FIDIS

Future of Identity in the Information Society

Title:	"D 2.1: Inventory of topics and clusters"
Author:	WP2
Editors:	Thierry Nabeth (INSEAD, France) Mireille Hildebrandt (Vrije Universiteit Brussel, Belgium)
Reviewers:	Denis Royer (University of Frankfurt, Germany) Claudia Diaz (KU-Leuven, COSIC) Mark Gasson (Reading University, UK)
Identifier:	D 2.1
Type:	[Deliverable]
Version:	2.0
Date:	Wednesday, 21 September 2005
Status:	[Final]
Class:	[Public]
File:	fidis-wp2-del2.1 Inventory_of_topics_and_clusters.doc

Summary

This deliverable represents the first results of a work aiming at specifying a conceptualization of the Identity domain conducted in the FIDIS project.

The objective of such a conceptualisation is to provide to both the experts and the non-expert a common and explicit understanding of the identity domain, facilitating the comprehension and the sharing of knowledge on this subject.

In this first version, the conceptualisation has consisted principally on the inventory of topics and concepts used in the Identity domain, and in the definition of key Identity concepts.

This document (complemented by a WIKI) is organised into sections providing:

- A methodological presentation of the approaches and principles used to specify a conceptualisation, and its application to FIDIS, in order to conceptualise the Identity domain.
- An overall presentation of key Identity concepts.
- An structured inventory of Identity terms
- A concluding section

Copyright Notice:

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	30.08.2004	Initial release (Thierry Nabeth)
		Continuous developments. Contribution to this document. - VUB: Mireille Hildebrandt. Contributions of some of the key concepts and rereading. - ICPP: Marit Hansen. Provided some content that was integrated in the key concepts. - IPST: Sabine Delaitre. Some contribution on the methodological aspects (Ontology, processes). - JWG: Denis Royer. Some technical aspects + Rereading. - KU-Leuven: Claudia Diaz: first reading. - Univ. of Reading: Mark Gasson: Final reading Contribution to the FIDIS WIKI. - participants from several organisations (JWG, TUB, VUB, VIP, ISRI, Reading, KU-Leuven, IPTS) have participated in providing definition in the Wiki
1.0	21.10.2004	Final version
1.1	8.08.2005	VUB: Mireille Hildebrandt: Clarification of the distinction between identity and identification; state of democracy.
1.2	25.08.2005 & 12.09.2005	VIP: Benoist, Emmanuel, Bernhard Arig, D.-O. Jaquet-Chiffelle: gobal feedback; addition of the concept virtual-identity.
2.0	19.09.2005	INSEAD: Thierry Nabeth. Major revision of section 2. in order to describe better the methods and processes of collection. Revision of the rest of the document.

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	VUB: Mireille Hildebrandt
All	KU-Leuven: Claudia Diaz
All	Univ. of Reading: Mark Gasson
All	VIP: Benoist, Emmanuel, Bernhard Arig, D.-O. Jaquet-Chiffelle
Some	JWG: Denis Royer
Chapter 3	ICPP: Marit Hansen (contributed to some content)
Methodology	IPTS: Sabine Delaitre

Table of Contents

Executive Summary

This deliverable (which consists in this document and in an associated WIKI) represents the first results of a work aiming at specifying a conceptualization of the Identity domain conducted in the FIDIS project.

The objective of such a conceptualisation is to provide to both the experts and the non-expert a common and explicit understanding of the identity domain, so as to facilitate the comprehension and the sharing of knowledge of the subject of Identity.

In this first version, the conceptualisation has consisted principally on the inventory of topics and concepts used in the Identity domain, and in the definition of key Identity concepts.

This document (complemented by a WIKI) is organised into sections providing:

- A methodological presentation of the approaches and principles used to specify a conceptualisation, and its application to FIDIS, in order to conceptualise the Identity domain.
- An overall presentation of key Identity concepts.
- An structured inventory of Identity terms
- A concluding section

The first section in **theoretical and methodological** and presents the main principles related to the specification of a conceptualisation, and **how it can be applied to the context of FIDIS**. The first part of this section first introduces the concept of Ontology (Ontology represents the discipline concerned with the specification of conceptualisations) and indicates the different forms of representation of this conceptualisation (from textual description, to simple taxonomies, or to very elaborated specifications involving semantic representations). It then presents the different methods, processes and tools that can be used to construct such a conceptualisation. In particular it indicates how a technology like a WIKI can be used for this construction by a community of users. Finally, it addresses the organisational and cognitive dimension (how to motivate people and groups to participate to this construction).

The second part of this section is about the application of these principles to the context of FIDIS to specify the conceptualisation of the Identity domain. In this first version, the conceptualisation of the Identity domain consists mainly in the definition of the key concepts and in the identification and the basic structuring (based on taxonomies) of the Identity terms.

The second section describes, using a textual representation, some of **the key Identity concepts** to the reader. It first presents the Identity domain and formulates what are the main Identity issues and concepts. In this description, it distinguishes a structural perspective and a process perspective. In the first case, **Identity** is considered related to the characterisation of a person via a set of attributes, their application in different situations, and how it relates to the person. Examples of concept presented include the distinction between the *ipse-identity* and the *idem-identity* (the inner identity of the person or its external projection), the concept of the *virtual person* or the relationship of the identity with a *territory*. In the second case, with the concept of identification, identity relates to the set of approaches, mechanisms and processes involved in the disclosure of the identity information, in the course of an interaction.

Examples of associated concept that relate to processes include anonymity, pseudonymity, observability, linkability, etc.

The **third section** presents an inventory that **categorises** and defines the different terms used in the Identity domain. It can be seen both as a dictionary and a map of the Identity domain, in which the reader can easily locate a term and get a definition. This section also details how the WIKI technology is being used in the project to support this process of conceptualisation. In particular, it proposes templates and approaches to be used to define the terms and their relationships, and indicate processes to be used for a community to collectively participate to this collection.

The **last section concludes and summarises** the work, indicate the principal issues that were identified, and gives some indications for future work. It is suggested that more elaborated representation will be initiated in subsequent versions of this conceptualising work, once more content will have been generated in the FIDIS project.

1 Introduction

1.1 Scope

This deliverable (document + WIKI), created in the context of the workpackage 2 of the **FIDIS Network of Excellence**, represents the first result of **the task aiming** at **specifying the conceptualisation of the Identity domain,** with a particular emphasis on a **multidisciplinary perspective**.

The objective of this conceptualisation is twofold:

- **to contribute to the understanding** of the many dimensions of the **Identity domain**. This is done in this first version via the identification and categorisation of the most important themes and topics of this field, and the description / definition of key Identity concepts;
- **to facilitate the sharing knowledge** on this domain amongst the different stakeholders. This is done through the definition of a shared vocabulary to be used in the Identity domain, and the setting-up of the conditions for the dynamic of exchange of this knowledge inside the FIDIS community.

The objective is also to provide a relevant contribution to the body of "scientific" knowledge, through a better semantic formalisation of the concepts associated with the Identity domain.

The **audience** of this document comprises both**:**

- **the non-experts** (citizens, employees, civil servants, …), who need to quickly understand the main issues in the domain of Identity or who wish to find a particular definition of a term that they may have encountered during readings or discussions;
- **the experts** who want to get a more holistic (multidisciplinary) perspective of this domain, a more unambiguous and agreed definition of the terms they are using, and better ways to communicate with others (non-experts or experts originating from different disciplines and thus having a differing background).

This document tries to be **comprehensible** (but not simplistic) even to the non-specialist. It presents the definitions simply, according to multiple facets, and is illustrated with concrete examples where possible. During the design of the document, the emphasis has been put on the comprehensiveness of the semantic of the terms and the desire to disambiguate the meaning(s) of the concepts, rather than on thoroughness and completeness of the definitions (more complete and "rigorous" definitions are to be found in more specialised documents).

Finally, the principal focus of this first conceptualisation work is on **setting-up** the solid foundations (both technical and organizational) for a lively **process** of continuous evolution of the specification of **conceptualization** of the identity domain in the consortium. It is assumed that this conceptualisation is going to emerge from the work that is conducted in this project and from the interaction of the members of the consortium. This conceptualisation is intended to be an "**active and living** thing", subject to continuous changes and additions as the field of Identity evolves (new concepts will progressively be integrated as the FIDIS project advances integrated). Its aim is also to help linking the different stated definitions of identity in the different perspectives (technical, social, legal, etc.) involved in this project, and / or present in the different workpackages. The final objective aims to capture and to reflect

the current state and **share understanding of the Identity field as seen by the "FIDIS community"**.

This deliverable (document + WIKI), aiming at facilitating the comprehension of the Identity domain, and the diffusion of awareness of the Identity issues, is complemented by two deliverables: (1) "D2.2 Set of Use Cases and Scenarios", which will provide more elaborated examples, cases, stories and scenarios, linking the different identity terms with one another; and (2) "D2.3 Identity models", which will provide a more formal and complete characterisation of how Identity is represented (in particular what is the data model of the identity of a person).

1.2 The Content and Structure of this Document

The first section of this document is methodological, and presents the approach that can be used to specify the conceptualisation of a domain in general, and how it can be applied in the context of the FIDIS project. It first provides a basic theoretical background related to definition of concepts and more precisely relates to the Ontology approach for explicitly specifying the semantic of a set of terms of a domain. It then indicates the set of processes and tools that can be employed to collect and construct the conceptualization. It finally presents how these principles can be applied in the context of FIDIS to specify the conceptualisation of the Identity domain.

The second section provides a relatively in-depth overview of the terms and topics that are considered as central to the Identity domain. Examples of the terms and concepts that are covered include: identity profile, multiple identity, digital territory and application context of identity, elements associated with the topic of identification such as privacy, anonymity, pseudonymity, etc.. The objective of this section is to introduce to the reader the main topics and issues relevant to the Identity domain. Most of its content is constituted by the content of existing documents originating from some of the members of the FIDIS consortium such as (Hansen and Pfitzmann, 2004), (Prime, 2004), or (Jaquet-Chiffelle *et al*., 2004), from the presentations at the first WP2 workshop (FIDIS WP2 workshop, 2003) and from the direct contribution from the FIDIS members.

The third section proposes an inventory of terms that are relevant to the Identity domain; it makes a tentative (tree-like) categorisation of these terms. It also proposes a set of templates to be used to systematically define terms and concepts, describes the WIKI infrastructure that has been set in place to support this identification process and the collection of the definitions of the terms and concepts. It also indicates the mechanisms and processes that have been set-up to encourage the participation of the members of the consortium in this construction.

The last section concludes the document providing a short assessment of the work presented, and some directions for future work.

2 Approach and methodology

The objective of this chapter is to describe the models and methodologies employed in WP2.1 to specify in FIDIS the conceptualisation of the Identity domain.

The first part of this chapter, which presents the global perspective, indicates how to approach the definition of concepts for a given domain. This part mainly relates to Ontology, i.e. the discipline aiming at the study of explicit specification of conceptualisations. It also indicates approaches and methods can be used to build these conceptualisations, as well as some tools (such as Wikis or the semantic technology tools) that can be employed for this purpose. It finally addresses the organisational and cognitive issues related to this process of construction (such as the motivational factors of the individuals and of the groups to participate in this construction).

The second part of this chapter presents how the principles of conceptualisation (or Ontology) can be applied in the context of FIDIS to construct a conceptualization of the Identity domain. More specifically, it indicates for FIDIS the model of the conceptualisation of the Identity domain, presents the methodology and the processes that have been adopted, and describe how the WIKI infrastructure can be used to construct collectively this conceptualisation in the FIDIS community. It also indicates how FIDIS addresses the organisational and cognitive dimensions.

2.1 Specifying a conceptualization (the Ontology concept)

"**Ontology**" (with a big "O") represents the "discipline" is concerned with **the study of how to specify conceptualizations**. The term "**ontology**" (with a small 'o') is used to refer to the concrete specifications of a conceptualisation of a given domain (Gruber, 1993).

In this section, we are going to present the theoretical perspective of this conceptualization, via a description of the concept of Ontology, what ontologies are used for and how to design them. We will also mention some of the technologies (WIKI and semantic web tools) that can be used for the collaborative construction of this conceptualisation. We will conclude by looking at the organisational issues such as roles assignments and motivational aspects.

Note 1: In the rest of this document, we use indistinctly the term "ontology" or the expression "specification of a conceptualization" to refer to the same concept. We have decided to keep using the expression "specification of a conceptualization" in some places just to keep this document more accessible to the non-specialist (the term "ontology" being sometime perceived as too complex for a "normal" audience).

*Note 2: The Ontology concept that is described in this chapter is intended for the specialist or for the non-specialist that would like to be aware of the theories of conceptualisation underlying this work. The **less expert readers** (or the expert who is only interested to get a quick overview) **are invited to skip this first part and directly go the second part (2.2)** which describes how these principles are being applied for defining the Identity concepts in FIDIS.*

2.1.1 Defining the Ontology concept

2.1.1.1 The specification of conceptualisation in History

The explicit and unambiguous specification of the concepts (i.e. the definition of ontologies) of a given domain has been the object of a lot of attention along History.

First in the Greece antiquity, the **philosophers** had come across the need to conceptualize the concepts in their aim of better understanding the nature of the Being. They invented term Ontology that they defined as the branch of metaphysics relating to the nature and relations of being. At this time this conceptualization was mainly done through writing and discourse. Since then, Ontology has in various times received the attention of the philosophers[1].

Then in the Middle-Age and later at the Renaissance, people have began to more systematically and explicitly specify the conceptualisation of a domain by using **dictionaries** and **encyclopaedias** (the reference first reference to the term dictionary be traced in the 13th-century[2], and the modern Encyclopaedia can be dated at the beginning of the 16th-century). Dictionary and encyclopaedia represents a way of specifying a conceptualisation that is based on definition, in alphabetical order, of the terms or words of a domain (dictionary), or on the subjects of a domain (encyclopaedia).

In the 19th century, **classification** played a key role in **Natural Science**, and one can cite the work related to the classification of species of Lamark, Buffon and Darwin that played a considerable influence in this area (and is at the root of genetics). Classification relies on the idea of conceptualising a domain based on the identification of a set of characteristics that can be own by an object and that is usually hierarchically structured (example of classification: the library classification of subjects[3]; or the classification of species in biology).

Computer Sciences has shown early an interest in the very explicit specification of concepts. The aim was at making the specification of concepts comprehensible by machines. For instance, as a necessary condition for conducting automatic operations and reasoning, the domain of Artificial Intelligence or Advance Computing started early trying to define explicit and formal specifications of knowledge (Aiii, 2004): Examples include Allen Newell's research on symbolic computation in the mid 50ies, then Ted Nelson's invention of Hypertext in the 60ies, then Marvin Minsky with the introduction of the concept of Frames in the 70ies, and later Douglas Lenat with his work on the Cyc framework aiming at representing common sense in the 80ies.

More recently with the advent of the Internet, the Computer Sciences field has generated a lot of activities around the Semantic Web (Berners-Lee, Hendler and Lassila, 2001). In this context ontology work relies mainly on the idea of conceptualising a domain in term of objects and semantic relationships. This trend towards the semantic web has dynamised

---

[1] See "Ontology: A resource guide for philosophers", by Raul Corazzon. http://www.formalontology.it/

[2] The first recorded use of the term "dictionary" to mean "word list" can be associated with the 13th-century Dictionarius of John of Garland; the first edition of the Webster dictionary of the English language was launched in 1806.

[3] For instance the Dewey Decimal Classification (DDC) system or the Library of Congress Classification (LCC) provide a dynamic taxonomical structure for the organisation of library collections. Note: These classifications should be distinguished from other classification in library such as the Dublin Core (http://dublincore.org/), which aim at defining the structure of the objects (books, authors, etc.) that are dealt in a library.

research and use in this domain, even to the extent of creating dedicated standards (such as OWL[4]) for representing ontologies.

In a parallel track, knowledge construction and categorisation has flourished, and new approaches have been invented such as combination hyper-textual and collaborative knowledge construction which is best exemplified by Wiki systems[5].

2.1.1.2 What is an "ontology"

As we have seen previously, we define ontology (with a small 'o') as a particular specification of a conceptualization of a given domain (Gruber, 1993)[6], and Ontology (with a big "O") as the discipline which is concerned with the study of ontologies.

Ontologies represents a broad concept that can take a variety of forms of diverse degree of complexity ranging from simple **textual descriptions** (aiming at the explicit description of a set of concepts), **inventories of terms and their definition** (such as dictionaries and encyclopaedias), **taxonomic hierarchies** (such as classification organised as trees)**,** and in the more complex case **semantic networks** (set of objects connected with one another with relationships) which structure is specified in a **meta-model** (specifying the structure of the objects, and providing some categorisation).

The different categories of ontologies differ in the level of deepness and formalisation they adopt for specifying the conceptualisation.

In the case of the simple text-based ontology, the level of formalisation is usually reduced to a minimum, and the description of the structures is totally implicit (in the best case the text provides a structure helping to make the organization of the concepts visible). More advanced ontologies are based on structural representations (lexical, syntaxic or semiotic). For instance the knowledge can be organised into synonym sets, each representing one underlying lexical concept. This later approach is typically adopted by the more advanced dictionary approaches (such as the Worldnet system[7]). In that case however, the structure is not used to represent the semantic. Classifications (taxonomical) start to use the structure to capture certain aspects of the semantic. For instance the different objects belonging to a same branch share some identical properties. Finally, the more sophisticated forms of ontologies (the ones that are promoted by the Semantic Web[8], and which aim at being interpreted by machines) operate directly at the level of the semantic. In this case "semantic ontologies" specify the concepts in term of semantic definitions of the objects that intervene in the domain and of the semantic relationships that hold amongst them. Besides, in order to make this specification more explicit, a distinction is generally done between the semantic network describing the domain (called the instance Ontology) and the meta-model used to describe the different classes of objects and relationships that form this network.

---

[4] OWL: (Web Ontology Language) is an XML/RDF based standard for representing Ontologies. http://www.w3.org/2004/OWL/

[5] For a definition of the WIKI concept, see "2.1.3.12.1.3.2 Terms collection tools with WIKIs"

[6] We use in this document the definition of Ontology that is used in the Information Sciences field (in particular in the Semantic Web area of research).

[7] WordNet® is an online lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory. http://wordnet.princeton.edu/

[8] The Semantic Web : http://www.semanticweb.org/

It appears legitimate now to ask the question of the most adequate form of ontology to use to specify a conceptualisation, in particular when you consider the important difference in complexity between the different forms of Ontologies. Indeed, very explicit conceptualisations like the semantic one can require a considerable amount effort that may not be justify, and can even in some case prove counter productive (by reducing the flexibility in the case of a domain continuously changing). On the other hand too shadow conceptualisations like the textual and can be ambiguous, and lead to partial understanding. They can also be more difficult to exploit by the information technologies.

The answer is not simple and varies according to the expectations in term of the quality of the ontologies (for instance their level of deepness and completeness in the specification of a conceptualisation), the nature of their exploitation (are they to be used in the context of an information system infrastructure?), but also the size of the domain to be conceptualized and the effort and expertise available for their design.

Actually, the different approaches for conceptualising a domain can be considered as complementary. Simplest ontologies (typically the textual ones) can be used in a first stage to clarify the domain. More sophisticated conceptualisations (categorisation, semantic representations, etc.) can later be used once the domain is better understood (thanks to the previous conceptualisation).

However, some too elaborated conceptualisation are not always desirable. They can indeed be unnecessary complex and create rigidity (for instance when a domain is still in a stage of continuous evolution). They may also require tools and computer resources that are currently not available. For instance, some applications (such as data mining or machine learning applications[9]) may have to manipulate huge amounts of data, making the use of in-depth ontologies unpractical and unnecessary.

2.1.1.3 What is an Ontology used for

It is important to question the reasons why Ontologies are built, and indeed what the usages of the Ontology actually are.

One of the main objectives of Science is to create very explicit models of the functioning of the world, that can later be applied without any other necessary or hidden knowledge (so that a phenomenon can be deterministically reproduced). Ontology work, by providing explicit and well understood definitions of the concepts, facilitates the specification of scientific models and in particular, contributes to the description of the scientific models in a way that is concise (no need to describe concepts that have already defined) and unambiguous (reducing the risk of multiple and / or erroneous interpretation).

The second reason, that is at least as important, is the creation of a common language that facilitates the sharing, exchange and reuse of knowledge amongst the community of people (researchers, practitioners, final users) who deal with these concepts. For instance, it is generally accepted that well formed ontologies can significantly improve communication and provide a basis for shared understanding and reuse of information (Ushold and Gruninger; 1996; Clark & Brennan, 1991).

[9] See for instance in (Canhoto and Backhouse, 2004) for an application that uses relatively simple representation , but that uses data mining to analysis suspicious transactions.

Finally, another function of Ontology work is to provide a mechanism to stimulate the construction of knowledge in the community, and a means for this community to develop its identity. Most specifically, making more visible the most important concepts and making people use the same terms to refer to them (versus many different terms to refer to the same concept) help to the development of a sense of belonging to this community and resemblance. This sense of belonging is important, since it leads to important outcomes by increasing organizational citizenship behaviour -loyalty, civic virtue, altruism, and courtesy-, motivation and commitment, and can be associated with involvement in community activities (Blanchard and Markus, 2002). Resemblance can contribute to the establishment and development of people relationships, since people tend to establish relationship with other people that are similar to themselves (Berscheid and Reis, 1998).

Note: As a consequence of this support for the community process, an Ontology should not be considered as a finished and static piece of information, but a living information body that is continuously growing (and in particular integrating new terms and concepts as they emerge) and adapting to the evolution of the focus of the community that is using it. It is therefore important that the mechanisms for adding vocabulary to this language are not seen as an external component of this ontology, but they are, on the contrary, directly built into them.

2.1.2 Ontology design

2.1.2.1 The ontology design challenge

The design of good ontological constructions is a non-trivial operation that typically requires a lot of time, effort and resources.

The design of Encyclopaedias and dictionaries are known to have represented a major effort, that had consisted in the **identification** of a very important number of terms and their **definition** or their **illustration** by examples (in encyclopaedia).

In Natural Science, the definition of classifications had required decades of work, and successful models (such as the classification of species) were considered as the major achievements of some famous scientists (Buffon, Lamarck, etc.). Scientist went through a **laborious and extensive collection of data** about animals that were later **classified** using empirical methods (typically finding **similarities** in the data that were collected).

More recently, Computer Sciences, and more particularly **the semantic web**, has defined some more systematic approaches, **methodologies** (Holsapple, and Joshi, 2002; Guarino and Welty, 2002, Denny, 2002; Prieto-Diaz, 2002; Noy and McGuinness, 2001) and tools **to conceptualise a domain** in a way that is interpretable by computers. However as Guarino and Welty (2002) points out, “The process of building or engineering for use in information systems remains an arcane art form” that would need to become a rigorous engineering discipline. (Missikoff, Navigli and Velardi, 2002) report some experiences in the designing of domain ontology taking several months. It also stresses the importance of designing Ontologies good enough to be *usable* and that have in particular a good level of: coverage (level of completeness), consensus (agreed upon), and accessibility (easy to use). Finally Noy and McGuinness (2001) acknowledge that “there is no single correct ontology-design methodology”.

To conclude, it appears impossible to identify an agreed and usable Ontology design methodology. In the next paragraph we will therefore give some indications of an approach for building Ontologies that we belie could be used in the FIDIS context.

2.1.2.2 The design process

Denny (2002) proposes the following steps to build an Ontology domain.

- Acquire domain knowledge.
  Assemble appropriate information resources and expertise that will define, with consensus and consistency, the terms used formally to describe things in the domain of interest. These definitions must be collected so that they can be expressed in a common language selected for the ontology.
- Organize the ontology.
  Design the overall conceptual structure of the domain. This will likely involve identifying the domain's principal concrete concepts and their properties, identifying the relationships among the concepts, creating abstract concepts as organizing features, referencing or including supporting ontologies, distinguishing which concepts have instances, and applying other guidelines of your chosen methodology.
- Flesh out the ontology.
  Add concepts, relations, and individuals to the level of detail necessary to satisfy the purposes of the ontology.
- Check your work Reconcile syntactic, logical, and semantic inconsistencies among the ontology elements.
  Consistency checking may also involve automatic classification that defines new concepts based on individual properties and class relationships.
- Commit the ontology
  Incumbent on any ontology development effort is a final verification of the ontology by domain experts and the subsequent commitment of the ontology by publishing it within its intended deployment environment.

Whilst this approach can be a source of inspiration, we believe this approach is not totally appropriate in the case of FIDIS, for which the domain is still considered as relatively fuzzy, and subject to evolution.

Besides, in WP 2, we are interested to make use of an approach which is more consistent with the WIKI tool, which we believe can appear to be particularly well adapted for the collection of resource from the all FIDIS community.

The construction process of the ontologies that we propose consists in a series of iterations involving the following operations:

- Identification of terms, concepts, etc. (corpus analysis)
- Categorisation of these terms
- Definition of their semantic
- Illustration (situating them with examples)
- Some tentative formalisation of the conceptual relations

The Identification of terms, concepts, etc... (corpus analysis)

The identification process of terms and concepts consisted in going through some content material belonging to the domain, and extracting the vocabulary of the terms that are the most significant.

The categorisation

The objective of the categorisation of the different terms, and the identification of the main concepts was the creation of a tree-like cognitive map of the domain, connecting the different terms with one another.[10] The tree representation is very intuitive and user-friendly. This type of representation allows user an easy navigation among the terms and an understandable way to handle the relation between the terms with one another.

The definition of the semantic

Each identified term and concept is progressively defined. This definition consists of providing a basic definition, structuring it, and beginning to link the term with other terms that appear to be related. The semantic is closely related to the term use context. This definition process is done via the extraction of content from different material, and from the active and collaborative contribution of the different participants.

The illustration (Situating the concept in concrete contexts)

As part of the definition process, a set of illustrative examples are provided in order to better link the term or concept to a concrete context, and thus making clearer to the reader what the different associated identity issues are.

Formalisation

Some tentative formalisation can be initiated in order to homogenise the definitions of the different terms and their relation with one another (and for instance, identifying the conceptual relation between the different terms). The formalisation however is expected to emerge (rather than be imposed from the outset), as more definition of terms are incorporated.

2.1.3 Some tools (ontology editors, WIKI, etc.)

Different tools can help this process of collection and specification.

2.1.3.1 Modelling tools (such as Ontology editors)

The first category of tools include editors that help to design Ontologies such as the Protégé system[11]. This kind of tool is only adapted for the representation of "semantic" ontologies. We will not detail this solution now, because we believe it is not currently adapted –at least at the initial stage- to the definition of a domain as broad and still subject to important evolutions.

---

[10] The use of Semantic technologies (Ontology tools) were envisaged, but has been temporarily postponed to a later phase in favour of Wiki techniques which offer more flexibility and are better able to support collaborative collection and definition of terms and concepts.

[11] The Protégé Ontology http://protege.stanford.edu/

Note: we can also mention graphical modelling tools that could be used to help represent objects and relationship. However, we are not aware of experiences using these tools to help in the design on sophisticated Ontologies.

2.1.3.2 Terms collection tools with WIKIs

Another more interesting tool is constituted by WIKI, which provide the possibility of a distributed group or community to contribute to the definition of connected terms and concepts, using hyper textual relationships.

A Wiki or WikiWikiWeb (pronounced [wicky] or [weeky] or [viki]) is a website (or other hypertext document collection) that allows users to collectively write documents using a web browser and a simple mark-up language for formatting these documents. One of the defining characteristics of wiki technology is the ease with which pages can be created and updated. Generally, there is no prior review before modifications are accepted, and most wikis are open to the general public - or at least anyone who has access to the wiki server.

WIKI systems have received an important level of attention recently, and are even considered sometime as the next big step following the blogging revolution (Rand, 2004; Cooper, 2005).

One of the important strength of WIKIs is that they allow a group of distributed people, to very easily collect information about topics, and to connect them with on another. We invite the reader interested to understand the functioning of WIKIs to have a look at an open and living WIKI such as Wikipedia[12]. Wikipedia is the most well known example of WIKI, and is a free online encyclopaedia comprising more than 732 000 articles as of September 2005 that has been authored by thousands of independent contributors without centralised supervision[13].

In the context of this section related to the specification of a conceptualisation, a WIKI represents a very powerful tool that can help a community to collect definition of terms and create very easily hypertextual relations.

2.1.4 The Organisational dimension

In this last chapter, we are going to mention the organisational and behavioural dimension related to the process of Ontology building. This aspect is particularly important, if we are in the case in which this ontology will have to be authored by several contributors.

The formulation of the organisational and behavioural dimension of Ontology building is the following: “How do you stimulate and motivate people in a group to participate in the construction of an artefact?”

The answer to this question is not new and has many different answers.

In this chapter we are just going to present the different issues that are related to the question of people participation in some collective actions. In chapter “2.2.2 The Organisational

[12] Wikipedia at http://www.wikipedia.org/

[13] Jimmy Wales, the founder of the Wikipedia, says, "The wiki model is different because it gives you an incentive when you're writing. If you write something that annoys other people, it's just going to be deleted. So if you want your writing to survive, you really have to strive to be cooperative and helpful."

dimension in FIDIS", we will provide indication about how these issues are been addressed in FIDIS in the context of the construction of the Identity WIKI.

2.1.4.1 The motives of people participation

People participation not because a collaboration infrastructure has been set-up (the experience has shown that the building of "an infrastructure" is never a guaranty that "people will engage into a collaboration", nor that they will sustain a collaboration), but because they have some reasons to do so, and only after a certain number of conditions are met. The theories of people participation in knowledge exchange processes, which states that people share they knowledge in order (Hall, 2001) to get a direct benefit, to increase their reputation, for internal satisfaction (altruism and efficacy), and for expected reciprocity, can provide us with some hints about the reasons that make people collaborate. First we can think that people participation because they perceive a direct personal benefit in this participation. But people can also participation for other less individualistic or direct reasons such as: the expectation to consolidate or develop their social capital, some altruistic and efficacy reasons (such as the satisfaction to work for the common good of the group), or because of some expected reciprocity ("if I provide some assistance to others, I can imagine that I will get assist from them when I will need help"). Of course, people may also participate because they have been ordered to do so, although this is something that may be difficult to enforce in an online setting.

2.1.4.2 Design principles to stimulate people participation

It appears possible to derive from these theories of participation some principles that can contribute to the establishment of the condition for collaboration inside a group and some strategies to activate such as: (1) the development of a climate of trust (Tung et al., 2001). For instance Jarvenpaa and Leidner (1999) have investigated the importance of trust in virtual team for the effectiveness of this collaboration. (2) The development of a sense of community (Blanchard and Markus, 2002; Koh and Kim, 2003). In the context of group collaboration, it may appear appropriate to invest some effort into developing a common understanding and share values inside the groups following the Clark (1996) common-ground theories of communication (Clark & Brennan, 1991). And (3) a feeling of recognition for the actions of their members (Chan et al., 2004). Obviously, mechanisms can be implemented to make the collaborative activities of the members, and their contribution to the group more visible.

2.1.4.3 Other barriers to overcome

Yet, having said that, a certain number of individual barriers still exists that may prevent people to "make the jump", abandon their old practices and adopt the new collaboration practice such as: (1) people may be afraid by the complexity of a domain they do not know, and be afraid to ask (I will look stupid if I ask about something that everybody seem to know, even my children!); (2) people may be unable to evaluate the cost of switching, and to enter in an unknown territory (will the new practices will really be more effective for me than the ones I am familiar with? What is the effort?); (3) people may be frightened not to be capable to learn these new practices (what will happen if I fail?); (4) people may wrongly believe that they already have these skills and not feel the need to learn (for instance some people consider that they already know. "already been there; done that, etc.").

2.2 The WIKI conceptualisation of the Identity domain in FIDIS

Defining "the perfect ontology" of a domain as complex and evolving as the Identity domain represents a major (and probably impossible) undertaking that the FIDIS project does not pretend to achieve:

First, because there exists many different possible categorisations of terms, each of them more adequate for the use in a particular context; Second, because the domain "Identity" is very dynamic and continuously evolving (new issues are emerging), and cannot be easily "institutionalised" yet; Third, and as a consequence, because the elaboration of a "perfect ontology" may not be a desirable goal: the objective of this ontological work is not so much to impose on a community a particular vision of the meaning of identity, but rather to give this community a tool (the shared vocabulary) that will help it to exchange and share ideas, and to collaborate in the creation of new knowledge and meaning (Wenger Etienne, 1999); Fourth, because this represents an unfeasible task: ontology building is still more of an art than an engineering discipline, and requires considerable time and effort.

The FIDIS "Identity" ontology has therefore been considered from the outset as a tool for the community, formed by the FIDIS network, to develop a shared understanding of the domain. The value of this "Identity" Ontology construction is therefore not only seen by FIDIS as simply its first output (the set of definition of terms), but also a way of providing a mechanism that helps this community to developed its own – dynamic and diverse - identity.

The FIDIS "Identity" Ontology is meant from the very beginning as a "living entity", constantly engaging the FIDIS members in some interaction, and continuously integrating the new findings of the "identity" domain. This collaborative process involves all the members of the FIDIS community in proposing terms, categorisations, definitions and illustrative examples.

The use of a WIKI appears to represent a tool of choice for supporting this vision of continuous process of conceptualisation, and was therefore adopted.

In the following chapter, we are going to present the different process of building the Ontology that will be "captured" into the FIDIS Identity WIKI (the description of process of building the textual conceptualisation of chapter 3 will not be presented, because it only consisted in the well know editorial work).

2.2.1 The construction process of the FIDIS Identity conceptualisation

In order to create the FIDIS Ontology, a bottom-up, constructivist, iterative and collaborative approach was adopted. This choice has been preferred in order to better manage the relative novelty of the Identity field, which is not yet stabilised, to allow more flexibility (and rearrangement), to facilitate the collaborative process and the progressive emergence of meaning. An alternative approach (that has not been adopted) would have consisted in a more linear process putting a stronger and earlier focus on the formalisation of the structure (definition of a "classification" and of the conceptual relations between the terms), and the

utilisation of this structure to "organise" the different terms (seen as instances of classes, connected to one another via well defined semantic relations).

Besides, this approach was supported by the use of a Wiki, provided as part of the FIDIS information infrastructure. A Wiki offers much flexibility, and is well adapted for concurrent and collaborative activities.

As we have indicated previously, the construction process of the FIDIS ontology consists of iterations of the following operations:

- Identification of terms, concepts, etc.
- Categorisation of these terms
- Definition of their semantic
- Illustration (situating them with examples)
- Some tentative formalisation of the conceptual relations

2.2.1.1 The Identification of terms, concepts, etc... (corpus analysis)

The identification process of terms and concepts has consisted of:

- The extraction of terms from this collection of Internet resources[14]
- The extraction from more formal identity domain documents (such as (Hansen and Pfitzmann, 2004), or the FIDIS deliverable elaborated in the project).
- A continuous process of collecting terms from the FIDIS members.

The first operation (of course very incomplete) of collection of Internet resources related to the concept of identity, was aimed at making a first state of the art description of this domain so as to get a first idea of the main identity issues, but also to collect a set of reference materials that could be used later to illustrate a particular identity issue. Examples of the categories of Internet resources include: sources of information, events, journal and magazines, articles, actors and categories of actors in this domain (companies, laboratories, etc.), projects and initiatives (research projects, consortium), etc. This collection of resources helped in particular to identify the terms, concepts, mechanisms and issues that are the more frequently found related to identity. Examples includes items such as phishing (criminal activity relying on the use of a fraudulent websites mimicking an official web site, which is designed to fool recipients), RFID tagging (which consists in the use of a rfid device to add identification attributes to an object), or adware (software pushing advertisements and that typically spy the activities of the users).

The extraction from more formal identity domain documents (which can be considered by themselves as text-based specification of conceptualisation) help to identify, and to find definition of the more elaborated identity concepts. For instance (Hansen and Pfitzmann, 2004) defines very thoroughly a set of identity concept such as: Anonymity, Unlinkability,

[14] The URL of this resource collection is currently available at: http://www.calt.insead.edu/fidis/information/news/.
Note: The content of this collection of resources will be moved to the public FIDIS WIKI identity definition space (to be available via the FIDIS main public web site) when this space will be opened.

Anonymity, or Pseudonymity. These concepts that can directly be imported and become part of the FIDIS conceptualisation (with the appropriate acknowledgment of the origin of the work done). This extraction process can be considered as even more important in the case of the use of the documents deliverable generated in the FIDIS project, since it provides the means to link all the different stated definition of identity that are formed in the different workpackages.

Note: the use of the FIDIS deliverable documents was very limited in the first version of the specification of conceptualisation generated by the workpackage WP2. Indeed, few of these FIDIS deliverables were available at the time of the elaboration of del 2.1., and besides the objective of this first version was principally oriented in laying-up solid foundations (principles and methodologies) for the elaboration of the FIDIS conceptualisation.

2.2.1.2 The categorisation

The categorisation in FIDIS has consisted in some process of aggregation of identity terms that had been identified (see "4.2 The Categorisation" for the result of categorisation). The idea was to be able to manipulate the different terms in a way that would brought a little bit more sense than would do an alphabetical organisation. For instance this categorisation helps distinguishing descriptive perspective of identity (presented in "4.2.1 Identity (as a person characterisation)"), a processes perspective of identity (presented in "4.2.2 Identification"), a tool perspective (see "4.2.3 Identity Management") or domains of application (see "4.2.4 Application (areas) and context"). This process of categorisation was mainly empirical, and did not rely of well defined methods of categorising a domain. This approach had the advantage of simplicity and did not require some sophisticate approaches or theories that are not really available in the context of this work. Automated clustering techniques represent a good illustration of more sophisticated mechanism of classification. They can be used to automate the acquisition of taxonomies or concept hierarchies from a text corpus (Cimiano, Hotho, and Staab, 2005) or the construction of Ontologies (Bisson, N´edellec, and Canamero, 2000).

Without having to go to the extreme of automated categorisation, we are aware that the use of more rigorous categorisations methods in FIDIS would be desirable, and will be investigated in the future.

2.2.1.3 The definition of the semantic

The definition of the semantic that was adopted in FIDIS and consisted in the very traditional definition of the term that is traditionally used inside WIKIs.

It has to be mentioned that the definitions of the terms have tried to conform to some templates (more details can be found "4.1.2.1 Model template") that provided some way to systematise and homogenise the representation. For instance, a term (entry in the WIKI) was seen as an object possessing attributes such as: textual definition related terms, issues, application domains, etc. (different categories of term objects exist).

Practically, this structure will only be enforced by guidelines and good practices, since these different terms will only be entered using text description in the WIKI system.

2.2.1.4 The illustration (Situating the concept in concrete contexts)

"Exemples or illustrations" were / are encouraged for each WIKI entry. For instance, the term anonymity may be illustrated by a set of examples of situations involving anonymity, and helping to illustrate the different anonymity issues.

Practically, this illustration may appear as an extension of the semantic definition that we have mentioned previously, and will only consist in adding an "example attribute".

Note: the cases stories and scenario of Del 2.2 represent typical examples that could be reference in the future in this section.

2.2.1.5 Formalisation

This step consists in the more explicit definition of the templates that we have mentioned, and aim at homogenising the definitions of the different terms and their relation with one another (and for instance, identifying the conceptual relation between the different terms). Beside some processes of consolidation and "cleaning" consisting in rewriting the description of the concepts can be done (this is particularly relevant if several authors have contributed independently to description of a single term).

2.2.2 The Organisational dimension in FIDIS

We have indicated in "2.1.4 The Organisational dimension" the different issues related to the contribution of a community in the population of terms.

Not surprisingly, the behaviour of the FIDIS community does not make an exception about these aspects. The contribution of the members of the FIDIS community has been unequal, and not up to the level that would be sufficient to generate in a sustainable way a high quality and lively shared Ontology adopted and used by everyone.

Practically, in the first version, we have observed a variety of behaviours related to the level of participation. For example the more "teckies" and the more junior people were the more important contributors. The more senior and knowledgable persons appeared to be the less likely to contribute (this is only true as a general trend). Several reasons were probably at the origin of these behaviours. The first reason is that the more senior people are the less inclined to adopt new practices (like WIKIs), since they consider their existing practices very effective, and have no reasons to change. The second reason is that they are the ones that may perceive the less benefit from an exchange process that they consider as unbalanced. Being much experimented, they are the more busy, but also less likely to benefit from the contribution of others. In other words in a open exchange, they may perceive they have more to lose (for instance their time) than to gain, and besides they have already have established efficient knowledge circuit (typically direct ones with other experts).

Several measures have been initiated to reverse this situation, and increase user participation. The first one is technical and ergonomical, and was aimed at reducing the cost of using the platform (in particular the time necessary to browse the content or to author content). Thus faster hardware was put in place, and the navigation and the esthetical aspects were improved.

A second measure was to increase the transparency of the system and in particular: to make visible the contribution of the different FIDIS members so that active authors would be

recognised for their efforts and their activities. Practically, a WIKI page logging the different actions is available making very visible what are the different contributions is available to all the visitors. Besides, each author of a contribution is invited to "sign" the entries in which they have contributed. Other measures are being considered, such as some action aiming at populating the content with enough material in order to reach a critical mass.

This organisational dimension is a very difficult one, and maybe this is the most important factors to make this WIKI a success.

We are very aware of the question, and we are working at it, and each new iteration of this WIKI is aimed at overcoming these difficulties, so as to reach as stage in which the WIKI will have become sustainable, and will not need exogenous stimulation (such as infusion of resources, effort and attention).

3 The Concepts of Identity and Identification

The objective of this section is to present to the reader an overview of the different issues and challenges of this domain (and indeed to express the importance of understanding this field), and then to present in more detail two of the related dimensions: The Identity dimension and the Identification dimension.

These two dimensions are defined as:

- The Identity dimension: set of characteristics representing a person
- The Identification dimension: set of terms, concepts, and mechanisms that relate to the disclosure of this identity information and the usage of this information.

This distinction between the two dimensions has been introduced to distinguish two different (and complementary) perspectives:

- A **descriptive perspective** referring to the representation of a person or thing in terms of a set of relevant attributes (third person perspective; objectification). In this case, the conceptualization of identity is done via the specification a set of attributes and associated states describing the characteristics of object (persons, groups, organizations) having an identity.
- A **process perspective** referring to the identification of a person or thing by uniquely differentiating him/her/it from all other persons and/or things (third person perspective; objectification), and by the usage of this information. In this case, this conceptualization of identity is considered in the context of identity related processes in which objects (persons, groups, organizations) possessing some identity are engaged, such as: the disclosure of identity information (authentication, profiling, etc.), and how this information is to be used (to give access to resources, to monitor and notarise the behaviours, etc.).

The following section provides more detailed, formal and structured definitions of the different terms and concepts used in the identity domain.

It is important to stress from the beginning that the terms Identity and Identification refer to two different concepts which are related but must not be confused. On the one hand, Identity is used to refer to a set of explicit relevant attributes (permanent or temporary) of a person in the context of practical activities. For instance, attributes related to the competency of a person will intervene in the working context, in a scenario in which competency represents an important factor of success in the accomplishment of a goal. On the other hand, Identification refers to the process used to link a person with an identity. Some criteria can be used for this purpose such as: the name of a person, her fingerprints, her genetic characteristics and her behavioural patterns. This reduction of attributes is necessary to produce easy, efficient and effective means to allow people access to restricted information or to find a specific person for reasons of security or commerce, for example.

Another distinction that should not be overlooked refers to the reductive characterisation of a person and to the identity of a living person (Hildebrandt 2005; Ricoeur 1992). The first, referred to as the idem-identity is static even if it is regularly upgraded, and is the only one

explicitly formalised and manipulated by information and identification technologies. The second one, referred to as the ipse-identity and representing who the person really is (from a philosophical point of view), is fundamentally fluid and indeterminate, and is out of the reach of the information and identification technologies.

3.1 The Identity and Identification Issues

Why is it important to understand the concept of identity?

3.1.1 Issues Concerning Many Facets of People's Life

Identity is not a topic only reserved to a small group of specialists. It intervenes very concretely in many facets of people's lives: their geographical mobility (dealing with the crossing of territories); their private life (dealing with their hobbies, romance, etc.); their family life (dealing with their marital status, their family structure); their social life (dealing with their friends, and their affiliation to groups); their work life (dealing with role, position, responsibility) and the way they conduct business activities (dealing with contracting, reputation, …); their life as a citizen (dealing with voting, and participation in communal life); their biological life (dealing with healthcare); their life as a customer (dealing with shopping and transaction); etc. Figure 1 helps to illustrate this multiple nature.

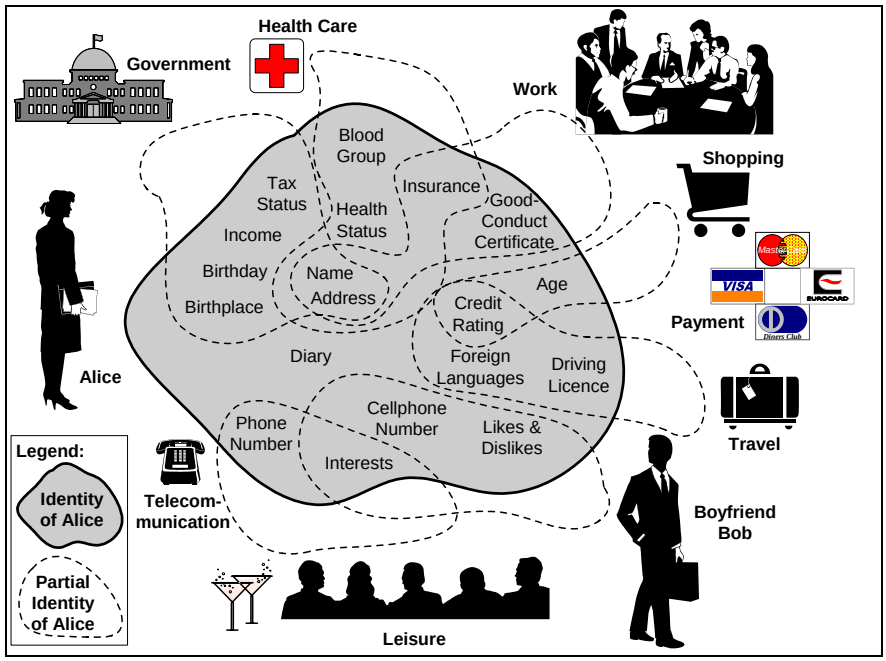

Figure 1 An example of the multiple facets of identity
(schema: Alice's Partial Identities (Clauß and Köhntopp, 2001), with the permission of Marit Hansen)

In practice, in each of these different portions of life, identity and identification issues can occur, and take different forms.

First, identity and identification issues can relate to the legitimacy of acting because of the affiliation to a particular group (country, company, social group) or given the prerogatives

(authority, right, etc.) attached to a particular accreditation (role in an organisation, diploma, recognised competence, bank account, etc.). For instance, citizenship can give you access to some social benefits or the right to travel and work in another country; a diploma or other such proof of competence can allow you to apply for a job position and later to exercise this profession; friendship opens up the possibility of asking for "and obtaining" free service from another person (the friend). Consequently, as individuals take on many different roles in the course of their life, different sets of characteristics, corresponding to these different roles, are used to represent their identity. Each of these "partial identity" includes both inherited "timeless" characteristics (such as nationality, gender, etc) and characteristics that they have acquired during their life (such diploma, competences, etc), or that they have been assigned or issued to fulfil this role (such as a position, some sort of authority, etc).

Another dimension is related to effectively proving (with different levels of reliability) that a person has indeed the affiliation or accreditation that they claim and that is required for the action. Examples of such elements can include an ID (passport, or business card), a key (proving to a technical infrastructure the right to access), a "parchment" (diploma), a social or competency clue (reflected in the attire or in the conversation), or a recommendation (for instance from an acquaintance).

Other aspects are related to the (partial) access of this identity information by others, their usage of this information and the question of the control (see for instance (Claessens *et al.*, 2003) for some discussions on anonymity control). The management of access to the information and of the control (by the person, by institutional bodies, by organisations, by commercial entities) is critical since it relates to the liberty of action of a person. For instance, the disclosure of information about the political opinion of a person (this person can be an activist or a Unionist) can seriously impact on the degrees of liberty of action of this person (in "the worse case" the person may be sent to prison, in other cases it may put the continued employment of this person in jeopardy). In particular, making the information too transparent can cause people to not act at all for fear of retaliation (from other people, from groups or from society). This can have negative consequences (people may fear denouncing unacceptable situations) or positive ones (preventing people from hiding revenues and paying less taxes or making people liable for a damage that they are responsible for). A more mundane aspect relates to the shameless exploitation of this information by third parties who consider it as a public resource. Spamming (direct marketing of mass emailing) represents one of the most irritating consequences of this.

3.1.2 Identity, Identification and the Information Society

Identity concepts and issues existed well before the advent of the information society. For instance, even before information technologies were available, people have had their privacy threatened, had to manage their participation in social circles, had to deal with the concept of national identity to cross frontiers, or had their profile used (sold, exploited, etc) by direct-marketing companies.

Still, the digitalisation of society, by augmenting the possibilities of Identity mechanisms and by opening new (digital) territories (going well beyond the imagination of Science-fiction novelists or even the wildest dreams of governments wishing to have a tighter control on their citizens), has considerably increased the importance of these topics and the associated risks and benefits. For example, research has been conducted on the implantation of RFID chips under people's skin, blurring the frontier between the real and the digital territories (Beslay

and Hakala, 2005), and a system for automated systematic facial recognition has been installed in the streets of Tampa, Florida. With sometimes a harsh return to reality: some people are strongly against the use of identification tags for humans (Michael Kanellos, 2004); and the face recognition system set in place in Tampa has thus far failed to identify one single crook or pervert listed in the department's photographic database, while falsely identifying 'a large number' of innocent citizens (Greene, 2002).

Currently more relevant to a large number of people is the variety of nuisances (spam, virus, spyware, phishing) which has invaded people's digital life, entering their mailboxes, destroying files, spying on online activities and trying to mislead them. In some cases this has even caused people to question the value that they are really getting out of these new digital territories (O'Brien and Hansell, 2004).

The information society has not invented many of the issues related to Identity that exist now, but it has significantly increased their importance. This change requires reassessment of the benefits and costs related to Identity issues, and for a renewal of the solutions which address them.

A high priority during this analysis is clearly the protection of freedom (free speech without fear of retaliation, privacy protection of opinion or medical information), enforcement of responsibility (liability and responsibility of your actions), human and society capability enhancing (the digital mechanisms help to reduce coordination and transaction costs, as well as to leverage the value of the social process).

3.2 The (self-) Identity concept

The notion of Identity is related to the characterisation and representation of a person (physical or moral) or of a group, and is concerned with the structure of this characterisation. For instance, Identity can be categorised according to different facets such as the personal Identity (personal), the biologic Identity (DNA), social Identity (membership), or the legal Identity and articulates them with their usage in different situations (such as leisure activities, transaction, work or social interaction).

The concept of Identity can be applied to a physical, a moral or an abstract person (such as an organisation or group). Notably, many different possible categorisations of identity information exist.

3.2.1 The I, the Implicit Me, and the Explicit Me

Without having to go too deep into the psychological realm, it may be useful to make a rudimentary distinction between:

- The I
  the indeterminate first person perspective
- The implicit me
  how a person perceives herself
- The explicit me
  how this person is perceived and represented (what is the image that this person provides to her environment).

These aspects establish the link between the living person, and her relation to the external environment (the explicit me), the two being modulated by the (un)conscious perceptions a person has of herself (the implicit me) (Rost, 2003).

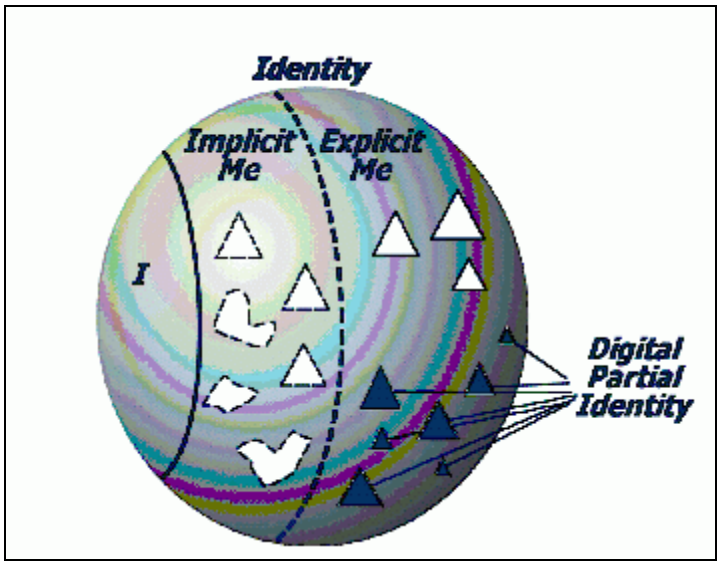

Figure 2: The I, the Implicit Me and the Explicit Me
(schema from (ICPP, 2003), with the permission of Marit Hansen)

This categorisation is important because it helps to raise and address several issues:

Acknowledging and addressing the Imperfection of the representation:

Firstly it demonstrates that the access and representation of a person is only imperfect (incorrect) and partial since it is always a reduction of the person to objectifiable attributes. As mentioned before, much care should be taken to acknowledge this. Indeed, conflicts and problems typically arise in the case of dissonance between the way a person perceives herself to be and the identity attributed to her. In real world situations, addressing this issue not always means just questioning the correctness of the information and providing some mechanisms for assessment, adjustment and correction, but also acknowledging that the objectified identity is never congruent with the living person. As to the correction of redundant or false information, European law imposes holders of personal information databases to explicitly provide some mechanisms allowing a person to rectify incorrect information.

The question of the Control:

The second issue is related to the control of this information: a person really only controls a limited part of her identity information. A large part of this information is externally controlled: by governments or institutions (the tax office, the healthcare organisms), by companies (for instance by the company employing this individual or by her bank), by commercial entities (such as marketing firms), or by the “public opinion” (such as newspapers or informal networks). Finding better ways to restrict an external entity from storing, manipulating and exploiting personal information may help address this issue. For instance, some mechanisms (legal, technical, etc.) can be used to enforce good practices when an entity

(governmental, commercial, …) manages personal information, such as defining what kind of information a certain category of entity is entitled to store (for example, companies may be forbidden to store medical information), what kind of operation can be conducted on the personal data file (for example, the police may be forbidden to access medical information), and how this information can be exploited (the commerce of some customer lists could be forbidden). Diverse (legal, technical, educational, etc) mechanisms (or a combination of mechanisms) could be used for this purpose. In the domain of law, it is important to note that the US and Europe have adopted different approaches on this issue, the US leaving the regulation of such matters, to a large extent, to private business enterprises (developing codes of conduct, good practices etc.), while Europe has tried to legislate on this issue (Agre and Rotenberg, 2001; Lessig, 1999).

3.2.2 True Identity, Assigned Identity, Abstracted Identity

A second categorisation of identity is as presented in the Three Tiers of Identity (Durand, 2002).

In this model, Andre Durand distinguishes three categories (or *tiers*) of identities:

- T1: The personal identity. (the inner and timeless identity)
  This is the true personal identity that is owned and controlled entirely by the person.
- T2: The corporate identity. (the assigned identity)
  This identity relates to a particular context (for instance a business relationship) and represents a temporary assigned or issued characteristic for the person such as: a job title, phone number, etc.
- T3: The marketing identity. (the abstracted or aggregated identity)
  This identity is more diffuse, and corresponds to some result of profiling. The person is not really considered as an individual (this person does not have a name), but as the result of filtering performed on a given set of characteristics. An example could be: "the customer belonging to the 'upper-level' social category, middle-aged, having a car more than three years old, playing golf, and living in one of the cities on the East-coast", who is contacted by a salesperson.

While this model may appear too simplistic to capture all the complexity of the Identity concept, it introduces several proprieties to Identity: its temporality, its conditionality, and its concreteness.

What is the impact of this model on the way we capture the Identity related issues?

Temporality & **Conditionality**

The Personal Identity represents an inherent property of the person and is both timeless and unconditional. The Corporate Identity is, on the contrary, conditional and temporary, and exists in a given context. This later identity can also be considered as attached to a person, rather than being part of the person. These concepts have some similarity with the Ipse and Idem identity of Paul Ricoeur, mentioned previously.

These properties of temporality and conditionality are important in the context of the management of the Identity because it allows a distinction between two facets that may be managed differently.

The first one is very important to the person and should therefore be controlled as much as possible by the person herself (or by very trustworthy third entities) and strongly protected.

Indeed threats on this "pervasive" identity (it intervenes in the many facets of life of a person) will have some more serious consequences for this person since it can potentially impact many parts of her life and for a long time. For instance, the thief of personal identity (done in the purpose of conducting illegal actions) has some impact in the reputation of the victim, who may suffer some consequences on her work (forbidding access to some jobs), her social life (isolating the person in society) or her personal life (destroying trust inside the family or in the circle of friends).

The second Identity is more linked to the role of the person in a given situation and can be more subject to control by a third party. Besides, the critical aspect of protecting the individual with the management of this identity may be oriented towards transparency and accountability rather than the privacy dimension. This could be relevant for mitigating the responsibility of the individual, for instance in the case of actions done as a representative of an organisation, and for isolating the representation of this identity in a specific area.

Concreteness

It is also interesting to note that an identity may not have a formal existence, and can, in particular, be abstract. For instance, the marketing identity does not explicitly represent the identity of an individual person, but an abstraction to which the person can *a posteriori* identify herself or be identified. Another abstract identity relates to the group or organisational identification: a person belongs to a group or an organisation not because of some formal and official status (explicit affiliation or contract), but via an implicit identification. A person believes she is part of a group or an organisation because she shares the same (assumed) attributes that characterises this group or this organisation (Dutton, Dukerich, & Harquail, 1994), or via a process whereby an individual's beliefs about an organisation become self-referential or self-defining" (Pratt, 1998: 175).

The **abstract nature** of this identity (marketing or organisational) does not prevent some **very concrete consequences** in the real life of the person: First, by becoming the target of direct-marketing campaigns (spamming) or psychological manipulation (advertising). Second, because this profiling (extraction of identity and categorisation) may reinforce the social structural rigidity, and may prevent people from gaining access to some resources (such as getting a loan to buy a house, or accessing jobs of high social status) because of belonging to some social categories. The management of identity should therefore be careful and put limits (given the performance of the technology, such as data-mining for profiling) on the uses that do not contribute to the well being of the person. On the more theoretical side, it may also support the transition between the social statuses of identities (Korotov, 2004).

Another emerging consideration is the possibility given by technology to "concretise" this implicit identity, with the advent of a whole range of applications enabled by technology. For instance, social computing services (Li, 2004) that explicitly represent and exploit the social network of a person are now proposed to help manage identity information that until now was only implicit and hidden. This is not without raising some serious new issues, such as the invasion of the "social private life" (Kahney, 2004) that identity systems will have to address, or the risks associated to a wrong perception or the real and substantive social position identity, and the biased social identity projected via the new information media (blogs, social networks, personal web pages). For instance, in the later case, this may mean displaying an "arranged identity" not really reflecting the reality, even unconsciously (for instance, people tend to identify themselves with organisations or groups with high social status or socially desirable features).

3.2.3 Virtual Person

So far, we have seen different aspects of persons and identities which however do not address some problems. In order to motivate this section, consider the following simple examples:

"Who closed the door?"
A simple question an observer might ask, which will usually be answered by a name. Yet there is also the possibility that something else is the actor in this action, for example the wind, a cat, some robot, etc.

"Who's the administrator of this website?"
Usually, there is a human administrator, say John Doe, which can be named, yet often, there is a whole set of persons which "hide" behind the term of administrator, eventually there might even be a computer program in charge, consider for example a chat room on the Internet where a program is ejecting people depending on some "bad" key words.

"Who reads this email?"
A typical situation, you sent an email to some address and the email may be read by some computer program replying to you that John Doe is out of office. Or John's secretary will read his mail.

What is in common in these examples? The actual observer does not know the very person which is being characterised; he might in fact in some cases even not be interested in knowing him or possibly even never really be able to know. Yet the three examples show that each question is a characterization of "something", maybe a physical person, a set of persons, a program, an animal, etc. This characterization is typically done by one of the four characteristics (cf. sect. 3.3.1.23.3.3.2). But in these examples, we cannot speak of identities neither of "explicit me" as for example the wind or several people might "hide" behind the mask.

So consider a new concept, the so-called virtual person. A virtual person might be perceived as a mask hiding what really is behind.[15] In the first example, the corresponding virtual person is characterised by the very sentence "Who closed the door?", hence the mask is defined by what it does. In the second example, one may think of characterizing the administrator by the persons (or possibly programs) knowing the secret password for accessing the system. So here, the virtual person is defined by something it knows.

In both cases, the virtual person is masking what is behind. In a concrete situation of an identification one might clearly be interested in making a link to some identifiable entity behind the mask. So for instance if closing or slamming the door has made some damage, there will be a process of identification being done for having someone pay the damage.

Going further, like in the case of connecting physical persons and identities, we can then speak of a virtual identity of a virtual person, which is – in some sort – an extension of the concept of digital partial identity (cf. section 3.2.1, Figure 2) modelling the "explicit us" rather than the "explicit me" (potentially different things/persons hidden by the mask).

---

15 This text is intendedly informal; for formal definitions of Virtual Persons, Subjects, etc. see Deliverable 2.2.

This concept of virtual persons helps also to clarify terms like pseudonyms, identification, etc. (cf. Deliverable 2.2).

3.2.4 Identities and Territories (Identity in Context)

Another approach of presenting the Identity field considers the concepts of "context" and "territory", which help structure the identity of a person. Territories (or spheres) represent classes of situations which pertain to a certain number of identity issues and involve only a particular subset of the characteristics of the person. In this case, the identity of an individual is no longer considered unique, but is rather constituted as multiple smaller identities that are activated in the context of each situation (see Figure 3).

For instance, a territory that corresponds to the social context of a person will shape a particular identity of this person (her social identity), involving a particular subset of the characteristics of this person (e.g., her social network) and will only concern some particular identity issues (e.g., reputation). The transaction territory considers this person according to a customer perspective, and is only concerned by the person's characteristics that are related to the conduct of a set of transactions (such as buying capacity, customer lifestyle, etc.) and associated identity issues (e.g., privacy).

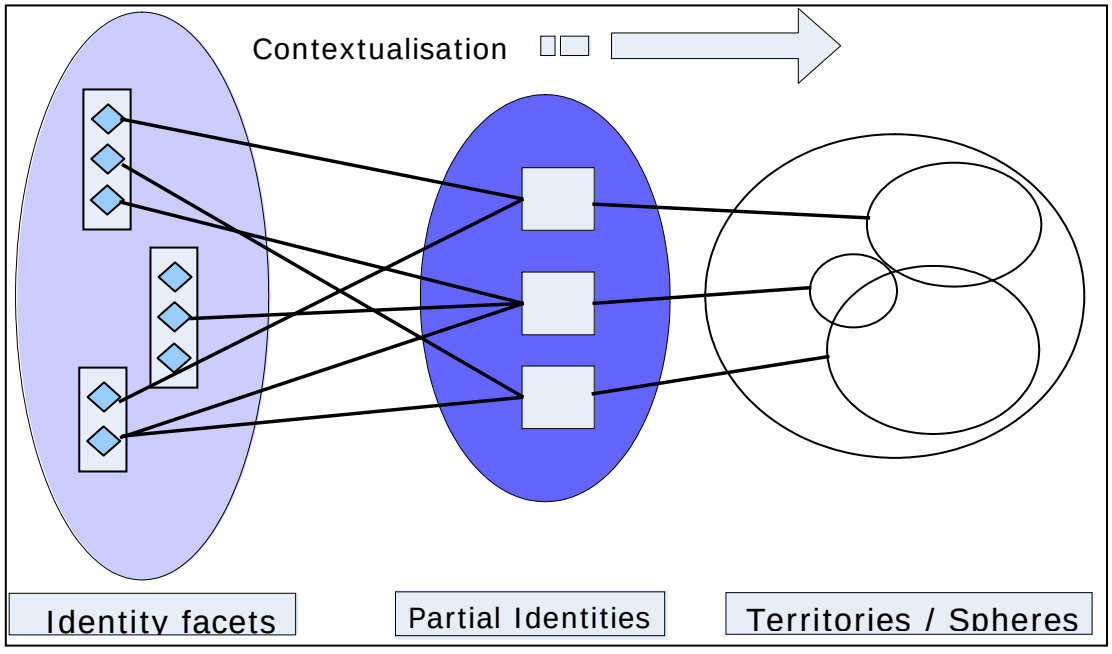

Figure 3: Mapping the Identity to the Territories

These different facets and spheres actually overlap with each other, raising a certain number of issues such as: (1) the risks originating from undesired leaking of information of one facet to another (e.g., from private spheres to public spheres); (2) the right of the person to isolate one sphere from another sphere (for instance the family, the work or the transaction sphere) in order to not make them too dependant on one another, preserving some space of freedom for the individual and limiting some forms of the 'domino effect'[16].

3.2.4.1 Mapping the Characteristics of the person into the Territories

The following example is an indication as to how the characteristics of a person can be mapped into a set of scenarios and, more generally, how the identity of this person is defined according to the context into which she is immerged:

- Personal identity (personal sphere)
  (subject, individual, person, patient, creature)
  - Name, date and location of birth
  - Location / geographical
    permanent location or instant location (mobile phone)
  - Biological and Health identity
    gender, DNA, medical information, alimentation style
  - Intimate identity (psychological)
    Aspiration, agenda, personality, attitude, motivation, cognitive style, …
  - Habits, style (dressing, …)
- Social identity (social sphere)
  social life
  - Family
  - Affiliation
    membership, projected image,
  - Friends, acquaintances, neighbours
- (Leisure) identity
  - Interest
  - Activities
    Hobbies, sport,
  - Role playing (persona -"playing a role in a mask"-, avatar)
- Organisational identity (work sphere)
  - Employee identity (work and organisational identity)
    Roles, position, formal authority.
  - Informal roles
    Leadership, power
  - Business identity
    reputation
- Citizen or townsman identity (society sphere)
  - Citizenship, nationality
    voting, becoming a representative, …

[16] Domino effect (Chaos concept): Mechanism, by which a single effect is propagated into a chain of other effects, resulting in dramatic consequences.

  - Legal identity
    police records, contracts, …
  - Political preference & orientation
  - Military information
- Customer or client identity (the transaction sphere)
  - Properties owned
  - Social category, life style
  - Buying capacity (real or perceived)
- Learning identity (the personal development sphere)
  - Knowledge capital (competency, …)
  - Certification (diploma, accreditation, etc.)
  - Aspirations
  - Learning style

3.3 The Identification Concept

In the previous section ("3.2 The (self-) Identity concept") we have presented the descriptive perspective of Identity, i.e. a conceptualisation relying on a set of distinguishing characteristics that an object owns and that forms its identity. In this section we are going to present a process perspective of Identity, i.e. processes in which this identity of an object is engaged and more specifically how the identity information is disclosed and used.

Identification concerns the set of approaches and mechanisms that intervene in the course of an interaction and which are very broadly related to the disclosure of Identity information (person characteristics and/or linking to a profile). For instance, it covers a variety of concepts such as anonymity (state of non disclosure of identity information), unlinkability (property of a system not to disclose information related to the relations that could exist between different items), or identifiers (information item that can be used to provide some level of authentication for an entity).

This identification is usually partial (disclosing only a partial identity), and is used in a specific context and for a specific purpose (granting access to a resource, delivering more personalised services, etc.).

Note: *Profiling* represents also a concept that can be related to identification. Profiling is defined as the process of constructing or applying a profile of an individual or a group. A profile consists of patterns of correlated data (Hildebrandt and Backhouse, 2005). The concept of profiling is specifically covered in workpackage 7 and will not be further described.

3.3.1 What is the identification used for

Identification intervenes in the following contexts:

- The access control to restricted resources or areas
- The exploitation of identity information
- The monitoring to enable accountability

3.3.1.1 Controlling access to restricted resources or areas (authentication)

One of the reasons for identifying a person is for controlling access to restricted resources or areas. This control comprises two different aspects: authentication and access management. Authentication relates to the verification of the identity of the person, and in particular ensuring that she is the person she claims to be. The management of the right of access relates to the role of this person in a context and the categories of operations granted to this person (for instance there may be some limitation related to the usage of a resource or to permitted operations in a particular area).

From an end-user perspective, identification may apply to another person or an organisation (moral person). Elements (factual, clues, etc) can also be used by a person to ensure that another person or organisation is indeed who she claims to be. For instance, an email address provides some level of identification related to the sender. The URL and the visual identity of a web site can also provide some means to authenticate an organisation.

3.3.1.2 Identification with the purpose of exploiting the identity information (knowing)

Another important reason for identifying a person is to allow access to relevant information that increases the impact of the interaction. The benefit of accessing this information concerns primarily the system accessing this information. It can, however, have some positive implications for the person herself, for instance by allowing a more customised and effective interaction between the system and this person. In other cases, the access to this information is a mandatory condition for the delivery of a particular service (examples include diagnostic services, or recommendation services that directly exploit this information).

Conversely, access to this information can also be harmful to the person. For instance, commercial companies (direct marketing) can exploit this information in order to better manipulate the target person. In other cases, this information provides the beneficiary of this information an unfair advantage over the person (advantage that can be exploited in a negotiation or in a job interview, for instance).

3.3.1.3 Monitoring and accountability

Monitoring and accountability relates to the ability to record and audit the actions of a person (and connect it to a partial identity).

This accountability can be used in a variety of contexts. In commerce, it can be used to help support different aspects of a transaction (payment, consumption, etc.). It can also help support some of the social aspects of electronic marketplaces, contributing, for instance, to the formation of the vendors or the customers reputation (for example, eBay utilises such mechanisms). Additionally, it can help monitor the general profile of the customer visiting a particular (web) site. In the domain of entertainment, monitoring may be used to record the downloading activities occurring in a peer-to-peer network (via the IP number of the computers involved in the exchange). In communication activity, this monitoring may consist of providing information (for instance login name, IP number …) that can be used to identify the author (or its virtual identity). In a security perspective, information may be logged in order to be used to identify suspect activities.

3.3.2 The risks associated to (incorrect or undesired) identification

It is important to mention that Identification (authenticating, knowing, monitoring) brings about a series of issues that can have some negative consequences for the person.

The problems can arise because of:

- incorrect identification
- undesired identification

3.3.2.1 Incorrect identification often relates to identity theft or identity fraud

Indeed one should be aware of the fact that the level of reliability of an identification is rarely absolute: for instance, login / password can be stolen, sender email address can be very easily forged (a practice often adopted by spammers), the visual identity of a web site can easily be imitated (for instance, phishing fraud consists of fooling the user by creating a copy of an official site), identity information spontaneously provided can be biased or obsolete. The consequences of incorrect identification can be serious, such as the disclosing of confidential information (for instance in the case of a break-in in a company information system), or the loss of important sums of money (e.g., if a phishing operation manages to convince a person to disclose her credit card information).

Note: The FIDIS workpackage "WP5 ID-Theft, privacy and security" specifically addresses the criminal aspects of Identity. Of particular interest for our concern is a task on ID fraud that is conducted in this workpackage which will comprise an ID fraud inventory.

3.3.2.2 The undesired identification directly relates to privacy issues

E-commerce, spyware and other similar mechanisms (such as tracing techniques employed by advertising companies) typically disclose information against the desire of the person and at her expense (this information can be used to manipulate the person and trigger a buying act).

In the workplace, undesired disclosure of information (for instance information about the affiliation to a union, a working practice, business contact, medical information) can severely harm the person and result in negative consequences (organisational pressure, job loss, etc.).

For the citizen, undesired disclosure of information (political opinions, expression of opinion) may have similar negative consequences.

In a later chapter of this document, we indicate mechanisms that can be used to reduce these risks and that achieve a better support for the authentication, the protection of the information, and the anonymity of a person.

3.3.3 The identification mechanisms

3.3.3.1 Explicit / implicit identification

Two different approaches to identification (of person characteristics and authentication) can be defined:

- The explicit identification
- The implicit (inferred) identification

Explicit identification relates to processes in which the person is aware, and even participates in this identification. This includes all of the explicit mechanisms that are used to authenticate a person such as: passwords, ID cards, biometric elements, business card, and presentation by another person (a social process). This also includes all the mechanisms that are used to collect explicitly the identity information (person characteristics) such as: questionnaires, ID cards (for the information they contain), etc.

Implicit identification relates to the processes that are used to authenticate the person and obtain the identity information without this person being aware. Implicit identification relies upon a series of available information (such as log files) from which the identity information is inferred or extracted. This can include identifiers attached to the person (such as RFID, IP number of the person, visual appearance, and social cues), or traces of characteristics (such as behaviour) that can be captured and analysed (for instance using profiling techniques such as data-mining).

As already indicated, not all of the identification mechanisms are equal and, in particular, the reliability varies considerably.

3.3.3.2 The different attributes used for the identification

The Identification processes (and more precisely authentication) can rely on different characteristics of a person such as:

- Something the person is
- Something the person does
- Something the person knows
- Something the person has

Identification can also be done via the use of a third (hopefully) trustworthy party, for example a certification authority.

The first identification mechanism (something the person is) relates to the characteristics directly associated to the user. For instance, identifiers based on biometrics represent a typical illustration.

The second identification mechanism (something the user does) relates to the behaviours that can be associated to the person, and can be considered as a particular case of the previous case. Characteristics can for instance refer to more implicit attributes such as the behavioural characteristics (behavioural patterns in a digital environment, attitudes in a social context) and that can be observed.

The next identification mechanism (something the user knows) relates to information that the user is supposed to know. This includes passwords, PIN or private information (e.g., the mother's maiden name).

The next identification mechanism (something the user has) relates to the possession of an artefact that is used specifically for the identification process. Examples include hardware-based tokens such as smart cards, software tokens such as digital certificates, or keys.

These different mechanisms differ in their usability (e.g., the access to some biometrics characteristics may be difficult) and their reliability (e.g., a key can easily be transferred or stolen).

3.3.4 Protecting from identification (and protecting privacy)

As previously indicated, identification is not always desirable (privacy has to be protected), and some mechanisms have to be proposed to help people protect themselves from undesired identification. This protection is becoming even more important since the advances in technology (such as profiling) can even have an impact on the state of democracy (Hildebrandt 2005). For instance profiling could be used via personalisation to manipulation people behaviour in a massive scale, reducing freedom of self-determination and personal autonomy, and therefore eroding societal freedom. In a doomsday scenario, personalisation services could put cultural and social diversity at stake: one political or religious message dominates the whole discourse (Hildebrandt and Backhouse, 2005).

Different concepts, means and mechanisms can be used to protect the privacy of a person (Hansen and Pfitzmann, 2004), which consist principally in obfuscating the identification process (hiding the user characteristics or traces, and thus making the authentication more difficult).

Examples of such concepts and mechanisms include:

- unlinkability
- unobservability
- encryption
- anonymity
- pseudonymity

Note: FIDIS Del 7.4 (Implications of profiling practices for democracy and rule of law) will focus on the implications of profiling on democracy and rule of law, integrating issues such as privacy and security but also posing the question: who is profiling who. This touches on issues such as equality (discrimination; dissymmetry) and transparency (invisibility of data processing; transparency of those that are profiled). The legal framework will be discussed in respect of building in checks and balances: facilitating opacity of individuals and transparency of data controllers/users.

3.3.4.1 The concept of Unlinkability

"Unlinkability of two or more items (e.g., subjects, messages, events, actions, ...) means that within this system, these items are no more and no less related than they are related concerning the a-priori knowledge." (Hansen and Pfitzmann, 2004)

This definition of Unlinkablitly is general, and deals with unlinkability of any sort of "items". (ISO, 1999) provides another definition that is more focussed on the user. It defines this concept as: "[Unlinkability] ensures that a user may make multiple uses of resources or services without others being able to link these uses together. [...] Unlinkability requires that users and/or subjects are unable to determine whether the same user caused certain specific operations in the system."

We can also differentiate an "absolute unlinkability" ("no determination of a link between uses") and "relative unlinkability" (i.e., "no change of knowledge about a link between uses").

Unlinkability of an item can in particular be partial, and "protect" only some operations associated to this item. For instance, unlinkability of an item can only concern the linking

with the originator of the item (such as the author of the message) or with the recipient of the item (the reader).

An example of an unlinkable item would be an anonymous message for which it is not possible to determine the identity of the author.

3.3.4.2 The concept of Unobservability

"Unobservability is the state of IOIs (the items of interest) being indistinguishable from any IOI at all." (Hansen and Pfitzmann, 2004)

(ISO, 1999) provides the following less general definition: "[Unobservability] ensures that a user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used. [...] Unobservability requires that users and/or subjects cannot determine whether an operation is being performed."

As seen before, our approach is less user-focused and thus more general. With the communication setting and the attacker model chosen in this text, our definition of unobservability shows the method by which it can be achieved: preventing distinguishability of IOIs. Thus, the ISO definition may be applied to different settings where attackers are prevented from observation by other means, e.g., by encapsulating the area of interest against third parties.

Unobservability is stronger than Unlinkability since it protects the content of an operation, and even its existence. Certainly, an unobservable item it unlinkable, since a precondition of linkability is the awareness of the existence of the item.

A similar concept is untraceability. The definition of the antonym is: "traceability is the possibility to trace communication between application components and as such acquire private information"; traceability is the ability to obtain information about the communicating parties by observing the communication context (e.g., through the IP address).

An example of an unobservable item message would be a secret message for which other parties cannot be aware of its existence, and *a fortiori*, of its content.

3.3.4.3 Concept and Mechanism: Anonymity

"Anonymity is the state of being not identifiable within a set of subjects, the anonymity set. The anonymity set is the set of all possible subjects[17]*. With respect to actors, the anonymity set consists of the subjects who might cause an action. With respect to addressees, the anonymity set consists of the subjects who might be addressed*." (Hansen and Pfitzmann, 2004)

Therefore, a sender may be anonymous only within a set of potential senders, his/her *sender anonymity set*, which itself may be a subset of all subjects worldwide who may send messages from time to time. The same is true for the recipient, who may be anonymous within a set of potential recipients, which form his/her *recipient anonymity set*. Both anonymity sets may be disjointed, be the same, or they may overlap. The anonymity sets may vary over time (and normally decrease, since we can make the assumption that digital systems do not "forget").

[17] I.e., the "usual suspects" :-) The set of possible subjects depends on the knowledge of the attacker. Thus, anonymity is relative with respect to the attacker.

It should be noted that this definition applies to any sort of subject, and not only to users.

(ISO, 1999) however provides a definition that only applies to a user: "[Anonymity] ensures that a user may use a resource or service without disclosing the user's identity. The requirements for anonymity provide protection of the user identity. Anonymity is not intended to protect the subject identity. [...] Anonymity requires that other users or subjects are unable to determine the identity of a user bound to a subject or operation."

Different levels of control of anonymity can be distinguished (Claessens *et al*., 2003):

- Unconditional anonymity (no revocation possible)
- User-controlled conditional anonymity
- Trustee-controlled conditional anonymity

In some applications it could be the case that the user wishes to revoke his anonymity. For example, in the case of a medical database, if a patient asks for his medical records he should be able to prove his identity.

In some other cases (trusted-controlled anonymity), the anonymity may be revocable by third parties under some specific conditions (e.g., in the context of fighting criminal activities).

An example of anonymity is a person connecting to a web site or using a peer-to-peer network: this web site, or external actors, do not normally have the ability (if we exclude, for instance, some spyware techniques) to determine the identity of the person. However, under certain conditions (e.g., in the context of fighting against piracy), a third party (a judge) may have the ability to disclose this information by linking the connection information (provided by the Internet Service Provider of the person) and the IP number trace of the person.

3.3.4.4 Mechanism: The use of Pseudonyms

Pseudonyms are identifiers of subjects, of sender and recipients (Hansen and Pfitzmann, 2004). The subject whom the pseudonym refers to is the *holder* of the pseudonym.

"Pseudonym" comes from Greek "pseudonumon" meaning "falsely named" (pseudo: false; onuma: name). Thus, it means a name other than the "real name". As the "real name" (written in ID papers issued by the State) is somewhat arbitrary (it can even be changed during one's lifetime), we will extend the term "pseudonym" to all identifiers, including all names or other bit strings. A pseudonym can be considered as a mapping of the identifier "real name" into another name. The "real name" may be understood as a pseudonym resulted from the neutral mapping. To avoid the connotation of "pseudo" = false, some authors call pseudonyms (as defined here) simply *nyms*. Although this is concise, in this document the usual wording, i.e. pseudonym, pseudonymity, etc, shall be adopted. However the reader may find nym, nymity, etc. in other texts.

On a fundamental level, pseudonyms are nothing more than another kind of attribute. But whereas in building IT systems, the designer can keep pseudonyms under his and/or the user's control; this is practically impossible with attributes in general. Therefore, it is useful to give this kind of system-controlled attribute a distinct name: pseudonym.

An example of a pseudonym is the name that a player chooses for himself when participating in an online game. Another example is the eBay vendors or buyers that want to transact without disclosing their identity.

Pseudonyms represent a particular indirect mechanism that helps isolating (and protecting) the identity of the person in the conduct of some activity. It should however be distinguished from anonymity in the sense that a pseudonym can have substance and persistence. For instance, some properties can be attached to a pseudonym such as reputation, which can be exploited by the owner of the pseudonym (e.g., for conducting some business, for developing social relationships or for playing).

3.3.4.5 Mechanism: Encryption

The final term is related to the protection of the item of information, in contrast to unobservability which was related to the protection of the exchange processes of this item. The circulation of an "encrypted" item may therefore be observable, and the sender and the recipient of the item may be visibly linked to each other. However, the nature of this link will remain unknown and private.

For instance, encrypting the content of an email (using the PGP[18] system) represents a typical example of a mechanism contributing to the protection of the identity of the person via content information hiding.

3.3.5 Other Mechanisms for managing the identification

The perspective presented has a relatively technical orientation, and is used primarily in the context in digital systems. The management of identification and of identity can also rely on more traditional approaches (that have to be adapted in order to take into account the context of digital systems). For instance, it can be legal: laws and rules can be elaborated to regulate and define the limit of what is permitted. It can also be educational: making people aware of the identity risks, and providing them (educating them) with good practices for improved management of their identity by themselves. This represents a very effective "tool", that is often overlooked (E&Y, 2004).

[18] PGP: Pretty Good Privacy is a public-key encryption application for exchanging files or messages with confidentiality and authentication.

4 An Inventory of Terms and some Categorisation

4.1 Introduction:

4.1.1 Using the Wiki

This section proposes an inventory of the different terms associated with the key Identity concepts that have been introduced in the previous section (Identity and Identity Management). This inventory is presented as a tree-like categorisation.

As indicated previously, this tree-like categorisation is not supposed to be perfect, but rather is a means to represent the different terms in a way that is easily understood by the user (including the non-expert), and easily further detailed. Importantly, the initial purpose of this inventory of terms is not to provide a final, formal and "approved" list or categorisation of "official" terms for FIDIS (this is something that can only be achieved at the end of the FIDIS project), but rather to create a structure that will be used to help the FIDIS network define a common vocabulary, and indeed to have a first tentative attempt at identifying the appropriate issues.

As the FIDIS "project" advances and the members develop a better shared-understanding of the Identity domain, this inventory of terms and the categorisation will progressively become enriched and refined by the FIDIS community.

The process of inventory and categorisation is conducted via the use of a Wiki of Identity terms, created specifically for this purpose.

A more in-depth description of the construction process has been detailed in a previous section. In this section, the initial results of this categorisation are presented. The categories presented in the next chapter represent a snapshot of the categories actually present in the Internal FIDIS Wiki for Identity definition, at the time of the writing of this document. This Wiki, which is currently only accessible by the members of the consortium, can be accessed at the following URL:

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

4.1.2 The definition of the terms

This Wiki has also started to collate, from the different members of the FIDIS consortium, the 'specification' for each of the terms that have been identified.

This specification consists of definitions, illustrations and references, as well as an indication of the different relationships connecting one term with another.

It should be noted that this specification provides a variety of perspectives for a better understanding of the meaning of the terms, and includes the information related to the contributors of this specification. This detail is important to note for several reasons: it provides some incentive for the people to contribute (acknowledging their work); it reinforces the perception of a collaborative nature of the specification and it provides additional means (social navigation) to connect this term with other related terms (via the information of the contributor).

4.1.2.1 Model template

The structure of the specification of the different term in the WIKI can follow a certain number of templates and guidelines. For instance each term can be specified using a certain number of facets that can be mandatory or not.

The facets that are available include:

- *definition*
  represents an abstract and very synthetic description of the term;
- *related terms*
  other terms that have some relationships such as synonymy;
- *Issues*
  present to list the different issues that can be associated to this term;
- *Application domain*
  application domain in which this term is the more likely to appear;
- *Examples*
  examples in which the term can intervene;
- References
  bibliographical reference relevant that can be used illustrate the understanding of a term;
- Authors
  Author having participated to the specification of this term
- Etc. (other facets will progressively be added)

4.1.2.2 Example

Example of the (preliminary) definition of the term Profiling:

Definition: *Profiling*

Online profiling is a complex concept that is subject to different definitions. It may mean the collection of anonymous transactional data that is used to create targeted advertisements; it may also mean the merger of clickstream data with personally identifiable information.

Related terms:

profile, data-mining, trace, abstracted identity,

Issues:

privacy, …

Application domain:

e-commerce, money laundering, …

Examples:

Privacy and Consumer Profiling, http://www.epic.org/privacy/profiling/

DoubleClick

Products:

CircBase® Prospect Plus Database includes active subscriber information from leading magazines willing to share their customers for non-competitive offers. Readership preferences provide a strong indicator of consumers' lifestyle, activities and interests.

http://www.experian.com/products/circbase_prospectplus.html

References:

CDT's profiling FAQ at http://www.cdt.org/privacy/guide/start/track.html

Greg Elmer; Profiling Machines (Mapping the Personal Information Economy); MIT Press, ISBN 0-262-05073-0, January 2004

In this book Greg Elmer brings the perspectives of cultural and media studies to the subject of consumer profiling and feedback technology in the digital economy. He examines the multiplicity of processes that monitor consumers and automatically collect, store, and cross-reference personal information.

http://mitpress.mit.edu/catalog/author/default.asp?aid=18786

Canhoto Ana Isabel and James Backhouse, (2004); Constructing categories, Constructing signs

- analysing differences in Suspicious Transaction Reporting practice; IS-CORE 2004 - 4th Annual SIG IS Cognitive Research Workshop; Submitted.

http://internal.fidis.net/154.0.html?&dir=/Papers&mountpoint=5

Contributors:

Lothar Fritsch (provided it in the WP7 forum)

Mireille Hildebrandt (provided it in the WP7 forum)

4.2 The Categorisation

An overall categorisation of the different Identity terms is:

- Identity (seen as the characterisation of the person)
  - Biological
  - Behavioural
  - Situational
  - Abstracted
- Identification
  - Identifier
  - Processes and properties
- Identity Management
  - Approach

    - Mechanisms for obfuscating the information
    - Mechanisms contributing to disclose the information
- Application (areas) and context
    - Application domain
    - Territory
    - Threat (associated to the contexts)
- Other

The following chapters will go into the details of this categorisation.

4.2.1 Identity (as a person characterisation)

General

- attribute,
- characteristic,
- hacker,
- identity (partial, digital, organisational, virtual, ...),
- organisation,
- person (real, moral, legal, ...),

Biological

- bio-Implant,
- gait/gesture (biometrics),
- gender,
- health, handwriting (biometrics),
- voice (biometrics),

Behavioural (psycho & socio)

- --- *psycho (internal identity)*
- aspiration, attitude,
- cognitive style,
- interest, ipse-identity, individual self,
- learning style,
- personality,
- --- *socio (external identity)*
- acquaintance, friendship, relationship,
- belonging,
- group, group identity,
- idem-identity,
- lifestyle,
- persona,
- social identity,

- trust,

Situational (assigned, issued, acquired)

- assigned identity,
- affiliation, avatar,
- citizenship, competency,
- membership, multiple identities,
- nationality,
- organisational identity,
- phone number, position,
- role,
- title,
- wealth,

Abstracted

- abstracted identity
- lifestyle category,
- profile (psychological, social, ...)
- reputation,
- social category,
- certificate/credential ownership
- attributes

4.2.2 Identification

General

- identification,
- privacy,

Identifiers

- access control, anonymity (relationship, ...), authorisation, authentication, ANI (Automatic Number Identification)
- certificate, certification authority, credential,
- key (private or public),
- identifier, ID, identifiability,
- linkability / unlinkability (absolute or relative),
- observability, unobservability,
- profiling, pseudonyms (person, role, relationship, role-relationship, transaction), pseudonymity,
- RFID,
- signature,

Anonymisers / Transparence

- logging, lure,
- proxy,
- relaying,
- social translucence,
- trace, traceable,
- application vs. communication anonymity
- anonymity metrics

4.2.3 Identity Management

Approach

- education means
  - good practices, codes of conduct, ...
- legal,
- police,
- security,
- technical mechanism (computer sciences, intelligent agent),

Mechanisms (hiding, restricting information)

- DRM (digital right management),
- encryption,
- PGP,
- Zero-knowledge protocols

Mechanisms (making the information more explicit)

- biometrics,
- datamining,
- forensic,
- sign-on,
- personality test

Others

- PKI,

4.2.4 Application (areas) and context

General

- application domain (health, commerce, government, ...)
- sphere,
- territory,

Application domain

- education,
  - distance education, training, ...
- banking,
  - banking operation, loan, credit card, payment,
- business,
  - contract
- commerce,
  - transaction
- government,
  - voting, social benefit, taxes
- healthcare,
- leisure (entertainment)
  - gaming, dating, music, email...
- work,
  - HRM (Human Resource Management), career management,
  - KM (Knowledge Management)

Territory (real or digital)

- CoP (Community of Practice),
- intercultural environment,
- homeland,
- workplace,
- --- 'digital'
- socialware,
  - blog, wiki, social network management system
- digital world,
- e-learning environment, e-marketplace (eBay), ...
- online game,
  - RPG (role playing game), MMORPG (Massive Multiplayer Online RPG), MUD (Multi-Users Dungeon),
- mixed environment,
- peer-to-peer, pervasive environment,
- virtual community, virtual reality, virtual world,

Threat

- big brother,
- loss of privacy
- loss of freedom of expression
- fear of retaliation
- credit card fraud,

- hijack,
- identity theft,
- malware (spyware, virus, worm, backdoor, ...), money laundering,
- phishing,
- scams, spam,

4.2.5 Others

This category collects all the terms relevant to the identity domain that have not yet been categorised.

- accountability, liability, attack, attacker, authentic / phony,
- communication, confidence, criminal,
- data protection, discipline (psychology, security, sociology, technology),
- ethical, entropy,
- information,
- misinformation,
- public, private,
- recipient, revocation, risk,
- secrete, secrecy, sender,
- transaction, transferability,
- uniqueness,
- validity,
- ...

4.3 Other categorisations

This chapter aims to be more concrete and present a categorisation that corresponds to the structure of attributes associated to the identity of a person.

Contrary to the previous categorisation which aims at defining the highest level concept for a better understanding of the meaning of the terms, this structure is to be used more as a formal representation that may be used in information systems.

In this document, the categorisation is given just as an indication, since it will be better addressed in a later deliverable (D2.3 models of identity). Besides, other aspects are also already being investigated by special groups and standards (such as HR-XML).

4.3.1 A (partial) definition of profile of the person

- profile representation (person characteristics)
    - individual profile
        - location
            - home location
            - work location
            - instance location (GPS/mobile)
        - biological
            - gender
            - eye colour

- height
- fingerprint

▪ job information
- organisation
- title
- role
- tasks
- salary

▪ career information
- title
- salary

▪ psychological profile
- 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

▪ sociological profile (affiliations)
- personal network (mates, friends, ...)
- family network
- professional network
- political network

▪ sociological profile (perception)
- reputation
- sociological category
- Team role (Belbin model)
  Plant, Resource investigator, Co-ordinator, Shaper, Monitor evaluator, Teamworker, Implementer, Completer

▪ preferences

▪ financial information
- banking information
- tax information
- incomes

o group profile

o organisation profile

4.3.2 Other representations

Similar (formal) structuring / categorisation / taxonomy could also be considered, for instance, for the representation of the identifier, the domain of application, the different authentication system used, categories of fraud, etc.

For example, a basic categorisation of the identifier concept is:

- identifier
  - o biometrics
    - DNA
    - finger print
    - retina

        - iris
        - face
        - gesture
        - …
    - electronic device
        - id card
        - RFID
        - …
    - digital id
        - login/password
        - token,
        - etc…

5 Conclusion and Future work

This document only represents 'the tip of the iceberg', and the first result of the work aiming at better understanding the identity concept and identifying the different terms associated with it.

Starting from the idea of creating yet another conceptualisation of the identity domain, we rapidly moved away this initial goal and set ourselves some more ambitious objectives. Indeed we wanted to make a more valuable contribution to the conceptualisation of the Identity domain and be able in particular to leverage the value of the Network of excellence in the identity domain that the FIDIS consortium represents. Besides, we also wanted to move away from the illusive objective of producing the perfect, official and complete reference of the terms that are present in the Identity domain (a probably impossible undertaking), in favour of a system less perfect but more able to capture and adapt quickly to the rapid evolution of the Identity concept. Thus, the first objective of this work would not be to produce the ultimate conceptualisation of the identity domain, but rather to produce a tool that would facilitate the circulation and the creation of new ideas related to Identity issues amongst the community of people interested by the subject.

5.1.1.1 What were we able to achieve?

The first part of this document is constituted of an approach and a vision defining the concept of collaborative Identity domain definition construction.

The approach adopted has consisted of adopting a Wiki-based approach for the collaborative collection of terms and definitions. The first version of this Wiki has been implemented, and a first round of data has been entered. Still some issues of dynamics of participations will need to be addressed if we want this WIKI to become sustainable.

The second part of this document provides a global overview of the Identity field, hopefully comprehensible by the non-expert, and while being very structured and synthetic, providing in some respect an informal categorisation of the field. Two different topics, Identity and Identification, were in particular covered. The content originated when possible from the existing work available from the FIDIS members.

The last part of this document represents a snapshot of the Wiki content and structure at the time of the writing this document, as well as some alternative categorisations (such as a tentative and partial representation of a user profile) which help to demonstrate the broadness and the complexity of the subject.

5.1.1.2 Where are we heading to?

As indicated, this deliverable should only be considered as the start of a continuous process that will help to identify, as they emerge, new Identity terms and concepts, define them (taking into account the different perspectives) and situate them (through the connection with other terms and the illustration with examples and cases), and all this in a collaborative and constructivist setting supported by the FIDIS Wiki of Identity.

Particular attention will be paid to stimulating the active participation of the maximum number of members in this process of knowledge construction, knowledge use and knowledge

exchange. The introduction of different social translucent mechanisms (Thomas, Kellogg and Erickson, 2001), such as visual tag helping to display the activity taking place in the Wiki space, and in particular finding a method (reputation) of rewarding the most active and valuable contributors.

Other deliverables, such as "Del 2.2 Set of Use Cases and Scenarios", provides more elaborated examples, cases, stories and scenarios, linking the different identity terms with one another; (2) "Del 2.3 Identity models", helps to continue this endeavour.

Besides, the new content will be generated in FIDIS (it has to be reminded that for this first version, little FIDIS content was available) will help in enriching the specification of the conceptualisation of the FIDIS domain. More specifically, the different specialised workpackages such as WP3 Identity Technologies, WP4 Interoperability, WP5 ID-Theft, WP6 Forensic, WP7 Profiling will provide some content from which will be extracted conceptualisation. In some cases these workpackage have even started their own classification that will need to be integrated (examples include "ID fraud inventory" started in WP4 or the Del. 7.2 "Inventory of actual profiling practices and techniques" to name only a few). In the other way round, these different workpackages will have to make use of the global conceptualisation that will have been elaborated.

Finally, the future direction will consist of consolidating the work by working on a more formal (but not rigid) semantic orientation, and investigating to which extent semantic web approaches and tools can be used for this purpose.

6 References

Aaai (2004); Brief History of Artificial Intelligence; AAAI AI Topics dynamic library of introductory information about Artificial Intelligence; web site: http://www.aaai.org/AITopics/bbhist.html

Agre Philip E. and Marc Rotenberg (ed.), Technology and Privacy: The New Landscape, Cambridge Massachussets, London England: MIT Press 1998

Berners-Lee, Tim; Hendler, James; Lassila, Ora. "The Semantic Web". Scientific American 284(5), pp. 34-43. May 2001 http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&ref=sciam

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert & S. T. Fiske & et al. (Eds.), The handbook of social psychology, Vol 2 (4th ed., pp. 193-281). New York, NY, US: McGraw-Hill.

Beslay Laurent and Hannu Hakala (2005); Digital Territory: Bubbles; Vision Book (2005); to be published; http://europa.eu.int/information_society/topics/research/visionbook/index_en.htm

Bisson G., C. N´edellec, and D. Canamero (2000); "Designing Clustering Methods for Ontology Building: the Mo'KWorkbench"; In Proceedings of the First Workshop on Ontology Learning (OL-2000) in conjunction with the Fourteenth European Conference on Artificial Intelligence (ECAI-2000), Berlin, 2000.

Blanchard A. and Markus L., (2002): Sense of Virtual Community-Maintaining the Experience of Belonging. Proceedings of the 35th HICSS Conference -Volume 8, Hawaii.

Canhoto Ana Isabel and James Backhouse, (2004); Constructing categories, Construing signs - analysing differences in Suspicious Transaction Reporting practice; IS-CORE 2004 - 4th Annual SIG IS Cognitive Research Workshop; Submitted.

Chan, C. et al. M., (2004). Recognition and Participation in a Virtual Community: A Case Study. Proceedings of the 37th HICSS Conference, Hawaii.

Cimiano, P., Hotho, A. and Staab, S. (2005); "Learning Concept Hierarchies from Text Corpora using Formal Concept Analysis"; Journal of Artifcial Intelligence Research 24 (2005)

Claessens J., C. Díaz, S. Nikova, B. De Win, C. Goemans, M. Loncke, V. Naessens, S. Seys, B. De Decker, J. Dumortier, and B. Preneel (2003); "Applications Requirements for Controlled Anonymity," APES Deliverable D7, 129 pages, 2003.

Clark, H.H., & Brennan, S.E. (1991). Grounding in communication. In L.B. Resnick, J. Levine, & S.D. Behrend (Eds.), Perspectives on socially shared cognition (pp. 127-149). Washington, DC: American Psychological Association.

Clauß Sebastian, Marit Köhntopp (2001); Identity Managements and Its Support of Multilateral Security; in: Computer Networks 37 (2001), Special Issue on Electronic Business Systems; Elsevier, North-Holland 2001; 205-219

Cooper Charles (2005); "Getting real about wikimania", CNET News.com, July 28, 2005 url: http://news.com.com/Getting+real+about+wikimania/2008-1082_3-5807538.html

Denny Michael (2002); "Ontology Building: A Survey of Editing Tools", XML.org, November 06, 2002, http://www.xml.com/pub/a/2002/11/06/ontologies.html

Durand Andre (2002); Three Tiers of Identity; Digital Identity World, March 16, 2002 http://www.digitalidworld.com/print.php?sid=26

Dutton J., J. Dukerich and C. Harquail (1994) Organizational Images and Member Identification. Administrative Science Quarterly, 39 (1994), pp 239-263.

E&Y (2004); The Ernst & Young Global Information Security Survey 2004; Ernst & Young, September 2004

FIDIS WP2 workshop, (2003); Taxonomy of Identity, Anonymity and Pseudonymity; First workshop of the Fidis WP2; Brussels, December 2-3, 2003

Greene Thomas C (2002); Face recognition technology a proven farce; The Register, 4th January 2002. http://www.theregister.co.uk/2002/01/04/face_recognition_technology_a_proven/

Gruber T. R.; A translation approach to portable ontologies; Knowledge Acquisition, 5(2):199-220, 1993 http://www-ksl.stanford.edu/kst/what-is-an-ontology.html

Guarino Nicola, Christopher A. Welty (2002); Evaluating ontological decisions with OntoClean. Communications of the ACM (CACM), Volume 45, 2002. pages: 61-65.

Hall, H., (2001): Social exchange for knowledge exchange. Paper presented at Managing knowledge: conversations and critiques, University of Leicester Management Centre, 10-11 April 2001.

Hansen Marit and Andreas Pfitzmann (2004); Anonymity, Unobservability, Pseudonymity, and Identity Management - A Proposal for Terminology; Working document. http://dud.inf.tu-dresden.de/Literatur_V1.shtml

Hildebrandt Mireille (2005); Privacy and Identity; in: Erik Claes and Antony Duff (ed.), *Privacy and the Criminal Law*, proceedings of the Conference on Privacy and the Criminal Law 14th-15th May 2004, to be published 2005

Hildebrandt Mireille James Backhouse eds. (2005); "Descriptive analysis and inventory of profiling practices"; FIDIS deliverable 7.2, June 2005

Holsapple, C.W. & Joshi, K.D. (2002) "A collaborative approach to ontology design". Communications of the ACM, 45(2), 42-47.

ICPP (2003), Independent Centre for Privacy Protection (ICPP) Schleswig-Holstein and Studio Notarile Genghini (SNG); Identity Management Systems (IMS): Identification and Comparison; study prepared under contract for Institute for Prospective Technological Studies, Joint Research Centre Seville, Spain, Sept. 2003; http://www.datenschutzzentrum.de/idmanage/study/ICPP_SNG_IMS-Study.pdf

ISO (1999); ISO IS 15408, 1999; http://www.commoncriteria.org/.

Jarvenpaa and Leidner (1999): Communication and trust in global virtual teams, Organisational Science, Vol. 10, Nov-Dec 1999, pp. 791-815

Jaquet-Chiffelle David-Olivier *et al.* (2004); Identity, Identifyability and Anonymity: concepts, definitions, applications; working paper, VIP (Virtual Identity and Privacy Research Center); to be made available at: http://www.vip.ch/papers/identity.pdf

Kahney Leander (2004); Social Nets Not Making Friends; Wired magazine, January. 28, 2004

Kanellos Michael (2004); RFID tags: The people say no; CNET News.com, September 7, 2004

Koh J. and Kim Y.-G., (2003): Sense of Virtual Community: A Conceptual Framework and Empirical Validation. International Journal of Electronic Commerce, Volume 8, Number 2, Winter 2003-4, pp. 75.

Korotov, Konstentin. (2004). "Neither Here not There" or "Both Here and There": Experiencing Liminality and Playing with Identity" Academy of Management Conference, New Orleans, August 6-11

Lave, J., Wenger, E. (1991). Situated learning: Legitimate peripheral participation. New York: Cambridge University Press.

Lessig Lawrence (1999), *Code and Other Laws of Cyberspace*, New York: Basic Books 1999

Li Charlene (2004); Profiles: The Real Value Of Social Networks, Forrester Research report, July 15, 2004

Missikoff M., R. Navigli, and P. Velardi (2002); The Usable Ontology: An Environment for Building and Assessing a Domain Ontology. Proceedings of the International Semantic Web Conference 2002, Springer, 2002, pp. 39-53.

Noy Natalya F. and Deborah L. McGuinness (2001); "Ontology Development 101: A Guide to Creating Your First Ontology" by, March, 2001, http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html

O'Brien Timothy L. and Saul Hansell (2004); Barbarians at the digital gate; The New York Times, September 19, 2004

Olson, Eric T. (2002), "Personal Identity", The Stanford Encyclopaedia of Philosophy (Fall 2002 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2002/entries/identity-personal

Pratt, M. G. (1998); To Be Or Not To Be: Central Questions in Organizational Identification; In D. A. Whetten, and Godfreey, Paul C. (Ed.), Identity in Organizations (pp. 171-207). Thousand Oaks: Sage.

Prieto-Diaz R (2002); "A faceted approach to building ontologies," presented at Information Reuse and Integration, 2003. IRI 2003. IEEE International Conference on, 2003.

PRIME (2004); Privacy and Identity Management for Europe Consortium: Framework V0, D14.0.a; 9 June, 2004;
http://www.prime-project.eu.org/public/prime_products/deliverables/pub_del_D14.0.a_ec_wp14.0_V4_final.pdf

Rand Matt (2004); “Extreme Blogging”; Forbes, december 2004.
http://www.forbes.com/best/2004/1213/bow001.html

Ricoeur Paul (1992), Oneself as another, Chicago and London: University of Chicago Press 1992

Rost Martin (2003), “An Introduction to the Concept of Identity”, Workshop on Taxonomy of Identity, Anonymity and Pseudonymity, FIDIS NoE – WP2 Kick-off meeting, Brussels, 2003-12-02. url: http://www.calt.insead.edu/fidis/workshop/workshop-wp2-december2003/

Thomas J. C., W. A. Kellogg, and T. Erickson (2001); The knowledge management puzzle: Human and social factors in knowledge management; IBM Systems Journal, Volume 40, Number 4, 2001

Tung, L., et al. (2001): An Empirical Investigation of Virtual Communities and Trust. Proceedings of the 22nd International Conference on Information Systems, 2001, pp. 307-320

Uschold, M., Gruninger, M., (1996), "Ontologies: principles, methods and applications", The Knowledge Engineering Review, Vol. 11, No. 2, pp. 93-136.

Wenger Etienne (1999). Communities of Practice: Learning, Meaning, and Identity Cambridge University Press, December 1999

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/ (to be later transferred at http://www.fidis.net/).