>
        <contentype>
                <referential>
                  <sourcedid>
                     <source>IMS_LIP_V1p0_Example</source>
                     <id>basic_6003</id>
                  </sourcedid>
                </referential>
        </contentype>
        <identification>
                <contentype>
                  <referential>
                     <indexid>identification_01</indexid>
                  </referential>
                </contentype>
                <address>
                  <typename>
                     <tysource sourcetype="imsdefault"/>
                     <tyvalue>Private</tyvalue>
                  </typename>
                  <contentype>
                     <referential>
                        <indexid>address_01</indexid>
                     </referential>
```

```
                    </contentype>
                    <street>
                       <nonfieldedstreetaddress>Towerview Apartment 3A, 34 Oxford St
                       </nonfieldedstreetaddress>
                       <complex>Towerview</complex>
                       <streetnumber>34</streetnumber>
                       <streetname>Oxford</streetname>
                       <streetype>Street</streetype>
                       <aptnumber>3</aptnumber>
                       <aptnumsuffix>A</aptnumsuffix>
                    </street>
                    <locality>West-end</locality>
                    <city>London</city>
                    <country>UK</country>
                    <postcode>SE23 2RR</postcode>
                    <timezone>GMT</timezone>
                    <geo>
                       <lat>57.01.49N</lat>
                       <lon>00.00.00E</lon>
                    </geo>
                 </address>
           </identification>
 </learnerinformation>
```

Note:

IMS LIP is based on XML, and is not intended for human reading. For other examples, see IMS Learner Information Packaging Best Practice & Implementation Guide Final Specification Version 1.0.

6.2.4 References

IMS Learner Information Packaging Information Model Specification; Final Specification, Version 1.0; IMS Global Learning Consortium, March 2001.
http://www.imsglobal.org/profiles/lipinfo01.html

For examples:

IMS Learner Information Packaging Best Practice & Implementation Guide Final Specification Version 1.0

http://www.imsglobal.org/profiles/lipbest01.html

6.3 HR-XML (xml for human resource management)

Human Resource represents another domain that has worked on the standardisation of the data that is manipulated in order to facilitate the design and the interoperability of Human Management systems.

Practically, the HR-XML consortium (http://www.hr-xml.org/) (a non-profit organisation) has produced a library of more than 75 interdependent XML schemas covering the major elements to be found in HR systems processes (such as Payroll, Healthcare, or Recruiting and Staffing).

6.3.1 Employees information in HR-XML

In HR-XML a variety of schemas that are used to represent *person's information* are defined such as:

PersonName Prescribes the form of the Person Name object

PostalAddress Prescribes the form of the Postal Address

JobAndPosition The Job and Position Header specifications are high-level entities that may be used within a variety of HRM models and business processes. Both entities may be categorised into fragments such as Duties and Responsibilities, Work Policy, Requirements, and Work Schedule.

Competencies (Measurable Characteristics) The competencies schema allows the capture of information about evidence used to substantiate a competency and ratings and weights that can be used to rank, compare, and otherwise evaluate the sufficiency or desirability of a competency.

Resume The Resume specification provides a definition for an XML Resume. The Resume specification includes modules for employment, education, and military history.

Benefits Enrolment The specification supports enrolment and maintenance of human resources in (1) tier-based coverages -- such as medical, dental and vision; (2) Spending Accounts -- more commonly known as flexible spending accounts (FSA); and (3) rate-based coverage -- such as life, short term disability, and long term disability.

Employee StockPlan The StockPlan schema defines the data used to create and maintain plan records for stock option, stock purchase, and stock award programs.

TimeExpense This specification for an "XML timecard" supports the reporting of time worked by the staffing resource performing work on the assignment over a particular time period as well as certain expenses that might be reported by contract or temporary staff.

6.3.2 The case of competency (Measurable Characteristics)

Let's describe a little bit more in detail the competency facet. (see "Competencies (Measurable Characteristics) Recommendation, 2004-08-02", http://ns.hr-xml.org/2_3/HR-XML-2_3/CPO/Competencies.html, for the complete description of this specification).

The aim of the competencies schema is to allow the capture of information about evidence used to substantiate a competency and ratings and weights that can be used to rank, compare, and otherwise evaluate the sufficiency or desirability of a competency.

HR-XML defines a competency as specific, identifiable, definable, and measurable knowledge, skill, ability and/or other deployment-related characteristic (e.g. attitude, behaviour, physical ability) which a human resource may possess and which is necessary for, or material to, the performance of an activity within a specific business context.

HR-XML defines several schemas to represent the different aspects of competency:

Competency

- **name** – (string)
- **description** – (string)
- **required** – (boolean)
- **CompetencyId**
  An identification code assigned to identify or classify the Competency.
- **TaxonomyId**
  A code that identifies the taxonomy.
- **CompetencyEvidence**
  CompetencyEvidence is used to capture information to substantiate the existence, sufficiency, or level of a Competency.
- **CompetencyWeight**
  CompetencyWeight allows the capture of information on the relative importance of the Competency or the sufficiency required.
- **Competency**
  Competencies can have relationships to one another. One way that the Competencies schema allows these relationships to be expressed is through the recursive nesting of one competency inside another.
- **UserArea**
  Attribute used to extend the description with additional attributes
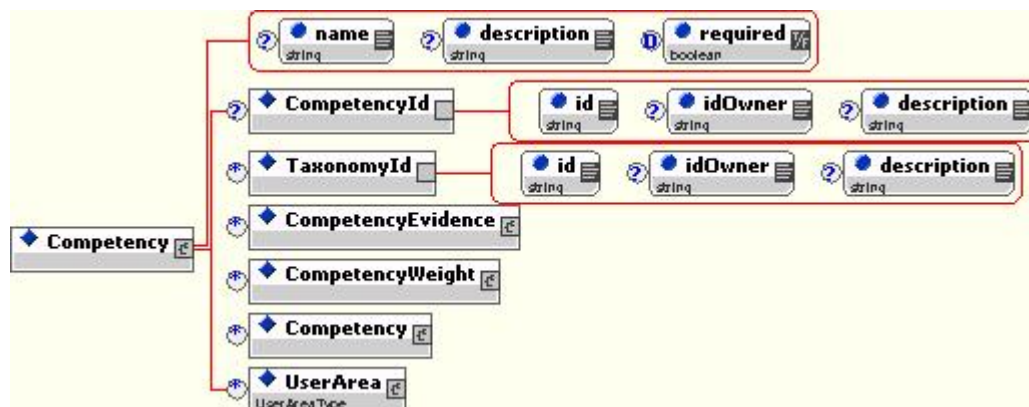
CompetencyId and TaxonomyId have the following structure:

- id – (string) – (required)
- idOwner – (string)
- description – (string)

6.3.2.1 CompetencyEvidence

CompetencyEvidence is used to capture information to substantiate the existence, sufficiency, or level of a Competency. CompetencyEvidence might include test results, reports, performance appraisals, evaluations, certificates, licenses, or a record of direct observation, such as a report given by a former supervisor or other employment reference.

CompetencyEvidence

- **EvidenceId**
  A code that identifies the CompetencyEvidence

- **name** – (string)

- **NumericValue** – (val, min, max)
  NumericValue is the required or desired level for the competency. The content of NumericValue is a rating value.

- **StringValue** – (string)

- **SupportingInformation** – (string)
  Contains additional descriptive information to substantiate or clarify a rating, measure, value, etc.

- **dateOfIncident** – (date)
  The date on which the CompetenceEvidence first establishes the existence of the Competency

- **typeDescription** – (string)
  A description of the type of CompetencyEvidence.

- **expirationDate** – (date)
  The identification of any applicable expiration date, such as the date that a license or certification expires.

- **typeId** – (string)

- **required** – (Boolean)

- **lastUsed** – (date)
  A requirement or assertion for the date on which the Competency was last used.
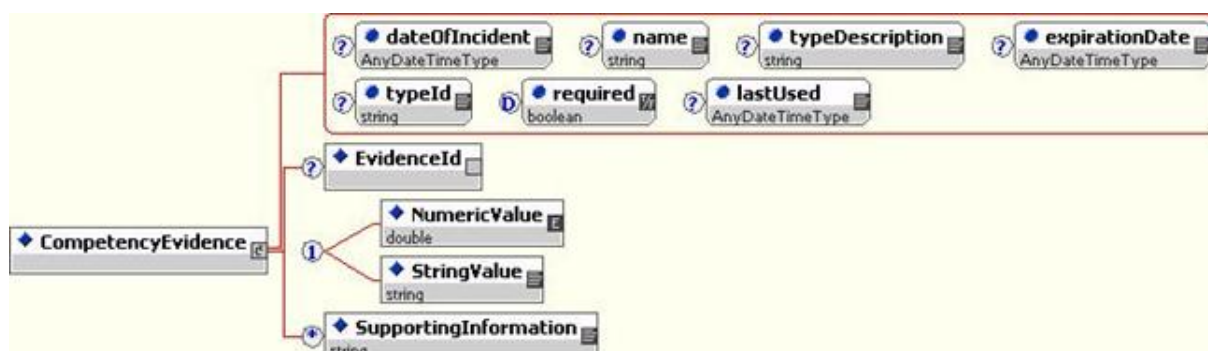
6.3.2.2 CompetencyWeight

CompetencyWeight allows the capture of information on the relative importance of the Competency or the sufficiency required.

CompetencyWeight

- **NumericValue**
  NumericValue is the required or desired level for the competency.
  - value – (double)
  - minValue – (double)
  - maxValue – (double)
  - description – (string)

- **StringValue** – (string)
  Same information as before, but represented as strings

- **SupportingInformation** – (string)
  Identifies the type of CompetencyWeight. Enumerated values are: levelOfInterest (A level of interest asserted or required for the competency) and skillLevel (A level of skill asserted or required for the competency).

- **Type**
  Identifies the type of CompetencyWeight. Enumerated values are: levelOfInterest (A level of interest asserted or required for the competency) and skillLevel (A level of skill asserted or required for the competency).
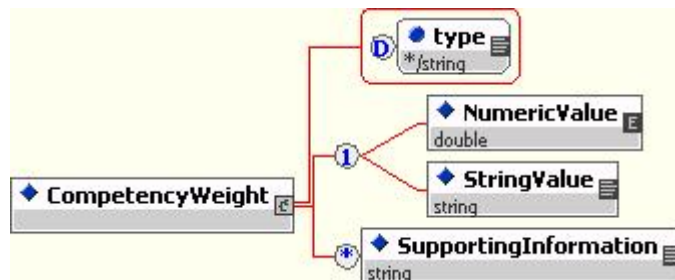
6.3.2.3 Some examples of representation

The following examples are aimed at indicating what kind of competency information (and evidence) can be typically represented:

Example: Competency with Years of Experience and Test Score
Describes an employee who has Java as a competency.
Acme Company, using their standard Java test, tested this employee. The candidate's test score was 89 on that test. In addition, it's recorded that the employee has 4 years of experience using Java and on a scale from 1-100, has a score of 90 for their level of interest.

Example: License as Competency Evidence
Describes a candidate's competency to drive a car.
The evidence used in this example is a valid driver's license that was obtained on 12/23/1986. The measured value in this case is simply 0 or 1 (1 meaning that the candidate has it, 0 meaning they don't).

Example: Education as Competency Evidence
Describes a candidate's competency fulfilling a requirement for Bachelors Degree.   The evidence used in this example is the Bachelors Degree itself, which was obtained on 05/21/1992. The measured value in this case is simply 0 or 1 (1 meaning that the candidate has it, 0 meaning they don't). For additional evidence, the Grade Point average for the degree is presented as additional evidence. In this case, the grade point average is 3.76.

Example: Recursive Competencies: Communication Skills
Communication skills can be made up of two skills: written and oral communication skills. Written and oral communication skills are the measurable and observable skills to which an employee/applicant will be measured. In order to assess whether a person has Communication Skills, it is necessary to evaluate the person's written and oral communication skills (the measurable and observable skills). This example weights oral communication skills higher (65 percent) than written communication skills (35 percent).

Example: Describing Language Skills
One type of skill that is important for many jobs is that of language. It is not sufficient to say that someone is fluent in a language. It is important, especially when looking at what an employee is required to do for a particular job, to look at how the language skill is being used. It may be very important for the individual to know how to read a particular language but there may not be any need to be able to write or to speak it. In other jobs, it may be important to be able to speak a particular language but there may not be any need to be able to write or read in that language.

Example: Direct Observation as Competency Evidence
There are certain skills that need to be assessed, evaluated and evidenced through direct observation. One example of this would be customer service. A manager may observe an employee dealing with customers and notice how the employee treats the customers. It may be difficult to test for customer service and it may be difficult to assess customer service based only on customer feedback.

Note: The xml description of these examples (that can be found in "Competencies (Measurable Characteristics) Recommendation, 2004-08-02"), have been omitted for reasons of simplicity, and also because the objective of this document is not to provide description of the specifications, but rather indicate what kind of information (categories of information, level of details) these specification can represent.

6.3.3 Copyright notice, the HR-XML Consortium

This copyright notice is provided for the content of the previous chapter that contains some parts of the HR-XML documents.

6.3.4 References

"Competencies (Measurable Characteristics) Recommendation, 2004-08-02", http://ns.hr-xml.org/2_3/HR-XML-2_3/CPO/Competencies.html,

6.4 Justice with the Global Justice Extensible Markup (JXDM)

Global JXDM (http://it.ojp.gov/jxdm/) provides a much elaborated schema definition that includes the possibility to describe with much details a person (see the document JXDM xls for a description of all the attributes).

Examples of types of person that have been defined are:

PersonType

- EnforcementOfficialType
- JudicialOfficialType
- CaseOfficialType
- JurorType
- MissingPersonType
- RegisteredOffenderType
- SubjectType

- WitnessType

Examples of other types of elements related to the representation of person's information that have been defined are:

PersonFacet

- PersonAssignedIDDetailsType
- PersonBiometricDetailsType
- PersonDentalType
- PersonMedicalDetailsType
- PersonNameType
- PersonPhysicalDetailsType
- PersonSocialDetailsType
- PersonToothType

Type of Object	Number of Attributes	Description
		Person
Person	77	Describes inherent and frequently associated characteristics of a person. Attributes: (77) Name, AlternateName, Residence, PrimaryContactInformation, Employment, DescriptionText, BirthDate, BirthLocation, BirthPlaceCode, DeathDate, AgeMeasure, AgeMeasure.Range, …
EnforcementOfficial	5+77	A person involved in the enforcement of law. This generally refers to a person designated by a public authority to keep the peace and arrest persons guilty or suspected of crime. May include, but not limited to: police, sheriff, marshal, federal authorities and military. Attributes: (5) TypeText, BadgeID, POSTLicenseID, UnavailableSchedule, Unit
Juror	7+77	A person who serves on a jury and listens to a case to determine the guilt or innocence of a person accused of a crime. This person is a member of a jury, including special or alternate jurors.
MissingPerson	17+77	Details about a person whose whereabouts are unknown. Attributes: (17) ID, DisappearanceTypeText, DisappearanceTypeCode, CircumstanceCode, CircumstanceText, DeclarationDate, DeclarationTime, DeclarationPerson, LastSeenDate, …
Subject	15+77	A person who is involved or suspected of being involved in an incident or criminal activity. This person may be the focus of an investigation or legal process, but is not necessarily the sole or primary focus. Attributes: (15) ID, CautionInformationCaveat, OffenderNoticeText, CautionInformationCode, OffenderNoticeCaveat, WarrantTypeDescriptionText, DisciplinaryAction, Supervision, Status, CriminalOrganizationInvolvementCode, InvolvementCriminalOrganization,

		InvolvementPerson, CriminalTraitDetails, DetainmentLocation, ArmedWeapon
JudicialOfficial	6+77	A person involved in a judicial area of government. May include, but not limited to: magistrate, all types of judges and justices, court clerks, and attorneys. Attributes: (6) TypeText, RegistrationID, BarMembership, PanelID, Court, Firm
VictimEntity	13+77	Details about a person, organisation, or other entity who suffers injury, loss, or death as a result of an incident. This is the object of a crime or tort. Attributes: (13) ID, Organization, Person, Property, TypeText, TypeCode, AffectedProperty, DispositionText, Injury, MedicalTreatmentRequiredIndicator, OffenseChargeText, OffenseChargeCode, SeeksProsecutionIndicator
Witness	8+77	Details about a person who has observed an incident. A person whose declaration under oath or affirmation is received or designated to be offered as evidence for any purpose. Attributes: (8) ID, AccountDescriptionText, LocationDescriptionText, Date, Time, ReportVerifiedIndicator, WillTestifyIndicator, SpecialConditionText
RegisteredOffender	10+77	Details about a person that is required to register their residential information with a local law enforcement agency due to having been convicted of a certain type of crime. Attributes: (10) ID, TypeText, Description, EntryDate, StartDate, EndDate, ReportingCriteriaText, ReportingProfessionName, CourtOrder, Conviction
	Facets of a Person	
Biometric	23	Information used to measurable a biological or behavioural characteristic, which can reliably recognise the identity, or verify the claimed identity, of a person Attributes: (23) BiometricID, ValueText, EncodingMethodText, DescriptionText, Status, Image, Binary, LabAnalysisText, CapturePerson, CaptureOrganization, CaptureDate, CaptureTime, …
Employment	23	Details about the current employment of a person. Attributes: (23) EmployerName, DepartmentName, PositionName, OccupationText, OccupationCode, OccupationTypeText, OccupationTypeCode, RankText, PersonID, Status, ContactInformation, Location, …
PersonAssignedIDDetails	14	A set of details about identifications issued to a person. Attributes: (14) SSNID, TaxID, DriverLicenseID, FBIID, StateID, AFISID, OtherID, RegisteredOffenderIndicator, FirearmSalesDisqualifiedIndicator, LicenseID, GeneralLedgerID, PersonHumanResourcesID, PersonVendorID, PersonNationalID
PersonBiometricDetails	29	A set of details about measurable biological or behavioural characteristics, which can reliably recognise the identity, or verify the claimed identity, of a

		person.
		Attributes: (29)
		DigitalImage, DigitizedSignatureImage, XRayImage, BodyXRaysAvailableText, BodyXRaysAvailableCode, BloodTypeText, BloodTypeCode, Saliva, Semen, Urine, …
PersonInjury	14	Details about an injury a person has received.
		Attributes: (14)
		TypeText, TypeCode, DescriptionText, Date, Time, LocationText, SeverityText, TreatmentText, TreatmentStartDate, TreatmentEndDate, Treater.Organization, Treater.Person, Causer.Force, Causer.Person,
PersonMedicalDetails	10	A set of details about the medical condition of a person.
		Attributes: (10)
		MedicalDescriptionText, MedicalCondition, MedicationRequiredText, LearningDisabilityText, PhysicalDisabilityText, VisionPrescriptionText, MentalStateText, Injury, Intoxication, MedicalFileIndicator
PersonPhysicalDetails	34	A set of details about the physical appearance of a person.
		Attributes: (34)
		HeightMeasure, HeightMeasure.Range, HeightDescriptionText, WeightMeasure, WeightMeasure.Range, WeightDescriptionText, EyeColorText, EyeColorCode, HairColorText, HairColorCode, …
PersonSocialDetails	32	A set of details about the social characteristics of a person.
		Attributes: (32)
		EthnicityText, EthnicityCode, PrimaryLanguageText, PrimaryLanguageCode.iso639-2b, PrimaryLanguageCode.iso639-2t, SecondaryLanguageText, SecondaryLanguageCode.iso639-2b, …

6.4.1.1 Reference

JXDM; "Microsoft Excel spreadsheet of types and properties"; jxdm-3.0.2.xls; Documentation of Release 3.0.2 of Global JXDM; http://it.ojp.gov/jxdm/3.0.2/

6.5 Mobile user (CC/PP & UAProf)

6.5.1 CC/PP

CC/PP (Composite Capabilities/Preferences Profile, www.w3.org/Mobile/CCPP/) is a standard based on Resource Description Framework (RDF). A CC/PP profile is a description of device capabilities and user preferences. For a lot of devices such as a PC, but also a mobile phone or a PDA, CC/PP provides a standard format of information description enabling then any Web terminal to communicate these capacities with a given server in a detailed way.

CC/PP is already used; the toolkit CC/PP intel contains, CC/PP Client Proxy, Client Profile Manager API, Server Profile Repository, Server Content Customization Module. Another interesting example of integration is the UPS (universal profiling schemata) profile package;

it is a Java package for the creation of valid CC/PP profiles (examples of profiles using the specific **fr.inrialpes.opera.upsprofiles** package are available on the web [14]).

The following list (drawn up from the above mentioned examples) shows some attributes useful for the categorisation of the users:

- **HardwarePlatform**:
  DeviceType, DeviceName, DeviceConstructor, screen, screenColor, RAMSize, ROMSize, etc.

- **SoftwarePlatform**:
  PlatformName, PlatformVersion, etc.

- **BrowserUA**:
  UsedPlayerName, UsedPlayerVersion, etc.

- **Resource**:
  OnlySupportedResources, NonSupportedResources, AdaptableResources, PrederredSupportedResources; in a sub-level of description, type, dimension, resolution, version, format, functionality, kind, adaptabilityValue, priorityLevel, requirement, etc.

6.5.2 UAProf

The User Agent Profile specification is concerned with capturing classes of mobile device capabilities and preference information. These classes include (but are not restricted to) the hardware and software characteristics of the device as well as information about the network to which the device is connected. The user agent profile contains information used for content formatting purposes. A user agent profile is distinct from a user preference profile that would contain application-specific information about the user for content selection purposes. For example, a user preference profile might designate whether the user is interested in receiving sports scores and, if so, the particular teams.

UAProf (User Agent Profile) is a specification elaborated by the Open Mobile Alliance (http://www.openmobilealliance.org/) and is based on the CC/PP specification.

Like CC/PP, an UAProf profile is a two level hierarchy composed of components and attributes. Unlike CC/PP, the UAProf specification also proposes a vocabulary – a specific set of components and attributes – to describe the next generation of WAP phones.

Information related to CPI is communicated in an XML document and covers the following attributes of the device[15]:

- Hardware Platform; screen size, audio, images, colour capabilities, manufacturer, etc.

- Software Platform; operating system, content types, etc.

- Applications/User preferences; mark-up languages and versions supported, scripting languages supported

---

[14] http://opera.inrialpes.fr/people/Tayeb.Lemlouma/MULTIMEDIA/CCPP/UPS-Package/UPSProfiles.html

[15] from both sources: http://www.w3.org/Mobile/posdep/ericsson-position-paper.htm and http://www.sourceo2.com/O2_Developers/O2_technologies/O2Guide_to_WAP/UAProf/How_UAProf_works.htm

- Network Characteristics; device location, capability e.g. GSM/GPRS capable, WLTS capable, bearer characteristics such as latency and reliability, etc.
- Browser; Browser name and version, xHTML version, JavaScript Support, etc.
- WAP Characteristics; WML version, script libraries, deck size, OMA download, etc.
- Push Characteristics; Push content types, application, etc.

6.5.2.1 References:

OMA (2003); User Agent Profile; Version 20-May-2003 ; OMA-UAProf-v2_0-20030520-C

6.5.3 Liberty Alliance Profile Service Specification

The Liberty Alliance defines a global framework for identity management. This framework relies extensively on XML technologies (format and services specifications). The Liberty Alliance provides in particular a certain number of schema (expressed as xml schema) to describe person's information.

More precisely, these schemas include:

- **Personal Profile**
  The ID-SIS Personal Profile specification defines a Liberty Identity Service that supports *person's information* regarding the Principal itself, be it in a private or work capacity. It is not intended to be a fully generic contact book and may not address all requirements of e-commerce applications. It is intended to be the least common denominator for holding.

- **Employee Profile**
  ID-SIS-EP provides basic employee information. It is a Liberty Identity Service that supports *person's information* regarding the Principal when in her job function.

- **Contact Book**
  The Liberty ID-SIS Contact Book (ID-SIS-CB) is a Liberty Identity Service that allows a Principal to manage contacts for private and business acquaintances, friends, family members, and even for herself. ID-SIS-CB supports communications applications, acting as Principal's phone or address book, allowing her quickly and easily to locate the information needed to contact people and businesses. Many other applications may also benefit, e.g., an e-commerce application may enable the Principal to ship to any address in the Principal's Contact Book.

- **Geo-location**
  The Liberty ID-SIS Geo-location (ID-SIS-GL) defines a web service. It offers geo-location information regarding a Principal. The geo-location information includes the position of a Principal, speed and direction related information, and information related to the quality of the position information.

- **Presence**
  The core meaning of "presence" refers to a Principal's availability for communications over a network. Examples include availability to talk over a traditional or mobile telephony network, chat over an instant messaging (IM) network, and participate in a video conference. In this core sense, presence is a boolean, "on/off" indicator of network availability. Over time, this core notion of presence has been extended to include other information about the Principal that changes quickly or that affects the Principal's interest in or ability to engage in communications. Examples of such "extended presence" include

Future of Identity in the Information Society (No. 507512)

the Principal's proximity to or interaction with a user agent (e.g., "away" or "do not disturb"), activity (e.g., "driving"), mood (e.g., "grumpy"), and date/time ranges for availability.

Extract of the XML schema for Personal Profile

```
<xs:element name= "PP" type="PPType"/>
  <xs:co mplexType name="PPType">
  < xs:sequence>
        <xs:eleme nt ref="InformalName" minOccur s="0" maxOccurs="1"/>
        <x s:element ref="LInformalName" minOccurs="0" maxOccurs="unbo unded"/>
        <xs:element re f="CommonName" minOccurs="0" m axOccurs="1"/>
        <xs:elem ent ref="LegalIdentity" minOcc urs="0" maxOccurs="1"/>
        <xs:element ref="EmploymentId entity" minOccurs="0" maxOccur s="1"/>
        <xs:element ref ="AddressCard" minOccurs="0" m axOccurs="unbounded"/>
        < xs:element ref="MsgContact" m inOccurs="0" maxOccurs="unboun ded"/>
        <xs:element ref= "Facade" minOccurs="0" maxOccu rs="1"/>
        <xs:element re f="Demographics" minOccurs="0" maxOccurs="1"/>
        <xs:el ement ref="SignKey" minOccurs= "0" maxOccurs="1"/>
        <xs: element ref="EncryptKey" minO ccurs="0" maxOccurs="1"/>
        <xs:element ref="EmergencyContact" minOccurs="0" maxOccur s="1"/>
        <xs:element ref ="LEmergencyContact" minOccurs ="0" maxOccurs="unbounded"/>
        <xs:element ref="Extension" minOccurs="0"/>
  </xs: sequence>
  <xs:attributeGr oup ref="commonAttributes"/>
</xs:complexType>
```

Liberty ID-SIS Service Specification

Personal Profile	Employee Profile	Contact Book	Geo-location	Presence
InformalName LInformalName CommonName LegalIdentity EmploymentId AddressCard MsgContact Facade Demographics SignKey EncryptKey EmergencyContact LEmergencyContact Extension	EmployeeID AltEmployeeID DateOfHire JobStartDate EmployeeStatus EmployeeType InternalJobTitle LInternalJobTitle OU LOU CorpCommonName CorpLegalIdentity ManagerEmployeeID SubalternateEmployeeID Extension	ID-SIS-CB is based on vCard	Eqop (quality) Geoinfo loc_type prio (priority) AreaComparison	Describe availability and desirability for communications over a network

6.5.3.1 References:

Liberty Alliance project http://www.projectliberty.org/

6.5.4 Biometrics (XCBF)

XCBF (OASIS XML Common Biometric Format) aims at providing a standard to describe information that verifies identity based on human characteristics such as DNA, fingerprints, iris scans, and hand geometry.

Biometric types defined in NISTIR 6529.

Biometric Types
Multiple Biometrics Used
Facial Features
Voice
Fingerprint
Iris
Retina
Hand Geometry
Signature Dynamics
Keystroke Dynamics
Lip Movement
Thermal Face Image
Thermal Hand Image
Gait
Body Odor
DNA
Ear Shape
Finger Geometry
Palm Geometry
Vein Pattern

6.5.4.1 References:

XCBF (OASIS XML Common Biometric Format
http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xcbf

NIST (2001); "NISTIR 6529, Common Biometric Exchange File Format (CBEFF)", January 3, 2001, http://www.itl.nist.gov/div895/isis/bc/cbeff/

6.6 Customer information with CIQ

CIQ (OASIS Customer Information Quality) aims at delivering global, application-independent, open XML specifications for party/customer information and profile management. It is still under construction.

The objective of CIQ (OASIS Customer Information Quality) is to deliver a set of XML specifications for defining and managing customer (also called "Party") information/profile (including customer/party relationships) that are open, vendor neutral, application independent and importantly, "Global".

The definition of "Global" here means that the CIQ family of specifications are designed to handle customer/party data (eg. name and address) of any country at an abstract (simple representation of data) or detailed (complex representation, i.e. breaking the data into atomic elements) level.

A customer/party can be one or more persons or organisations. An organisation can be a private/public company, educational institute (e.g., school, college, University), association, club, not for profit, government, consortium, group, etc.

To achieve interoperability of customer/party information within and across an organisation, the optimal approach is to use a "single base customer/party information standard" throughout the organisation to define and represent customer/party data that can support different application requirements (e.g., parsing, matching, validation, verification, postal services, profiling, billing, shipping, customer views, customer relationships, customer segmentation, etc). This standard can then be extended to support specific application needs.

6.6.1 Specification

The committee has developed three XML Standards for Customer Information/Profile Management:

- xNAL: extensible Name and Address Language
  To describe the name and the address of the person
- xCIL: extensible Customer Information Language
  To describe the characteristics of the customer.
- xCRL: extensible Customer Relationships Language
  To describe the relationship with this customer (such as transaction history)

6.6.1.1 xNAL: extensible Name and Address Language

Name and address, as a data type, have unique characteristics, which make it very difficult to manage them. This data is often volatile: customers come and go, addresses change, names change. This data is often cluttered when entered. Name and address fields on front-end screens are usually free format and ripe for users to enter comments and extra data, without any edits.

Future of Identity in the Information Society (No. 507512)

6.6.1.2   xCIL: extensible Customer Information Language

Although name and address data is the key identifier of a customer, other data helps to uniquely identify a customer. Customer addresses frequently change and it is not trivial to link the customer across multiple addresses with just name information.

In the example below, a customer can have two completely different addresses and it is nearly impossible to uniquely identify the customer with the name alone. Customer centric data such as telephone numbers, e-mail addresses, account numbers, credit card numbers etc. will be necessary to achieve this. This helps in achieving single customer view, customer relationship management strategies, understanding customer profiles, etc.

Following are the customer data elements that xCIL Standard supports:

1.   Customer Name and address Details
2.   Customer Identifier
3.   Organisation Details (Branches, Stocks, etc)
4.   Birth Details
5.   Age Details
6.   Gender
7.   Marital Status
8.   Language Details
9.   Nationality Details
10.   Occupation Details
11.   Qualification Details
12.   Passport Details
13.   Religion Details
14.   Ethnicity
15.   Telephone Details
16.   Facsimile Details
17.   Cellular Phone Details
18.   Pager Details
19.   E-mail Details
20.   URL
21.   Financial Account Details
22.   Identification card Details
23.   Person Physical Characteristics
24.   Tax number Details
25.   Vehicle Information Details
26.   Family Member Details
27.   Income Details
28.   Reference Contact Details
29.   Hobbies
30.   Habits
31.   Residency Details
32.   Visa Details

6.6.1.3   xCRL: extensible Customer Relationships Language

Customer relationship management is the key to build effective customer relationships.

Customer relationships could be categorised into the following:

- Organisation to organisation relationship
- Organisation to person relationship, and
- Person to person relationship

A standard way to represent customer relationship helps to achieve interoperability between different systems, processes and platforms and in building effective single customer views.

There are no standards for representing customer relationship and hence, this work attempts to define a standard in XML to capture and represent such relationships.

Some examples of **Person to Person relationships**
Mrs Mary Johnson and Mr. Patrick Johnson, where Mary is the "Wife" of Patrick and Patrick is the "Husband" of Mary
Mrs Mary Johnson and Mr. Patrick Johnson "IN TRUST FOR" Mr. Nick Johnson, where Mary and Patrick are the trustees of Nick and Nick is the beneficiary
Mrs. Mary Johnson, Care of Mr. Patrick Johnson, where Mary is dependent on Patrick
Complete Organisation Structure (Employee-Employee Relationship)

Some examples of **Person to Organisation relationships**
Mrs. Mary Johnson and Mr. Patrick Johnson "DOING BUSINESS AS" Johnson & Associates, where Mary and Patrick are persons who are jointly doing a business under the name of a company called Johnson & Associates
Mrs and Mr. Johnson "IN TRUST FOR" Mr. Patrick Johnson "DOING BUSINESS AS" Johnson & Associates
Mrs and Mr. Venkatachalam "IN TRUST FOR" Mr Ram Kumar and Mr Laxmana Samy "ADMINISTRATORS OF" Sakthisoft Pty. Ltd "TRADING AS" Mantra Corporation
Mr. Ram Kumar, Care of MSI Business System Pty. Ltd, where Ram is the person and MSI Business Systems is the company.

Contact Management

Examples of Contact Management could be, a person maintaining a list of personal contacts, an account manager of an organisation maintaining a list of potential and or existing business contacts, a list management service provider maintaining a list of customers subscribed to their services, etc.

6.6.2 Reference

http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ciq

6.7 Social information with FOAF, XFN, FTML

6.7.1 FOAF (friend of a friend)

FOAF allows the expression of personal information and relationships, and is a useful building block for creating Information Systems that support online communities.

FOAF is an RDF format intended to provide "a way of representing information about people in a way that is easily processed, merged and aggregated" and is primarily concerned with allowing an author to provide a detailed personal description, as well as provide machine-readable links to, and information about, other people.

6.7.1.1 Example:

The following listing represents an example of the description of the Edd Dumbill person, as well as the establishment of the acquaintance of people Edd Dumbill with Simon St.Laurent, using the foaf:knows property.

```
<foaf:Person>
    <foaf:name>Edd Dumbill</foaf:name>
    <foaf:mbox                        rdf:resource="mailto:edd@xml.com" />
    <foaf:nick >edd</foaf:nick>
    <foaf:workplacehomepage        rdf:resource="http://xml.com/" />
    <foaf:depiction                rdf:resource="http://heddley.com/edd/images/edd-shoulders.jpg" />
     ...
    <foaf:knows>
      <foaf:Person>
        <foaf:mbox                 rdf:resource="mailto:simon@xmlhack.com" />
        <foaf:name>Simon St.Laurent</foaf:name>
      </foaf:Person>
    </foaf:knows>
</foaf:Person>
```

Note: the relationship between Edd and Simon is done via Simons mailbox address.

6.7.1.2 Reference:

http://www.foaf-project.org/

6.7.2 XFN (Xhtml Friends Network)

XFN is a simple way to represent human relationships using hyperlinks. In recent years, blogs and blogrolls have become the fastest growing area of the Web. XFN enables web authors to indicate their relationship(s) to the people in their blogrolls simply by adding a 'rel' attribute to their "a href" tags.

6.7.2.1 Example:

Adam and Brad have met each other through mutual acquaintances, and had a few interesting conversations at parties where they found they had several interests in common. They have linked to each other in their respective blogs as follows:

```
// in Adam blog:

<a href="http://brad-log.example.org/" rel="met friend">Brad</a>


// in Brad blog:

<a href="http://adam-log.example.org/" rel="met acquaintance">Adam</a>
```

6.7.2.2 References

http://gmpg.org/xfn/

6.7.3 Other

Other specifications helping to capture social information can be mentioned.

For instance the FTML (Family Tree Markup Language)[16] is used to represent genealogical information, and therefore the relationships that can exist between family members.

6.8 Human characteristics representation with HumanML

The primary purpose of the OASIS HumanMarkup Technical Committee, and thus, the HumanMarkup initiative overall, is to represent human characteristics through XML and interoperable public standards designed for structured information processing, and to enhance the fidelity of human communication, most specifically in digital Information Systems.

HumanML (http://www.humanmarkup.org/) represents a holistic approach for representing human characteristics. This is very much to be considered as a work in progress (and probably not very active).

HumanML is creating schemas to represent the following human characteristics:

- Emotion
- Intention
- Gesture
- Political

Practically the human schema is the following:

ATTRIBUTE GROUPS: (non-cap camel)
- humlTemporalAtts
- humlIdentifierAtts
- humlCommAtts
- physicalDescriptors
    - height
    - weight
    - hairColor

---

[16] FTML (Family Tree Markup Language) http://www.geocities.com/SiliconValley/Lab/6125/genealogy/

- o eyeColor
- o build
- o scarsMarksTattoos
- age
  - o dateOfBirth
  - o dateOfDeath
- gender
  - o genderAtBirth
  - o currentGender
  - o impersonator
- bodyPart
- [arm, leg, head]
- intensity

COMPLEX TYPES (Capitalized camel)
- Address
  - o [postal, residential, email, previous, current]
- Artifact
  - o humlIdentifierAtts
  - o humlCommAtts
  - o humlTemporalAtts
- Belief
  - o humlIdentifierAtts
- BodyLocation
  - o bodyPart
  - o location Locator
  - o humlIdentifierAtts
- Channel: {sight, hearing, touch, taste, smell, kinesthetic}
  - o humlCommAtts
  - o strength of signal

7 Glossary

Acronym		comment
Standards & Technologies		
CC/PP	Composite Capability/Preference Profiles	mobility
CIQ	Customer Information Quality	commerce
FOAF	Friend Of a Friend	Social networking
FTML	Family Tree Markup Language	Genealogy
GPS	Global Positioning System	location
HR-XML	Human Resource XML	Human resources
HumanML	Human Markup language	Holistic
IMS	Identity Management System	
IMS/LIP	IMS Global / Learner Information Packaging	e-learning
JXDM	Global Justice Extensible Markup	justice
LDAP	Lightweight Directory Access Protocol	directory
Liberty Alliance **ID-SIS-CB** **ID-SIS -GL**	Liberty Alliance Identity Service Interface Specifications -CB (Contact book); -GL (Geo-Location)	Identity management
LMS	Learning Management System	e-learning
P3P	Platform for Privacy Preferences	
PKCS	Public-Key Cryptography Standards	authentication
PKIX	Public-Key Infrastructure (X.509)	authentication
SAML	Security Assertion Markup Language	authentication
SIP	Session Initiation Protocol	
SSN	Social Security Number	
UAProf	User Agent PROFile	mobility
vCard	Virtual Card ?	Electronic business card
XFN	XHTML Friends Network	Social networking
XCBF	XML Common Biometric Format (OASIS)	biometric
XML	Extensible Markup Language	Information technology
xNAL	Extended Name and Address Standard	
Organizations		
ANSI	American National Standards Institute	http://www.ansi.org/
EDUCAUSE	(Association promoting the intelligent use of information technology in education)	http://www.educause.edu/
IETF	Internet Engineering Task Force	http://www.ietf.org/
IMS Global	Instructional Management Systems Global Learning Consortium	http://www.imsglobal.org/
ISO	International Organization for Standardization	http://www.iso.org/
Liberty Alliance	Liberty Alliance (identity management)	http://www.projectliberty.org/
OASIS	Organization for the Advancement of Structured Information Standards	http://www.oasis-open.org/
W3C	World Wide Web Consortium	http://www.w3.org/