FIDIS

Future of Identity in the Information Society

Title:	"D16.3: Towards requirements for privacy-friendly identity management in eGovernment"
Author:	WP16
Editors:	J.C. Buitelaar (TILT, the Netherlands), M. Meints (ICPP, Germany), E. Kindt (ICRI, Belgium)
Reviewers:	J. Backhouse (LSE, United Kingdom), J. Vyskoc (VaF, Slovakia)
Identifier:	D16.3
Type:	Deliverable
Version:	1.1
Date:	Sunday, June 14th 2009
Status:	[Final]
Class:	[Public]
File:	2009_06_14_Fidis_D16.3_Reqs_PF_eGov_v1.2_final

Summary

This report describes in a multi-disciplinary way requirements for privacy-friendly identity management in eGovernment. The cooperation among the large number of disparate entities is compared with so-called 'circles of trust', whereby identity and service providers have to agree on procedures and conclude agreements, including on the allocation of their roles and responsibilities within the eGovernment context. The use of authoritative sources, the importance of an authorisation management and the authentication and assurance mechanisms are hereby further identified as basic legal approaches for privacy-friendly IMS. Basic technologies that support the fulfilment of these requirements are presented and discussed.

The deliverable also discusses various advanced technical approaches, which may prove valuable for eGovernment, in particular techniques for the management of identities in networking infrastructures. This includes Private Information Retrieval, DC networks and MIX networks. The BBox architecture which may provide a secure logging system under certain conditions is also described. Finally, an organisational framework for privacy policy handling is suggested in combination with technical approaches to support privacy policy handling. Various issues, however, are still open for further research.

The FIDIS NoE receives research funding from the Community's Sixth Framework Program

Copyright Notice:

PLEASE NOTE: This document may change without notice – Updated versions of this document can be found at the FIDIS NoE website at www.fidis.net.

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 (ICPP)	Germany
5. Institut Europeen D'Administration Des Affaires (INSEAD)	France
6. University of Reading	United Kingdom
7. Katholieke Universiteit Leuven	Belgium
8. Tilburg University[1]	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 (MU)	Czech Republic
14. VaF Bratislava	Slovakia
15. London School of Economics and Political Science (LSE)	United Kingdom
16. Budapest University of Technology and Economics (ISTRI)	Hungary
17. IBM Research GmbH	Switzerland
18. Centre Technique de la Gendarmerie Nationale (CTGN)	France
19. Netherlands Forensic Institute (NFI)[2]	Netherlands
20. Virtual Identity and Privacy Research Center (VIP)[3]	Switzerland
21. Europäisches Microsoft Innovations Center GmbH (EMIC)	Germany
22. Institute of Communication and Computer Systems (ICCS)	Greece
23. AXSionics AG	Switzerland
24. SIRRIX AG Security Technologies	Germany

[1] Legal name: Stichting Katholieke Universiteit Brabant
[2] Legal name: Ministerie Van Justitie
[3] Legal name: Berner Fachhochschule

Versions

Version	*Date*	*Description (Editor)*
0.1	26.09.2008	- Drafting of Table of Contents (ICCP, K.U.Leuven, KU)
0.2	19.10.2008	- Insertion of texts from D16.1 (TILT)
0.3	21.12.2008	- Texts from ICRI, ICPP, TUD, Freiburg included
0.4	18.02.2009	- Integration of final versions of the contributions
0.5	27.02.09	- Structure of the TOC revised, introduction, conclusions and glossary included
0.6	06.03.09	- Executive Summary added (ICRI) - Integrative editing throughout the deliverable (TILT) and especially in chapter 5 (ICPP)
0.7	10.04.09	- Comments reviewers processed. Due to discussion at workshop at Fidis General Meeting at Frankfurt, substantial restructuring and other changes carried out. (TILT)
0.8	24.04.09	- Additional editing, changes to TOC and integration (ICRI)
1.0	07.05.09	- Finalisation (TILT)

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)*
1 Executive Summary	Els Kindt (ICRI)
2 Introduction	Martin Meints (ICPP)
3 Requirements for privacy-compliant IdM in eGovernment	Brendan Van Alsenoy and Sigurd Vandebuerie (ICRI), Martin Meints (ICPP), Hans Buitelaar (TILT)
4 Basic Technical Approaches	Stefan Berthold and Stefan Köpsell (TUD), Martin Meints (ICPP)
5 Advanced Technical Approaches	Stefan Berthold and Stefan Köpsell (TUD), Maike Gilliot (ALU-Freiburg)
6 Combined Technical and Organisational Approaches	Martin Meints (ICPP), Simone Fischer-Huebner (KAU)
7 Summary and Conclusions	Hans Buitelaar (TILT), Brendan Van Alsenoy (ICRI)
8 Bibliography	Hans Buitelaar (TILT)
9 Annex: 9.1 Glossary and 9.2 Definitions	Brendan Van Alsenoy (ICRI), Hans Buitelaar (TILT), Martin Meints (ICPP), All

Table of Contents

1 Executive Summary

This report describes multi-disciplinary requirements for privacy-friendly identity management (IdM) in eGovernment. The cooperation amongst the large number of disparate entities is compared with what has been described in IdM literature as 'circles of trust', whereby identity and service providers agree to adhere to certain procedures and agreements, particularly with regard to identification and authentication mechanisms. Consideration is also given to the allocation of roles and responsibilities for the many additional tasks which require allocation within eGovernment. Particular attention is given to the use of authoritative sources, both as a means for ensuring data accuracy and as integral component of user- and access- management. The report further highlights the importance of authorization management, and the use of appropriate assurance levels and authentication mechanisms as basic legal requirements for privacy-compliant Identity Management Systems (IMS). Security and data handling policies, as well as the management of logs and auditing processes are similarly tied in with these requirements. Reference is also made to the i2010 eGovernment Action Plan, in which the European Commission considers the main purpose of electronic identification for public services as easing access and offering personalised and smarter services.

Besides the general technical requirements which have been described in Fidis D.16.1, the present document reminds the readers of the basic technical fundamentals of identity management, before analysing in the following pages more advanced technical approaches. The advanced technical approaches discussed refer to various techniques for the management of identities in network infrastructure on the one hand, and techniques for secure logging on the other. Reference is made to identity obfuscation techniques, with a focus on reaching a degree of anonymity on connection level rather than on content level. Various techniques, such as Private Information Retrieval, DC networks and MIX networks are elaborated, discussing their advantages and flaws. Private Information Retrieval and DC networks offer a very high level of protection, but require many resources in terms of computational power and bandwidth. Proxies on the other hand require that the operator of the proxy is trusted completely by its users. Some of these systems, however, especially MIX-based systems, are still under development and have not yet achieved a quality of service level appropriate for use in large scale applications. As to secure logging, it is known that log data are vulnerable to various attacks, leading to a potential loss of integrity and authenticity. The so-called 'BBox' architecture, which may provide a secure logging system under certain conditions, is described in the chapter on advanced technical approaches.

In the last chapter, an outline of an organisational framework for privacy policy handling is suggested in combination with technical approaches to support the handling of privacy policies. In particular, the tasks of the organisational process are discussed, such as the initialisation, the implementation and the checking of the privacy policy and the technologies to enhance privacy in eGovernment. The semi-automated support for privacy policy handling, such as P3P and the architecture introduced and developed in the EU Prime project are herein further elaborated as possible tools for privacy-friendly eGovernment.

This deliverable describes, from a legal, technical and organizational point of view, which privacy preserving measures can be applied in the context of governmental IMS. It indicates which of these measures may be considered as standard requirements (i.e., necessary for compliance), and which measures might be considered as 'advanced'. It also puts forward some issues for future research, relating to, intera alia the use of PKI in eGovernment as an

adequate solution and to appropriate technical solutions to ensure that the data, obtained after lawful authorization, are further processed for a legitimate purpose.

2 Introduction

FIDIS Deliverable D16.1 (hereafter 'Fidis D16.1')[4] provided a general overview of requirements that governmental Identity Management Systems (IMS) need to meet. In this document general aspects of data protection compliance and privacy friendliness were also presented. It analysed selected implementations in member states and shed some light on possibly concurring requirements typically put forward by different stakeholders, such as governmental officials and citizens. The deliverable also presented how different European member states implement IMS in a local, national and transnational context.

The current deliverable aims at giving an insight in privacy preserving measures in the context of governmental IMS, based on the analysis of requirements relevant to achieve privacy friendliness. The authors of the deliverable understand privacy friendliness as

- compliance to national data protection legislation including privacy awareness and
- the application of best practices such as Privacy Enhancing Technologies (PETs) going beyond compliance.

The focus of this deliverable is put on organisational and technical measures that are state-of-the-art to achieve data protection compliance (summarised in the chapters 3 and 4) and those which are – from the perspective of the authors - currently exceeding the state-of-the-art. As solutions exceeding state-of-the-art are mostly designed for very specific purposes and specific technical environments, they are not largely implemented in today's governmental IMS. Some approaches discussed in this deliverable are still ongoing research and lack the maturity for immediate large scale application in governmental IMS.

In the context of this deliverable identity management is understood very broadly. As already elaborated in previous FIDIS deliverables[5] in some cases identification and identity management may be carried out based on information collected from e.g. the networking infrastructure. Today's governmental IMS typically do not take these aspects into consideration. In the context of some governmental services such as e-voting[6] or e-petitions they nevertheless may be very relevant. The presented advanced technical approaches in some cases may provide solutions for such eGovernmental services.

This deliverable is structured as follows: following this introduction (chapter 2) requirements for privacy compliant governmental IMS are summarised. This chapter is followed by a summary of - from the perspective of the authors - most relevant basic technical and organisational approaches required from a legal perspective, combining recommendations for general technical and organisational measures to achieve privacy compliant governmental IMS. The fourth chapter "basic technical approaches" contains relevant technical solutions for the same purpose. In the following two chapters advanced solutions are presented, first on a technical level (chapter 5), then on a combined technical and organisational level (chapter 6).

[4] Buitelaar, H., Meints, M. and Van Alsenoy, B. (eds.), *D16.1: Conceptual Framework for Identity Management in eGoverment, FIDIS deliverable*, 2008, available at <www.fidis.net>, last consulted 15 February 2009 (hereafter 'Fidis D16.1').

[5] E.g. Alkassar, A. and Hansen, M. (eds.), *D3.8: Study on protocols with respect to identity and identification – an insight on network protocols and privacy-aware communication, FIDIS Deliverable*, 2008, available at <www.fidis.net>, last consulted 15 February 2009.

[6] See Alkassar, A. and Volkamer, M. (eds.), *E-Voting and Identity*, Springer, Heidelberg, 2007.

The advanced solutions presented in these chapters clearly go beyond compliance with data protection legislation.

3 Requirements for privacy-compliant IdM in eGovernment

In this chapter fundamental requirements for privacy friendly IMS in eGovernment are summarised. They are based to a large extent on the findings of Fidis D16.1. For the general privacy framework relevant to eGovernment, we therefore refer to Fidis D16.1.

The Article 29 Data Protection Working Party has expressed its concern with regard to some specific and complex data protection issues involved in the development of various types of e-government solutions.[7] The involvement of the national DPAs in eGovernment initially mainly concerned implementation of security measures, such as measures relating to identification and authentication of users as well as of agents or professionals accessing applications of online administrations, the encryption of the data and the implementation of logging functionalities.[8] This has evolved to attention to other aspects, in particular the use of the identifier, the unique entry point (portal) and interconnections of public databases.[9]

The use of the electronic identity card to enable access to online administrative procedures has recently become another point of focus in the debate. On this issue, it is worthwhile to recall that the European Commission considered in its i2010 eGovernment Action Plan, that electronic identification management is to be among the "critical key enablers" of eGovernment. However, in the same communication, the Commission stated that in its view, (biometric) national ID cards and electronic identification management for public services are markedly different: *'national ID cards serve public security, for example by facilitating integrated border management and supporting the fight against terrorism, whereas electronic identification for public services is intended to ease access and offer personalised and smarter services*'. [10]

As the Commissioner for Human Rights and the Council of Europe pointed out in their recent report of December 2008, this would mean that 'in data protection terms, this should make it imperative to separate ID cards from eIDM products, and to isolate the databases behind these different products'.[11]

The scope of this contribution is to define requirements and recommendations to achieve privacy friendly eGovernment and to identify the elements which might be used to address those needs.

[7] Article 29 Data Protection Working Party, Working Document on E-Government, *WP* 73, 8 May 2003, p. 18.

[8] Article 29 Data Protection Working Party, Working Document on E-Government, *WP* 73, 8 May 2003, p. 4.

[9] Article 29 Data Protection Working Party, Working Document on E-Government, *WP* 73, 8 May 2003, p. 2.

[10] European Commission Communication of 25 April 2006, *i2010 eGovernment Action Plan - Accelerating eGovernment in Europe for the Benefit of All* [COM(2006) 173 final, (i2010 eGovernment Action Plan), of which a summary is available at <http://europa.eu/scadplus/leg/en/lvb/l24226j.htm and full text (12 p.) at http://ec.europa.eu/information_society/activities/egovernment/docs/highlights/comm_pdf_com_2006_0173_f_en_acte.pdf>, last consulted 15 February 2009, p. 9.

[11] Council of Europe and Commissioner for Human Rights, *Protecting the right to privacy in the fight against terrorism,* December 2008, CommDH/IssuePaper (2008)3, 6.

3.1 Establishing Circles of Trust

eGovernment applications often require the cooperation of a large number of disparate entities. For such collaboration to be successful, agreements need to be made regarding the exchange of identity information among the communicating entities. Such a form of co-operation resembles what has been described in IdM literature as a "Circle of Trust" (CoT), whereby a group of service providers and identity providers share linked (partial) identities and have pertinent business agreements in place regarding how to do business and interact with identities.[12]

Deciding at the level of a CoT that every participant must adhere to certain procedures and policies, may help to significantly limit the operational risk of each participant when he seeks to initiate its own application or data exchange. This approach also has the advantage of minimizing the problems associated with bi-lateral negotiations and multiple contracts with many interdependencies.[13] Finally, it allows privacy considerations to be taken into account during the design of the system.

The basic foundation of a CoT is the reaching of an agreement on how identification and authentication will be organized. Fidis D16.1 described the elements to be taken into account in the management of identity life cycles. Most eGovernment identity management systems have put mechanisms in place for identifying and authenticating their users, most notably by the provisioning of identity documents.[14] However, there are many additional issues concerning information use and governance which need to be addressed in order to create both compliant and successful applications. In the following section we provide an overview of 'elements of trust', which need to be present to ensure compliance with both data protection and functional requirements.

3.2 Support for trust decisions

Trust is a concept that crosses disciplines, so the focus of the definitions differs. In identity management, trust is typically understood in its operational sense. An entity can be said to trust a second entity or a system, when it makes the assumption that the second entity or system will behave exactly as it expects.[15]

[12] See FIDIS 13.3, p. 22 (note 41). Based on Rössler, T., *Identification and Authentication in Networks enabling Single Sign-On*, available at <http://www.iaik.tugraz.ac.at/teaching/11_diplomarbeiten/archive/roessler.pdf>, last consulted 15 February 2009, p. 33 et seq., and J. Hodges (ed.), Liberty Technical Glossary, available at http://www.projectliberty.org/specs/draft-liberty-glossary-v2.0-05.pdf, 9 June 2006, p. 7..

[13] Deadman, S. (ed.), *Circles of Trust: The Implications of EU Data Protection and Privacy Law for Establishing a Legal Framework for Identity Federation*, 2005, available at <www.projectliberty.org>, last consulted 15 February 2009, p. 5.

[14] See also the recent document of Commissioner for Human Rights and Council of Europe, *Protecting the Right to Privacy in the Fight Against Terrorism*, 17 November 2008 mentioned above.

[15] Based on Lead Study Group on Telecommunication Security, Security Compendium Part 2 - Approved ITU-T Security Definitions, available at <http://www.itu.int/ITU-T/studygroups/com17/def005.doc>, last consulted 10 March 2009, p. 51 and ZUCKER, L.G., 'Production of trust: Institutional sources of economic structure, 1840-1920', In, B.M. STAW and L.L. CUMMINGS (ed.), 'Research of organizational behavior', JAI Press Inc., Londen, 1986, p. 53-111, and Slone, S. (ed.), Identity Management. A white paper, 2004, available at < http://www.opengroup.org/onlinepubs/7699959899/toc.pdf>, last consulted 15, February 2009.

Trust may apply only for certain specific actions. A trusted entity can violate the trust with which it was endowed, either by performing actions which it is not supposed to do, or by failing to perform actions which it is expected to.[16]

Establishing trust, especially in the private sphere, often proves quite difficult in practice. Making trust decisions is particularly difficult in a digital environment, when people want to interact with people and organizations they have never met and have little time to get to know at a personal level.[17]

Trust is a crucial aspect of information systems that implement online interactions and transactions, both from the user perspective as well as from the service side perspective: both parties want to be confident, that the transaction will be completed to their mutual satisfaction.[18]

We believe the following trust requirements are essential to the success of eGovernment:

- Trust in identification, authentication and non-repudiation mechanisms: it requires inter alia trust in the accuracy of identifiers, certificates, authentication and digital signature providers, … ;
- Trust in accuracy and integrity of data: this notion of trust refers to the accuracy and integrity of assertions or any other type of digital claim made with regards to individual (or group of) entities also (said to be offered by identity providers in the broad sense);
- Trust in the reliability, availability and performance of the (identity management) systems and protocols of other governmental entities involved in any particular communication;
- Trust in compliance with established policies, including data protection and privacy policies: this notion of trust refers to the expectancy that each party will properly adhere to agreed or stated policies such as data handling policies, access control mechanisms, pseudonym management etc.

As we will elaborate over the next chapters, there are several mechanisms to meet the trust requirements mentioned above. Which mechanisms are appropriate depends on a large number of technical and administrative factors, as well as the cost associated with each mechanism. The needs may vary dramatically according to the application envisaged.[19] For instance, with regards to identification and authentication mechanisms, the level of identity assurance needed may range from minimal (where no or practically no identity assurance is

[16] Slone, S. (ed.), *Identity Management. A white paper*, 2004, available at <http://www.opengroup.org/onlinepubs/7699959899/toc.pdf>, last consulted 15 February 2009

[17] Slone, S. (ed.), *Identity Management. A white paper*, 2004, available at <http://www.opengroup.org/onlinepubs/7699959899/toc.pdf>, last consulted 15 February 2009, p. 7.

[18] HUYSMANS, X. and VAN ALSENOY, B., ‘Conceptual Framework for Identity Management in eGovernment and Requirements Study’, Deliverables 1.1 and 1.3 of the IBBT project ‘IDEM’ (Identity Management for eGovernment), 2007, p. 103.

[19] ITU-T SG17 Focus Group for Identity Management, “Report on Requirements for Global Interoperable Identity Management”, September 2007, available at <www.itu.int/ITU-T/studygroups/com17/fgidm>, accessed 4 December 2007, p.16.

needed; possibility of anonymity), to pseudonymity, to very high assurance level where significant consequences may follow from entity provisioning.[20] (see also section 4.2)

In order to realize trust in the operational sense, one needs trusted parties. If trust services are offered by an entity that is not alien to the internal relationship, we call it a trusted party. A typical example of a trusted party is an entity that acts as an intermediary for eGovernment data exchange but which also has an interest in the exchange, such as the Belgian Crossroads Bank for Social Security.[21]

In large-scale identity management systems, trust services are often offered by Trusted Third Parties (TTPs), i.e. an entity which is trusted by one or more other entities to perform one or more specific actions within a specific context and which is alien to their internal relationship.[22] Trusted Third Parties are – obviously – typically also service providers, either (joint) data controllers or data processors.

From a citizen perspective, we believe that trust (in the broad sense) depends on the perception of whether or not the requirements mentioned at the beginning of this section are seen to be fulfilled. Additional properties which we believe may enhance this trust include: mutual authentication mechanisms, transparency and remote monitoring mechanisms, independent certification and auditing, and privacy-friendly identity management in general and user control in particular.

Before the members of a CoT can count on proper implementation of any of the elements enumerated above, agreements need to be in place by which the participants agree to adhere to certain policies and practices. In the following sections we seek to address certain elements which require particular elaboration from a privacy perspective, and to provide some additional guidance as to how these requirements might be implemented. This list is likely to require additional elements according to the application at hand. Our point of departure here are the demands Directive 95/46/EC places upon large-scale identity management systems. In chapter 4, we provide further elaboration on the technical approaches which can be used to accommodate these requirements. Afterwards we explore alternative mechanisms which might be useful once the appropriate level of maturity is reached.

[20] ITU-T SG17 Focus Group for Identity Management, “Report on Requirements for Global Interoperable Identity Management”, September 2007, www.itu.int/ITU-T/studygroups/com17/fgidm, accessed 4 December 2007, p.16.

[21] HUYSMANS, X. and VAN ALSENOY, B., ‘Conceptual Framework for Identity Management in eGovernment and Requirements Study’, Deliverables 1.1 and 1.3 of the IBBT project ‘IDEM’ (Identity Management for eGovernment), 2007, p. 93.

[22] Based on eGovernment Unit, eGovernment Unit, *Modinis IDM Terminology Paper,* available at <https://www.cosic.esat.kuleuven.be/modinis-idm/twiki/pub/Main/GlossaryDoc/modinis.terminology.paper.v2.01.2005-11-23.pdf>, last consulted 10 March 2009, pp. 11-12.

3.3 Determination of roles and responsibilities

Directive 95/46, which is the general Data Protection Directive,[23] sets out certain roles to which it attaches certain responsibilities. In an eGovernment CoT, any participant (service provider, mediator,[24] integrator,[25] authoritative source, ….) might be acting as a controller, processor or third party depending on the application at hand.

As discussed in Fidis 16.1, every processing operation in eGovernment requires as a rule a legal basis to legitimize the processing. When a law mandates a certain form of processing, it should in principle indicate which entity shall act as a controller. Where legislators are not explicit in this regard, but merely entrusts the processing to a particular governmental entity, it may be assumed that the latter will be responsible for the processing operations that are performed pursuant to this legal basis.[26]

However, it is possible that there are still cases in which these qualifications are difficult to make. For instance, several governmental entities might be charged with complementary tasks of public interest. This, in turn, might require multiple governmental entities, each within their respective domain, to carry out certain processing operations. If there is no clear specification in the law as to which entity shall act as a controller, their respective roles are determined by the general criteria of the Directive (purposes, means). However, we do wish to bring into remembrance here legislators' obligations under art. 8 ECHR to provide a legal basis, which is sufficiently clear and precise.[27]

In any event, the collaborating entities should specify in a written agreement which entity will take up which role vis-à-vis the processing, wherein the different obligations of the parties are appropriately indicated. The allocation of responsibilities in eGovernment CoT should include, but also go beyond the mere distinction of controller, processor and third party that is made by the Directive. This is important. Note that within an eGovernment CoT tasks shall also need to be distributed among multiple co-controllers.[28] Furthermore, certain entities are often charged with offering services that exceed one particular operation.

[23] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, *O.J.* L 281, 23 November 1995, pp. 31-50.

[24] "A mediator is an entity that manages data traffic from and to authoritative sources. Mediator services include (1) routing, (2) transporting, (3) transforming or (4) granting access to the authentic data to authorized users. The latter implies prior authentication." IDEM glossary (p.20-21), available at https://projects.ibbt.be/idem/uploads/media/2007-12-27.idem.glossary.v1.07.pdf (slightly refined during FIDIS glossary workshops).

[25] "An integrator is a mediator that integrates, orchestrates and/or aggregates services from multiple authentic sources and delivers the result to the authorized requesting entity." IDEM glossary, available at <https://projects.ibbt.be/idem/uploads/media/2007-12-27.idem.glossary.v1.07.pdf>, last consulted 15 February 2009, p. 20-21. (slightly refined during FIDIS glossary workshops).

[26] Bot, D. de, *Privacybescherming bij e-government in België. Een kritische analyse van het Rijksregister, de Kruispuntbank van Ondernemingen en de elektronische identiteitskaart* [Protection of privacy in the e-government of Belgium. A critical analysis of the National Register, the Crossroadsbank for Entrerprises and the electronic ID card], Vandenbroele, Brugge, 2005, p. 35.

[27] See also Fidis 16.1, p. 38.

[28] See also recently the Commission Recommendation of 26 March 2009 on data protection guidelines for the Internal Market Information System (IMI), 2009/329/EC, O.J. L 100/12-28, 18 April 2009, at 17.

Directive 95/46 mandates the inclusion of the following elements in a contract between a controller and processor (art. 16-17):

- a provision that the processor may only act on behalf of the controller and as authorized by the controller (unless required by law);
- a stipulation that the processor is bound by the same obligations as those to which the controller is bound;
- an indication of the technical and organizational measures that will govern the processing;
- the liabilities of the processor vis-à-vis the controller.

However, there are many additional tasks which need to be allocated within an eGovernment CoT in order to realize an appropriate division of roles and responsibilities, such as:

- which entities are authorized to act as data providers for which data sets;
- which entity shall perform which authentications, authorizations and checks (as well as the corresponding liabilities related thereto);
- which entity will be charged with the maintenance of logs for which operations[29];
- which entities shall act as trusted parties to which transactions;
- which entities will be charged with the updating of technical policies in accordance with legislative developments and possible authorizations issued by data protection authorities;
- which entities shall serve as a front-office to accommodate the rights of data subjects such as the right of access and correction;
- which entities shall serve as a point-of-contact in the event of a security breach; and
- which entities shall be charged with regular verification of policy compliance.

3.4 Authoritative sources

Every controller must be able to ensure the accuracy of the data he processes (art. 6d Directive 95/46/EC). For that reason, it is crucial that every processing operation is based on information which is sufficiently reliable and up-to-date. To achieve data accuracy, several Member States rely on what can be characterised as 'authoritative sources'.[30] An authoritative source can be described as a data repository, which is managed by one or more entities that are functionally responsible for the collection, validation and updating of data originating from the actual source of the information (e.g. a citizen, a governmental entity, national

[29] See also Belgian Privacy Commission, *Recommendation nr. 01/2008 of 24 September 2008 concerning user- and acces management in the governmental sector,* 24 September 2008, 3, available at <http://www.privacycommission.be/nl/docs/Commission/2008/aanbeveling_01_2008.pdf>, last consulted 19 December 2008.

[30] See FIDIS D16.1, p. 45.

regulatory bodies of professions, …), and which are recognized as being the most qualitative source of such information.[31]

From a public policy perspective, the purpose of authoritative sources is mainly to realize principles of single collection and re-use of data. [32] If implemented and deployed properly, authoritative sources may also be very useful in advancing several privacy principles. In the first place, authoritative sources can help to minimize the amount of data stored centrally, which may dramatically reduce the gains for potential attackers. It also advances the principle of data minimization in the sense that it can avoid unnecessary duplication of the same data. Finally it increases the probability, that the most accurate information is being processed, provided of course that it is sufficiently validated and updated in due time.[33]

Prior to commencing any application or data exchange in the eGovernment domain, participants should establish which the relevant authoritative sources for that application are. In other words, for each data item or data set (e.g. basic identification data, social security status,…) that is needed for a particular application, the authoritative source that will be used must be designated in advance.

Authoritative sources can (and should) also play an important role in user- and access management.[34] The authorization profiles of individual entities are often dependant on attributes of the requesting entity, which may vary over time. For instance, access may be dependant on professional qualifications (e.g. health professional, attorney), membership of a group (e.g. employee of a particular governmental agency), or mandates (e.g. from a citizen to his accountant), which may become revoked or expire. By verifying the relevant characteristics through authoritative sources, users' privileges may more easily be kept up-to-date.

To make such a system work an inventory of authoritative sources and the information they contain is indispensable. Data registries and reference directories can be employed to point out where data needed for a particular application is kept and to enable subsequent data exchange. Seeing as such directories give rise to vast data aggregation capabilities, specific safeguards must be in place to prevent abuse of their functionalities. After all, although the data no longer needs to be maintained centrally, the use of discovery services[35] in turn creates centralized data aggregation opportunities. Discovery services allow an IdM network to locate

[31] Definition based on art. 1, 1° of the Royal Decree of 26 June 2003 implementing the Crossroads Bank of Social Security Act (Belgian Official Journal, 1 July 2003) and art. 2, 2° of the Flemish Act of 18 July 2008 concerning administrative electronic data exchange (Belgian Official Journal, 29 October 2008).

[32] See e.g. Deprest, J. and Robben, F., eGovernment*: the approach of the Belgian federal administration*, available at <http://www.ksz.fgov.be/En/Como/2003%20%20EGovernment%20paper%20v%201.0.pdf>, last consulted 15 February 2009.

[33] See also Belgian Privacy Commission, Recommendation nr. 01/2008 of 24 September 2008 concerning user- and access management in the governmental sector, 24 September 2008, p. 4, available at <http://www.privacycommission.be/nl/docs/Commission/2008/aanbeveling_01_2008.pdf>, last consulted 15 February 2009.

[34] See also Belgian Privacy Commission, Recommendation nr. 01/2008 of 24 september 2008 concerning user- and access management in the governmental sector, 24 September 2008, 6, available at <http://www.privacycommission.be/nl/docs/Commission/2008/aanbeveling_01_2008.pdf>, last consulted 19 December 2008, p.6.

[35] Discovery services provide the capability to inter alia locate identity resources of an entity (i.e. credentials, identifiers, and other attributes) from different sources. For more information regarding discovery services see International Telecommunication Union - Telecommunication Standardization Sector. Focus Group on Identity Management. *Report on Identity Management Framework for Global Interoperability*, p. 18.

a network entity's identity resources, and may include credentials, identifiers and attributes. For this reason, careful consideration should be given to which entities are charged with maintaining and managing and operating data registries and reference directories. Such roles should in principle only be bestowed upon Trusted Third Parties (TTP), who have the ability to act as independent intermediaries towards the requests for data exchange they may receive. Ideally, such TTPs would operate under close supervision of national data protection authorities, and be subject to regular audits. We recommend that eGovernment developers also consider maintaining some form of functional separation among the entities that manage directory services, e.g. by taking into account the different contexts and sectors (such as health, finance, social security, …) that exist within government. Each intermediary should of course then only manage references to the extent that there are legitimate bases requiring its retrieval and exchange of this information.

To maintain the accuracy of the information contained in authoritative sources, certain policies must be in place. Agreements and procedures must be in place regarding how each authoritative source will validate data upon collection, on how inaccuracies will be tracked, reported and dealt with.[36]

Also technical measures that detect and prevent unauthorized manipulation are a necessity. In first instance, this requires a restriction of modification rights (cf. *infra*; authorization management). Secondly, data to and from authoritative sources should be authenticated through use of the appropriate data origin authentication protocols (which in turn also serve to establish integrity during transmission). Finally, it is recommended, that the metadata of the information contained in authoritative sources provides some indication of the 'level of confidence' of the information (e.g. date of collection, last update, etc).

3.5 Authorization management

It is a basic principle of data protection that data access should be restricted to authorized entities and at the same time be limited to the data needed so that an authorized entity can execute its task adequately. Confidentiality is generally understood as keeping the content of information secret from all but those authorized to access it.[37]

There are numerous approaches to providing confidentiality, ranging from physical protection to the use of access control and cryptographic algorithms.[38] There are however several other security objectives and privacy principles which need to be taken into account besides confidentiality in order to create a privacy-friendly IMS. For instance, confidentiality does not refer to other important privacy aspects, such as data accuracy, linkability or the restriction of further processing capabilities. In the following sections, we shall elaborate on some key points of attention in realizing privacy policy enforcement.

[36] See e.g. Deprest, J. and Robben, F., eGovernment: the approach of the Belgian federal administration, available at <http://www.ksz.fgov.be/En/Como/2003%20%20EGovernment%20paper%20v%201.0.pdf>, last consulted 15 February 2009, p.7.

[37] Menezes, A.J., Van Oorschot, P.C. and Vanstone, S.A., *Handbook of Applied Cryptography*, CRC Press, Boca Raton, 1997, p. 32. See also IDABC*, IDABC Glossary*, available at <http://europa.eu.int/idabc/servlets/Doc?id=1348>, last consulted 10 March 2009, p. 5.

[38] Menezes, A.J., Van Oorschot, P.C. and Vanstone, S.A., *Handbook of Applied Cryptography*, CRC Press, Boca Raton, 1997, p. 32.

As we just highlighted, ensuring that access to personal data is limited to duly authorized entities is a standard requirement. In addition, the processing capabilities (read, write, modify …) of each entity should be limited to that which is necessary to realize the goals of the processing. This follows from a combined reading of the controller's security obligation and the proportionality principle. These requirements apply not only at the level of each governmental entity but also at the level of each individual user.

The aforementioned requirements can (in part) be accommodated through implementation of technical policies, as elaborated in section 6.3. In this section, we discuss elements that should be included and the organizational measures which need to be taken to transform such policy languages into privacy policies.

Prior to any application or network connection, the available resources and services need to be documented. The personal data contained in data repositories should be categorized in a generic fashion (e.g. contact data, age, social security status, date of application …). After an overview has been made of the types of information that are needed, business processes and information flows must be mapped out so that the function and role of each possible user can be clarified. The access and processing capabilities of each entity can then be determined, but must be defined according to that which is strictly necessary for the requirements of the application or service. This should result in an overview of valid recipients for each object that qualifies as personal data, as well as a list of the actions they are allowed to perform upon these resources.

Given the scale of eGovernment, it appears to be neither sufficient nor practical to adequately manage users' rights entirely under a model of role-based access control. This holds particularly in instances where users need to be authorized across multiple domains. As indicated earlier, processing rights are often dependant on a wide variety of attributes, such as mandates, group membership, professional qualifications etc. which may change in time. By basing authorization decisions on relevant attributes instead of roles, users' privileges become more manageable and may more easily be kept up-to-date. This in turn serves both the security and proportionality of the processing.

When setting up a new application or network connection, developers should carefully consider precisely which attributes need to be present to justify authorization. Authorization policies should specify in which capacity(ies) a resource or service is accessible to users, as well as the situation (i.e. for what purpose) and the time-frame.[39] Where intermediaries (e.g. mediators, service integrators) are used, these policies should in first instance be managed and enforced at that level. The technical authorization mechanism used must of course also allow for sufficient granularity as to the permissions of every possible requesting/ asserting entity.

In order to mitigate the privacy risks associated with the use of a single unique identifier in eGovernment, several Member States have introduced a system of prior checking and authorization which is performed by their national Data Protection Authority. Such a model is highly recommendable,[40] even when several distinct identifiers are employed. Where it exists, the technical authorization mechanisms that are used should have the ability of verifying whether or not such a prior authorization has been issued. This of course requires that the

[39] See also Deprest, J. and Robben, F., *eGovernment: the approach of the Belgian federal administration*, available at <http://www.ksz.fgov.be/En/Como/2003%20%20EGovernment%20paper%20v%201.0.pdf>, last consulted 15 February 2009, p. 45.

[40] Reason for this is that the element of human intervention allows for an evaluation of the intended operations in a manner which cannot be achieved through technical means alone.

authorizations (or legal exemptions thereto) are systematically updated to the system's policy information points in order to maintain the functionality of the system.

Although the approach described above provides a relatively high level of protection, it still displays an important flaw: while it does verify the presence of a prior authorization, it exercises no control over the purpose for which the action takes place. In other words, once the condition of authorization has been met (y/n), there is no mechanism that verifies whether the data is in fact being processed (requested) for a legitimate purpose in that instance. This implies that the described model still allows for 'free searches' over all the referenced data repositories, enabling access to 'as much data as possible' as long as the requested data falls within the scope of the authorization profile of the requesting entity.

Therefore it may also be recommendable that technical solutions are implemented which enable a verification (or at least registration) of the purpose of the request, so that only the data needed for the processing are disclosed, even if the requesting entity's authorization profile in principle permits access to greater amounts of data. This would not only be an important additional safeguard, both with regards to the finality principle as with regards to the principle of data minimization,[41] it could also have the advantage of rendering the audit trail more intelligible (cf. *infra*; section – 3.9 Logging & auditing).[42] Seeing as information relating to the purpose for which data is accessed may in itself also reveal sensitive information, careful consideration must also be given to which entity shall be trusted with registering and/or verifying the purpose of individual operations.

3.6 Authentication and assurance levels

Not all authentication and assurance levels need to be of the same robustness. In this section a brief summary is given of the procedures that may be followed in establishing the assurance level of an authentication process. This process should of course take into account the relevant principles of data protection.

In general, it can be said that the determination of the appropriate level of authentication assurance should start off with a risk assessment of the application or system. This should also be mapped with the identities individual entities may assert when accessing the application. Subsequently the identified risks are to be mapped against the applicable authentication assurance level. Authentication assurance is defined as the degree of confidence in an asserted real-world identity (determined inter alia by the policies controlling identity proofing) and·

[41] See also Langheinrich, M. and Roussopoulos, M. (eds.), , *Technology-Induced challenges in Privacy & Data Protection in Europe*, ,available at <http://www.enisa.europa.eu/doc/pdf/deliverables/enisa_privacy_wg_report.pdf>, last consulted 15 February 2009, pp. 25-26.

[42] When implementing a purpose specification mechanism, special consideration should be given to the entity that will register the asserted purpose. If the purpose were simply to be recorded by the entity that is being queried, this entity is likely to learn additional information with regards to the data subjects involved; without this being necessary. In order to avoid unnecessary "leaking" of potentially sensitive information, the registration of purpose could be performed by an entity other than the queried entity, e.g. by a trusted intermediary. Note however that this will not completely remove the issue but rather shift the problem to the level of the intermediary.

the degree of confidence in an electronic identity presented to a service provider by means of a credential (i.e. proof of possession).[43]

The authentication assurance levels should be layered according to the severity of the impact of damages that might arise from misappropriation of a person's identity. The more severe the consequences, the higher degree of confidence/trust in an asserted identity will be required prior to allowing an action to take place.

The literature commonly defines four assurance levels:

· Level 1: ....Minimal Assurance

· Level 2: ....Low Assurance

· Level 3: ....Substantial Assurance

· Level 4: ....High Assurance

Authentication errors with potentially worse consequences require higher levels of assurance. Business process, policy, and technology may help to reduce risk.

Several methods can be used to determine the level of risk. In the United States, the OMB advises agencies to follow a five-step process in determining the appropriate assurance level for their applications:

- . Conduct a risk assessment for e-authentication of the system. The risk analysis measures the severity of potential harm and the likelihood of occurrence of adverse impacts to the system if there is an error in identity authentication.
- . Map identified risks to the applicable assurance level. After all of the risks have been identified, agencies should tie the potential impact of the risks to the proper level of authentication to be used.
- . Select technology based on e-authentication technical guidance. [44]
- . Validate that the implemented system has achieved the required assurance level. A final validation is needed to confirm that the system achieves the required level of assurance, and that the selected authentication process satisfies requirements.
- Periodically reassess the system to determine technology refresh requirements. Reassessments ensure that the authentication requirements continue to be valid as technology and requirements change.

[43] IDA Authentication Policy. *Basic policy for establsihing the appropriate authentication mechanisms in sectoral networks and projects*, available at <http://ec.europa.eu/idabc/servlets/Doc?id=19281>, last consulted 15 February 2009, p. 18.

[44] In December 2003 the US Office of Management and Budget (Memorandum M-04-04, E-Authentication Guidance for Federal Agencies) advises that agencies refer to the technical guidance issued by National Institute of Standards and Technology, US Department of Commerce (NIST). Vide Sh. Radack, Electronic Authentication: Guidance for selecting secure techniques, ITL Bulletin, August 2004, available at: http://carc.nist.gov/publications/nistbul/August -2004.pdf.

Assessment should be made of potential impacts of an authentication error by taking into consideration:

- Inconvenience, distress, or damage to standing or reputation;
- Financial loss or agency liability;
- Harm to agency programs or public interests;
- Unauthorized release of sensitive information;
- Personal safety; and/or
- Civil or criminal violations.

In an eGovernment setting separate entities are often required to collaborate in different sectors. An important requirement for the successful collaboration of these entities is the drafting of an agreement concerning the procedures that will be followed in the identification and authentication phase. A useful point of reference for such an agreement can be found in the mutual recognition agreement as proposed in the IDA Authentication Policy for establishing the appropriate authentication mechanisms in sectoral networks and projects.[45]

Considering that personal information plays a central role in most authentication solutions, privacy should therefore be a key factor in determining the most appropriate implementation and operation of authentication solutions. This can be achieved by paying due attention to the principles of data minimisation and proportionality. It is helpful to make a distinction between the entity and its attributes. Only those attributes are recorded in this process that are necessary in the light of the assurance level required.[46] The authentication principles[47] require that matters such as personal choice (opt-in) and privacy be given equal weight to considerations such as cost.[48] These are to be referred to, and considered as part of, any government agency authentication initiative.

Allocation of liability if things go wrong must also be considered. If, for example, there is a serious system failure or other problem with the solution that leads to loss, it is important that all parties know who can and cannot be held liable for any wrongdoing, and what the implications of that liability are. In general, these concerns can be handled either through statute (i.e. legislation defining liabilities and the extent of each type of liability), via common law such as the law of negligence and/or through contract (such as a contract between the Client and the service agency). Many existing online commercial services, such as Internet

[45] Enterprise DG, *IDA Authentication Policy. Basic policy for establishing the appropriate authentication mechanisms in sectoral networks and projects*, available at <http://ec.europa.eu/idabc/servlets/Doc?id=19281>, last consulted 20 October 2006.

[46] Enterprise DG, *IDA Authentication Policy. Basic policy for establishing the appropriate authentication mechanisms in sectoral networks and projects*, available at <http://ec.europa.eu/idabc/servlets/Doc?id=19281>, last consulted 20 October 2006, p. 41: "Identity information shall include at a minimum full legal name, date and place of birth, current address, identity numbers of any documents checked in the registration process such as passport etc".

[47] Cf Enterprise DG, *IDA Authentication Policy. Basic policy for establishing the appropriate authentication mechanisms in sectoral networks and projects*, available at <http://ec.europa.eu/idabc/servlets/Doc?id=19281>, last consulted 20 October 2006, pp. 33-35 which proposes guiding principles for the IDA Authentication Policy.

[48] Cf., eGovernment Unit of New Zeeland, *Authentication for e-government. Best Practice Framework for Authentication*, available at<http://www.e.govt.nz/services/authentication/authentication-bpf/bpf.pdf>, last consuled 15 February 2009, p. 26.

banking, rely on the latter legal provision and ask that individuals accept terms and conditions as part of registering for an online service. In the case of a CoT, however, focus is more on allocation of responsibilities by Members of a CoT among each other, rather than the terms of use that exist between the individual citizen and a particular service provider.

3.7 Security policies

Access control policies play a crucial role in protecting privacy. They are however generally limited to challenging a requesting entity to produce the appropriate credentials, and subsequently evaluating their processing rights in light of the applicable policy. On the other hand, data is most often transmitted across public networks, which introduces additional security risks (interception, MITM etc) which cannot be resolved by access control policies alone.

Security policies allow specifying how the data to be exchanged with another actor, should be protected. These policies indicate which security levels should be applied for which type of transactions. They can refer to both authenticity as well as to confidentiality.[49]

Encryption is a technique which transforms data from a readable form (known as plain text or clear text) to one that is unintelligible (referred to as cipher text). It may be applied during transmission as well as during storage. This helps to maintain confidentiality even in instances where data has been intercepted or the security of a database has been compromised. Further elaboration on how encryption may be used in and outside IdM systems is discussed in section 4.4 (Encryption Schemes & Secrecy).

In order to ensure data accuracy, it is of major importance to have a sufficient level of certainty as to the identity of the information provider. Parties involved in an exchange must after all be able to establish whether the information emanates from a qualitative and authorized source. Data to and from authoritative sources should therefore be authenticated through use of data origin authentication protocols (which also serves to establish their integrity during transmission). Relying parties should only be permitted to process personal data further if there is sufficient certainty as to its origin and integrity (i.e. upon verification that it emanates from the intended source and has not been subject to manipulation). Certain cryptographic techniques and models such as Public Key Infrastructure (PKI) have been developed to help secure these objectives. They are elaborated in section 4.2.

In short, it is required that appropriate security policies are adopted to specify how data will be protected when it is exchanged among actors, particularly where authoritative sources are involved. It is also required to make similar agreements that will govern the electronic exchange of the results of executed authentications and verifications performed by involved parties.[50]

[49] See also Robben, F., *Een voorstel van informatiebeveiligingsbeleid bij de uitbouw van E-government door de federale overheidsdiensten,* available at <http://www.law.kuleuven.ac.be/icri/frobben/publications/2005%20-%20Voorstel%20van%20informatieveiligheidsbeleid%20bij%20de%20uitbouw%20van%20E-government.pdf>, last consulted 10 March 2009.

[50] See also Belgian Privacy Commission, Recommendation nr. 01/2008 of 24 september 2008 concerning user- and access management in the governmental sector, 24 September 2008, p. 4, available at <http://www.privacycommission.be/nl/docs/Commission/2008/aanbeveling_01_2008.pdf>, last consulted 19 December 2008.

3.8 Restrictions and obligations

The adaptation of privacy policies should not only be thought of in terms of access control and data release policies. The parties involved in a data exchange should also agree in advance how the receiving entity may further process the data. An additional measure to ensure that the finality principle is respected consists in imposing restrictions and obligations on the receiving entity.

Data handling policies can be described as policies that define restrictions on the secondary use of personal data once it has been received by a particular entity. If the eGovernment framework is based on authoritative sources as described above, further transmission or release of personal data by a data recipient should as a general rule be denied (unless of course that entity is acting as an intermediary or data is being processed within a given value chain).

As the reader is aware, data may only be collected for a specific purpose and kept in a form which permits identification for no longer than necessary to achieve the purpose(s) of the processing. Prior to initiating an application, the storage duration of each data category should be specified for every entity involved. There should also be a clear understanding on how information will be deleted once the goals of the processing have been achieved. Sanitization policies are policies that define how long each data item may be retained and give directives on how the information should be removed. The length of data retention, of course, varies upon the purpose of the processing.

Finally, certain processing operations might be particularly sensitive and merit closer follow-up. To duly alert relevant entities (e.g. supervisors, security officers), an automated notification service could be installed. Similarly, certain processing activities (e.g. "breaking the glass") should be investigated immediately and should therefore trigger notification. As an additional transparency enhancing measure, developers should also consider incorporating the possibility of forwarding notifications to the data subject into their application.

3.9 Logging & Auditing

With the aid of logging and monitoring, it is possible to investigate after the fact whether the established policies have in fact been adhered to. From a privacy perspective, it is important to log every action or every set of actions that are performed with respect to resources involving personal data. This creates an audit trail ("who did what and when"), which can later be reviewed for policy compliance.[51] In that way logs also assist in creating accountability.

The members of an eGovernment CoT need to have agreements in place as to which entities will manage which logs and how audits will be organized.[52] In particular, they need to establish how supervisory entities shall be able, either at their own initiative or pursuant to a

[51] Koorn, R.(ed.), *Privacy Enhancing Technologies – White Paper for Decision-Makers*, written for the Dutch Ministry of Interior and Kingdom relations, available at <http://www.dutchdpa.nl>, last consulted 22 May 2007, p. 35.

[52] Robben, F., *Gebruikers- en toegangsbeheer: beschikbare diensten*, available at <http://www.law.kuleuven.be/icri/frobben/presentations/20061108.ppt>, last consulted 15 June 2007.

complaint, to perform a full tracing of processing operations performed upon personal data.[53] In this regard it is important to have readily available documentation of the business processes and information flows so that abnormalities can more easily be detected.

In Fidis D14.6, an overview is provided of international security related standards which deal with logging.[54] The same deliverable also discusses the importance of guaranteeing certain security objectives when creating an accountability framework through logging. In first instance, it is important that a logged event cannot be altered or deleted without this being noticed.[55] This implies appropriate measures to ensure the integrity of the logs. Other requirements which map with security objectives in this context include authenticity of logged events and ensuring the completeness and uniqueness of the logs.[56]

Data contained in logs generally also qualify as personal data themselves. Consequently the logs should only be available to authorized entities. Obviously the processing of such data (logs) should also adhere to data protection principles. Confidentiality of logs can be achieved by implementing additional access control mechanisms and by applying encryption techniques.

[53] Belgian Privacy Commission, Recommendation nr. 01/2008 of 24 september 2008 concerning user- and access management in the governmental sector, 24 September 2008, p. 4, available at <http://www.privacycommission.be/nl/docs/Commission/2008/aanbeveling_01_2008.pdf>, last consulted 19 December 2008.

[54] Müller, G. and Wohlgemuth, S., *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009, p. 22 et seq.

[55] Wouters, K. et al., ‘Secure and Privacy-Friendly Logging for eGovernment services’, in Ares *2008 - Proceedings of the Third International Conference on Availability, Security and Reliability*, pp. 1091-1096, IEEE Computer Society, p. 1091-1092.

[56] See Müller, G. and Wohlgemuth, S., *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009, p. 21.

4 Basic Technical Approaches

This section introduces the technical fundamentals behind identity and information security management. It provides further elaboration on several of the techniques highlighted in the previous chapter. In particular we focus on authentication, authorisation and access control mechanisms, as well as the encryption schemes which are used in this context e.g. to ensure secrecy of data and on digital signatures.

4.1 Authentication, Authorisation, Access Control (AAA)

Authentication techniques are, in contrast to encryption schemes, conducive to the integrity and accountability of a user of a system or of a message. They can be combined with encryption in order to achieve both, integrity and secrecy.

Access control

According to the classification of Identity Management Systems (IDM) by Bauer et al.,[57] access control is one of the fundamental building blocks of Type 1 IDM (Account Management Systems). Though not explicitly mentioned, the regulation of access is required for Type 2 IDM (Profiling Systems) and Type 3 IDM (User-Controlled Context-Dependent Role and Pseudonym Management Systems). While the authors of D3.1 refer to role-based access control, generally all forms of access control can be applied in IDM scenarios. These forms differ in the way they authenticate the subject and the granularity level of access permissions or the expressiveness of the underlying access control logic, respectively. In literature about access control, there are two established terms, subject and object. The term “subject” refers to the person which seeks to get access to an “object”. The object might be any resource which is worth to be protected against unauthorized access.

As to the authentication, early access control systems depend on a login-based authentication. That is, each subject has a pseudonym which is only accepted by the access control system in combination with the corresponding password. This allows managing a limited number of subjects with different permissions on a fine-grained level. The permissions can be expressed in a binary matrix where there is a row for each subject and the entries in the columns denote the permissions. In eGovernment this could be applied, if data access is to be regulated between several (well known) governmental institutions. The institutions would be the subjects and the data of each institution would be the object.[58]

A much more flexible way of managing access permissions are policies. A policy applies to a set of subjects and further attributes of the context in which access should be granted or refused. For instance, in data protection, the purpose of accessing personal data plays an important role and is therefore evaluated in many privacy policy languages. However, the flexibility of policies may lead to contradicting policies which makes it necessary to deal with conflicts. Policy conflicts are either resolved by simple strategies, for instance, only the first matching policy is relevant, or more sophisticated strategies, for instance policy hierarchies or

[57] Bauer, M., Meints, M., and Hansen, M. (eds.), *D3.1 Structured Overview on Prototypes and Concepts of Identity Management Systems, FIDIS Deliverable,* 2005, available at <www.fidis.net>, last consulted 15 February 2009.

[58] This describes a simplified approach of what would actually be necessary in most cases to enable access control between governmental institutions. For instance, roles could facilitate an abstract seperation of duties.

combining the results of each matching policy by logical operators. The choice of combining strategies depends a lot on the policy language and on the corresponding implementation of the access control system.

Policy languages have been surveyed by Alkassar and Hansen.[59] Two of the most prominent examples are the Enterprise Privacy Authorization Language (EPAL) and the eXtensible Access Control Markup Language (XACML). EPAL[60] has been developed by IBM in order to support enterprises in managing their (internal) access control requirements. The language is designed to express privacy policies that need to be enforced within large companies. XACML[61] is an OASIS standard language for general access control policies. It has a wide range of applications, that includes not only privacy policies of enterprises, but also for instance user-side privacy policies.[62] Even though, EPAL was famous, when it was submitted to the W3C for standardisation in 2003, recent developments tend to favour XACML over EPAL for privacy policies. A reason is that XACML already is a standard while the standardisation process of EPAL is still pending. But there are also problems with resolving policy conflicts in EPAL, which do not appear in XACML due to a more sophisticated choice of policy combination algorithms. A comparison between the two languages can be found in.[63]

Apart from pseudonyms, authentication can be exercised by means of credentials. This allows to regulate the access depending on the properties a subject has. In particular, credential-based access control does not require to identify or re-recognise the subject. It is sufficient that the subject proves some of its more or less common properties. Depending on the specificity of the properties, this kind of authentication still allows access control decisions while the subjects remain rather anonymous. This could be applied, if a service is to be offered only to citizens with specific properties, for instance, all citizens of a city or all clerks of an institution.

The advantages of both, using policies instead of specific permission assignments for each subject and working with credentials for authentication without (necessarily) identifying subjects, can also be combined. This has been explored by Ardagna et al..[64]

Apart from instant policy checks by means of access control facilities, there are researchers advocating for policy compliance checks after the fact. This requires that system properties (or system behaviour) of the data processing system can be proved. If so, it is possible to drop access control facilities at the site of the data provider in favour for the (provable) self regulation at the site of the data processor.

[59] Alkassar, A. and Hansen, M. (eds.), *D3.8: Study on protocols with respect to identity and identification – an insight on network protocols and privacy-aware communication, FIDIS Deliverable,* 2008, available at <www.fidis.net>, last consulted 15 February 2009.

[60] IBM, Enterprise Privacy Authorization Language (EPAL), W3C Submission Request, Nov. 2003, available at <http://www.w3.org/Submission/2003/07/>, last consulted 3 February 2008.

[61] Moses, T., *eXtensible Access Control Markup Language (XACML) Version 2.0*, OASIS Standard, Feb. 2005, available at <http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-core-spec-os.pdf>, last consulted 3 February 2008.

[62] Moses, T., (ed.). *Privacy policy profile of XACML v2.0, OASIS Standard*, Feb. 2005, available at <http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-privacy_profile-spec-os.pdf>, last consulted 3 February 2008.

[63] Anderson, A., *A Comparison of Two Privacy Policy Languages: EPAL and XACML*, Proceedings of the 3rd ACM workshop on Secure web services, ACM New York, 2006, pp. 53-60.

[64] Ardagna, C. A. et al., *Exploiting cryptography for privacy-enhanced access control*. To appear in Journal of Computer Security, 2009.

There are two ways of achieving proofs of policy compliant data processing. The first approach, which is described in detail in Müller and Wohlgemuth,[65] is to record the system behaviour and secure the records in a way, that they cannot be modified by unauthorised persons. The only authorised person is the auditor, who can check the records and therefore check that the system behaved according to the policies. This approach is known as generating privacy evidences by means of secure logging. Secure logging is also the basic mechanism in section 5.2. The technological details, such as the use of encryption and hashes, are thus discussed in that section. The second approach, which has been discussed only briefly in Müller and Wohlgemuth,[66] produces the same kind of evidence, but proves the policy compliance of the system by means of provable system properties, such as known from trusted computing.[67] In this approach, there is also an auditor, or to be more precise there have to be two kinds of auditors at least. The first auditor checks that (general) proofs about the required system properties hold for the system. Then, the system is sealed such that it can be used, but any system modification after the first check can be made obvious by the second auditor. The second auditor continuously checks the seal. If the seal is broken, then the system is *maybe* not working according to the policies. In any case, this requires that the system properties are checked again by the first auditor. Any data, which has been processed by the system in the meantime, between the last successful check of the seal and the discovery of the broken seal should be considered as insecure or leaked from regulation.

In total, we see that there are three options. The first one is instant access control and needs to be installed at the site of each data provider. The other two are secure logging for privacy evidences and trusted computing. They would be needed to be implemented at the site of the data processor.

Out of these three options, only the second allows to check for policy compliance with respect to policies that are changing over time. This is necessary, if for instance the policies are not entirely clear at system construction time. This advantage can be achieved for any of the other approaches by complementing it with secure logging for privacy evidences.[68]

The trusted computing approach has the advantage, that only the first auditor has to spend much effort in the check. The checks of the second auditor are rather simple, that is they cause almost no additional effort. The price of secure logging and trusted computing is, that both approaches allow a time of uncertainty about the policy compliance between the checks of the auditors. This can be mitigated by raising the frequency of checks.

None of the approaches alone provides a satisfying solution for policy enforcement. Access control at the site of the data provider cannot be used to control the processing of the data at

[65] Müller, G. and Wohlgemuth, S., *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009.

[66] Müller, G. and Wohlgemuth, S., *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009.

[67] See also Alkassar, A., Husseiki, R., *D3.9: Study on the Impact of Trusted Computing on Identity and Identity Management*, FIDIS Deliverable, 2008, available at <www.fidis.net>, last consulted 15 February 2009 and Müller, G. and Wohlgemuth, S. (eds.), *D14.3 Study on the Suitability of Trusted Computing to support Privacy in Business Processes*, *FIDIS Deliverable*, 2008, available at <www.fidis.net>, last consulted 15 February 2009.

[68] See also the outline in Müller, G. and Wohlgemuth, S., *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009.

the site of the data processor. Thus, personal data which is disclosed in compliance with the policy can unnoticeably be leaked by the data processor to any other processor. Trusted computing has been designed to protect the properties of a few against misuse in the masses. However, the same mechanisms applied in data protection are required to protect the property of masses. Several works have pointed out, that the trusted computing mechanisms we know at the moment do not scale well for this task. This issue has been briefly discussed Müller Wohlgemuth,[69] and in detail in Korba and Kenny[70] as well as Böhme and Pfitzmann.[71] Unfortunately, it is also necessary to apply trusted computing mechanisms for creating privacy evidences. This is for ensuring, that the records about the system behaviour are correct (at the time of logging) and complete, that is none of the records has been dropped before being secured by cryptography in the logs.

Factors of Authentication

As already introduced by Leenes[72] for authentication of a user three factors can be used. The strength of authentication is influenced by the quality requirements for the factor or factors, and the number of factors required. The factors are:

- Knowledge: Something I know, e.g. a PIN or password; the length and complexity of the password have an influence of the strength of authentication.
- Possession: Something I have such as a token or chip card that can not easily be copied.
- Biometric features: Something that physically belongs to me such as a fingerprint, the hand or face geometry etc.

All of these factors are used in governmental IMS. Knowledge for example is used for authentication purposes of governmental employees, possession in the context of electronic signing (signature smart cards) and biometrics in the context of Machine Readable Travel Documents.[73]

Message Authentication Codes

For authentication of citizens in the digital world, it is necessary to authenticate messages, since they are the vehicles for the authentication data of citizens. In (large) networks where the authenticity of messages cannot be guaranteed by the physical infrastructure, it is common

[69] Müller, G. and Wohlgemuth, S., *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009.

[70] Korba, L. and Kenny, S., *Towards meeting the privacy challenge: adapting DRM*, available at < http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.8617>, last consulted 10 March 2009.

[71] Böhme, R. and Pfitzmann, A., 'Digital Rights Management zum Schutz personenbezogener Daten?'[Digital Rights Management for the protection of personal data?], Datenschutz und Datensicherheit, vol. 32, issue 5, 2008, pp. 342-347.

[72] Leenes, R. (ed.), D5.2b: ID-related Crime: Towards a Common Ground for Interdisciplinary Research, FIDIS Deliverable, 2006, available at <www.fidis.net>, last consulted 15 February 2009.

[73] See Meints, M. and Hansen, M. (eds.), *D3.6: Study on ID Documents, FIDIS Deliverable,* 2006, available at <www.fidis.net>, last consulted 15 February 2009.

to deploy message authentication codes (MACs) for message authentication. Note that, as a side effect, MACs also guarantee the integrity of messages.

Message authentication codes can be seen as the counterpart of symmmetric encryption schemes. That is, similar to symmetric encryption schemes where the same key is used for encryption and decryption, in message authentication codes, the same key as used for creating and verifying the MAC value and thus the integrity of the message. Figure 1 shows the conceptual framework for MACs.

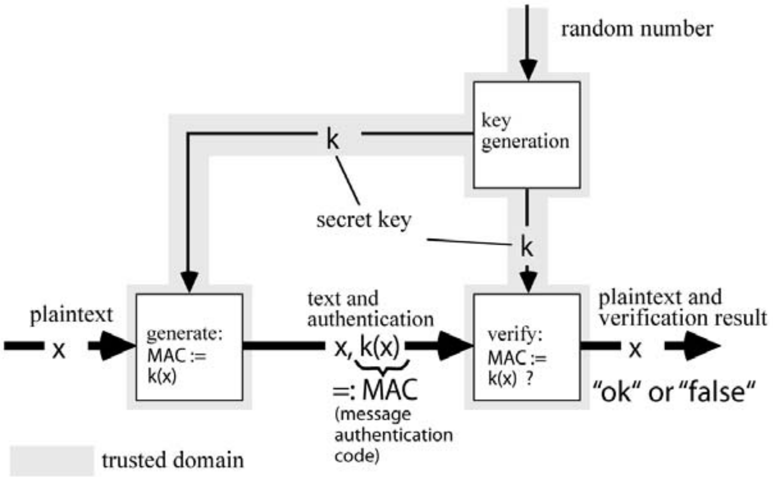

Figure 1: Symmetric message authentication scheme

A common implementation of MACs is the usage of keyed hash functions (HMAC). A hash value of a given message is often referred to be a "(digital) fingerprint" of that message. Conceptually a hash value is the result of applying a so-called hash function (also known as digest algorithm) to a given message. This hash value is constant in size for a given hash algorithm and typically much shorter than the message itself. Usual sizes of hash values are, between 160 bits and 512 bits. Fundamental properties of cryptographic hash functions are that it is hard (i.e. needs unreasonable computing power and storage) (a) to find a message that matches a given hash value and (b) to find so-called "collisions", i.e. two different messages, which have the same hash value.

Today, the two hash functions MD5 and SHA-1 are widely used. But especially MD5 can be seen as insecure and should not be used anymore. Moreover recent cryptanalysis found weaknesses in SHA-1, too. So for new applications the more secure versions SHA-256 and SHA-5125 should be used.

To be useful for MACs, the hash function has to be initialized with the secret key. Thus the receiver of a message can be sure, that a message has been sent by a particular sender and has not been modified on the way from the sender. However, the receiver cannot convince a third party of the authenticity of the message, since the receiver is in possession of the key and could, therefore, have faked the authentication code himself.

Kerberos

Based on research result of the Athena-Project at the Massachusetts Institute of Technology (MIT) in 1993, John Kohl and Clifford Neuman developed the current version five of the Kerberos network authentication service. Kerberos allows based on so called tickets the mutual (multi way) authentication of (a) the Kerberos-service itself, (b) multiple application services and (c) multiple users against a central repository in a so-called realm (realm database).[74] As authentication messages and tickets are strongly encrypted, Kerberos is considered to be safe against various forms of Man-in-the-Middle-attacks (e.g. sniffing, spoofing, and dictionary and replay attacks). From a user's point of view, Kerberos also supports single-sign-on, as an initially so called "ticket granting ticket" can be reused to acquire service specific authentication tickets as a background service.

Cryptographically Kerberos V. 5 supports symmetric cryptographic algorithms such as 3DES, AES and RC4 and hash algorithms such as MD5 (not recommended any more as collision attacks are possible), HMAC and SHA-1.

The following figure shows the basic workflow of Kerberos authentication in a simplified way:

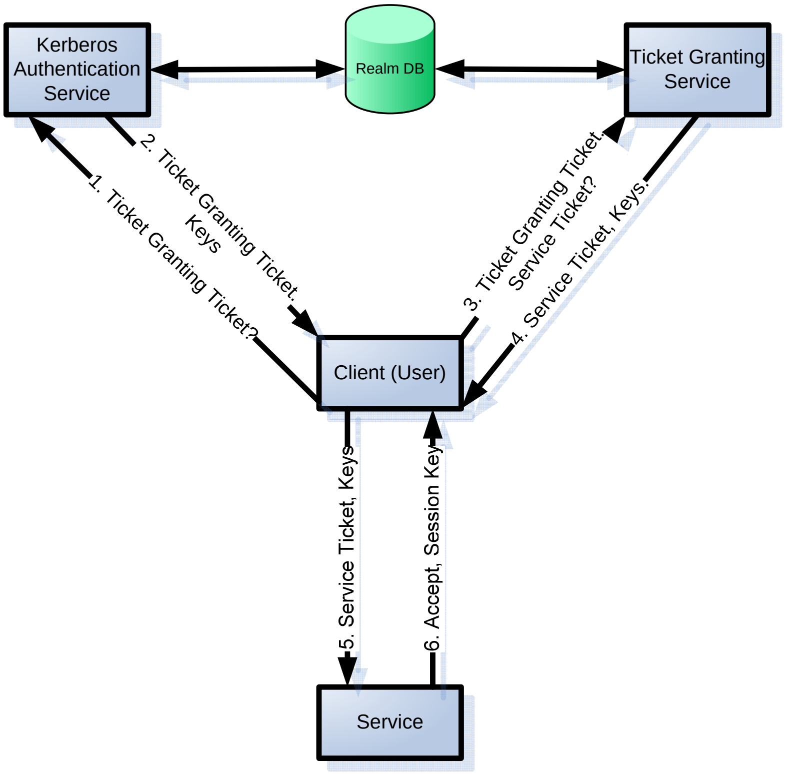
Kerberos Authentication Service
Realm DB
Ticket Granting Service
1. Ticket Granting Ticket?
2. Ticket Granting Ticket, Keys
3. Ticket Granting Ticket, Service Ticket?
4. Service Ticket, Keys.
Client (User)
5. Service Ticket, Keys
6. Accept, Session Key
Service

Figure 2: Simplified scheme of Kerberos authentication[75]

Since 2001 Kerberos is largely distributed in the Directory Service "ActiveDirectory" included in Microsoft Windows Server and clients operating systems (Windows Server 2000

[74] See e.g. http://www.kerberos.info/

[75] Figure based on http://de.wikipedia.org/wiki/Kerberos_(Informatik)

and 2003, Windows client 2000 and XP). For many Unix operating systems and applications (e.g. Apache web service) also Kerberos implementations are available.[76]

LDAP-Directory Services

Directory services are central repositories for the management of resources in a network domain. Resources managed in this context typically are:

- Users (user names, assigned roles, authentication data, and other attributes such as e-mail addresses etc.)
- Services in the network such as applications, printers, file services etc. and rights for the use of these services relying on the roles defined
- Configuration data for hosts (clients and servers) in the network
- Etc.

For directory services, typically the Lightweight Directory Access Protocol (LDAP) specification, currently standardised as RFC 2251[77] by the IETF are used. This specification includes a structure for the storage of attributes in the directory and protocols for the access and exchange of directory data.

LDAP directories are frequently used as backend technology. They can be used as realm databases for Kerberos network authentication services and due to the compatibility to the X.500 framework, as backend technology for Public Key Infrastructure (see section 4.2). Thus directory services are largely used as backend technology for governmental identity management systems.

4.2 Public Key Infrastructure

One of the biggest obstacles to the adoption of modern cryptographic algorithm and protocols is, from an organizational and usability point of view, the burden of key distribution. If one wants to use symmetric algorithms this is more obvious, as a trustworthy (i. e. secrecy-protecting) channel is needed for the transportation of the secret keys. But even in the case of asymmetric cryptography, where public keys are used and therefore no secrecy-protecting channel is necessary, one still faces the problem of integrity and accountability when distributing keys.

Public key infrastructures (PKIs) are an approach to solve these problems. Though, PKIs require complex key management facilities, they are usually used in large organisations for key and trust management, since the effort of setting up a PKI is paid off by its use in large scale. An application area for PKIs is the WWW where web servers establish their authenticity by means of certificates managed in a PKI. The users only need to trust a single root certificate and the service which is issuing further certificates in the PKI. The root certificate and the trust in the issuing service are called the trust anchor. This setting matches well with the current authentication setting in government where a central (governmental) organisation is issuing certificates for a vast number of clients (the citizens). A single citizen

[76] See e.g. http://de.wikipedia.org/wiki/Kerberos_(Informatik)

[77] See http://tools.ietf.org/html/rfc2251

can prove his or her identity by showing the certificate (for instance a passport), whereas the instance which requires the proof does not need to trust the particular certificate of the citizen, but merely all certificates issued by the government.

In a PKI, public keys are reliably assigned to persons by means of digital signatures and a certification authority (CA). A certification authority is an organization or institution which attests that a given public key belongs to a given entity. The entity is usually a human being but could also be a machine, e.g. a web-server. The assignments are also known as (digital) key certificates. These certificates are digitally signed by the certification authority. Figure 3 exemplifies the functionality of a PKI.

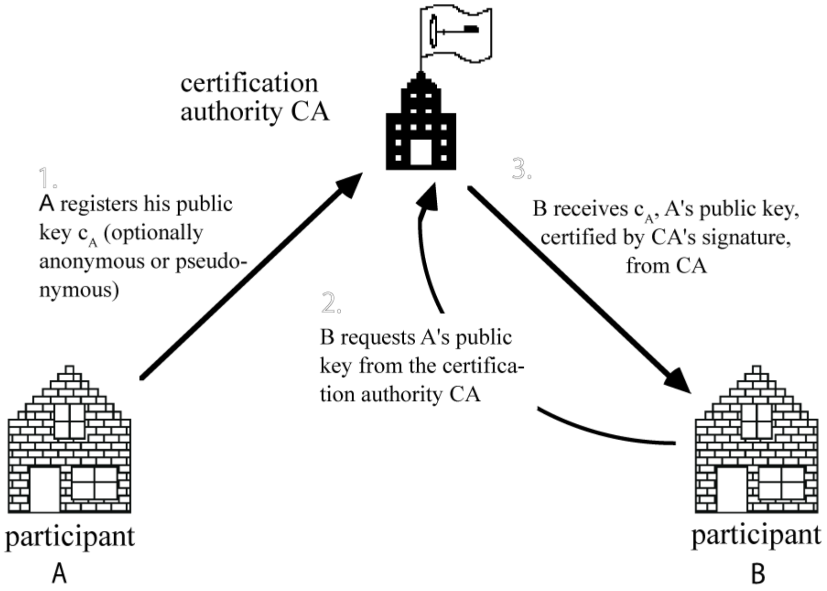
certification authority CA
1.
A registers his public key c_A (optionally anonymous or pseudonymous)
2.
B requests A's public key from the certification authority CA
3.
B receives c_A, A's public key, certified by CA's signature, from CA
participant A
participant B

Figure 3: Basic functionality of a certification authority

A typical use case for digital key certificates is to link a public key to the real identity of an entity named within the certificate. But it is also possible to issue digital key certificates for pseudonyms. In this case the certification authority in fact knows the real name of the entity for which it issued a pseudonymous key certificate. This way the CA can reveal the true identity if necessary, e.g. if required by law. Another type of certificates is the attribute certificate, which binds a set of arbitrary attributes to an entity. Thus it can be seen as a generalized form of a digital key certificate as the public key can be seen simply as an attribute of the related entity.

PKIs usually form a hierarchy of certificates where the higher level certificates are used to attest the integrity and validity of the lower level certificates. The integrity and validity of the root certificate (the certificate on the highest level) has to be checked manually, e.g. by comparing the hash value of that certificate with a publicly known value which is published, for instance in newspapers or governmental communications.

One weakness in this hierarchical concept is the large tree of implicit trust it spans. This becomes more obvious if one considers, that different certification authorities might have slightly different policies with respect of the steps required before the CA will sign a certificate. One CA might demand an official document proving the identity of the key owner before it signs the certificate while other CAs might not. To give just one example in January 2001 the company VeriSign, Inc. - one of the world's leading CAs - issued two digital certificates to a person who fraudulently claimed to be a representative of Microsoft Corporation. The issued certificates allowed the person to sign software in the name of Microsoft.[78]

Another weakness of current PKIs is the way they deal with revocation. Certificates may get lost due to accidents or burglary, for instance. The common way is to provide a certificate revocation list (CRL) in order to keep every user informed about the validity of certificates. The distribution of such CRLs, however, requires users of PKIs to be online and up-to-date whenever they intend to use a certificate, since they would need to check it before usage. This is quite inconvenient, since PKIs without revocation would not require the user to be online. In fact, there are several approaches to improve the distribution of certificates and trust chains. However, these approaches fail in being significantly better than the naive algorithms.

One measure to limit the size of a revocation list is to limit the validity of a given certificate to a certain period of time (typically one or two years). The validity period is also an appropriate measure to deal with the uncertain development of cryptanalysis. New achievements in that subject may compromise the security of certificates. The validity period is kind of a lower bound of the prediction about the time it would take to break the certificate (or key, respectively).This validity period is encoded in each digital certificate. But as now digital certificates can become outdated one has to renew them from time to time. This implies additional effort for the users of digital certificates.

All these processes - the registration process, the care of the revocation list and the renewal certificates - cause costs which need to be covered by the users of a certification authority if this authority is operated by a private company. Therefore the users typically have to pay an annual fee. Naturally this is a disadvantage of PKIs - especially if the benefits of using them do not compensate for the costs.

From a practical point of view there are even more problems, which are related to interoperability—although there exist a whole series of standards related to public key infrastructures. In 1988 the International Telecommunication Union (ITU-T) publishes the X.509 standard titled "The Directory: Public-key and attribute certificate frameworks" within their X.500 information technology related standards, which focus on open systems interconnection. Most digital certificates today (e.g. The Netherlands, Belgium, Spain, Germany; are conforming with the current version 3 of the X.509 standard.[79] This version introduces extensibility by means of profiles. One of the (if not *the*) most important profile is developed by the Public-Key Infrastructure (PKIX) working group of the Internet Engineering Task Force (IETF). The goal of this working group, which was established in 1995, is to develop standards for a public key infrastructure to be used on the Internet. The group produced more than 40 so-called Request For Comments (RFCs), which are Internet standards.

Not only the "correct" implementation of all these "standards" is a hard task—as there is always room for interpretation—but also the inherent flexibility and extensibility of X.509 supports application or domain specific extensions, which hinder a global interoperability.

[78] http://www.verisign.com/support/advisories/authenticodefraud.html

[79] See Fidis D 16.1 and Meints, M. and Hansen, M. (eds.), *D3.6: Study on ID Documents, FIDIS Deliverable,* 2006, available at <www.fidis.net>, last consulted 15 February 2009.

4.3 Credentials

Credentials, in the very meaning of the word, provide a way to attest personal properties.[80] We distinguish between the party which is issuing the credential, the party which is receiving the credential, and the party whose personal property it concerns. In line with data protection acts, we refer to the latter party as the data subject, to the first one as the issuer, and to the remaining one as the receiver. In information technology, credentials appear as signed representations of the personal properties.

The signature may be issued by the data subject herself. This is the basic case and can be enhanced to provide anonymity against the receiver, if the data subject is creating new signing keys for each credential. These signing keys can be seen as pseudonyms of the data subject. However, the data subject might create and sign any property that way. There is, particularly, no authority which assures the receiver, that the property personally belongs to the data subject. Thus, from the perspective of the receiver, authorization of credentials is necessary.

Authorization of credentials can be achieved, if there is a third party, which is willing to issue credentials with its own signing key and is believed to be trustworthy by the receiver. Then, this third party can check and sign the personal properties of the data subject. This, however, eliminates the straightforward solution for anonymity by means of using different pseudonyms for each transaction. The issuer is most likely in possession of identification data of the data subject and both, the issuer and the receiver, learn the current pseudonym of the data subject from the communication. If, in addition, both collaborate, they can identify the data subject by linking the data to the pseudonym.

The receiver could also act as an issuer, in order to obtain a simplified and more practical setting. The trustworthiness of the issuer would, in that case, not be in question for the receiver. On the other hand, such a setting would, in fact, make the collaboration between issuer and receiver most likely to happen and, therefore, finally prevent the user from acting anonymously.

In order to overcome the shortcoming of being tracked by pseudonyms, convertible credentials can be used. Such credentials can be issued for one of the user's pseudonyms and can then be converted by the user to another one of her pseudonyms. Thus, convertible pseudonyms enable the use of different pseudonyms for the communication with the issuer and the receiver, respectively. If the receiver of a credential provides a service, as it is used to be the setting for identity management, then he would not be able to link the service request of the user to the corresponding identity data from the issuing process, even if the receiver collaborates with the issuer of the credential. This isolation of the issuing process of a credential from the release process enables the user to perform different actions unlinkable to each other while providing data, which is authorized for the receiver, at the same time.

[80] Bauer, M., Meints, M., and Hansen, M. (eds.), *D3.1 Structured Overview on Prototypes and Concepts of Identity Management Systems, FIDIS Deliverable,* 2005, available at <www.fidis.net>, last consulted 15 February 2009.

4.4 Encryption Schemes & Secrecy

4.4.1 Expectations & Limitations (Scope)

Encryption can be used to provide secrecy of message contents. In a setting where a person intends to transmit messages over a channel which can be considered insecure such that the possibility of eavesdroppers cannot be eliminated, encryption can provide a way to establish a secure channel inside the insecure one such that eavesdroppers are not able to obtain communication contents. Encryption can, however, not be used to cover the existence of conversations. Eavesdroppers can, in fact, learn about who is communicating with whom and when. Thus encryption cannot conceal the communication circumstances.

An adversary who is capable of manipulating transmissions is even capable of interrupting the transmission of valid data. This change from valid to invalid data would not necessarily be recognizable by the recipient, since encryption does not protect the integrity of messages.

Confidentiality or secrecy are low-level requirements of e-Government and thus quite fundamental in that context. In Germany, protocols for e-Government are more and more standardised. So for instance, the transport protocol for e-Government in Germany is the "Online Services Computer Interface" (OSCI Transport). OSCI Transport is designed to achieve confidentiality, integrity, availability, authenticity, non-repudiation, and accountability of data transfers in several e-Government scenarios. Confidentiality is in that standard separated into two main requirements, (a) the confidentiality of content data and (b) the confidentiality of communication data. In OSCI Transport, three levels of confidentiality are considered, normal, high, and very high. As to the normal confidentiality level, no encryption is necessary. The latter two confidentiality levels require asymmetric encryption. For the confidentiality of content data, all three levels of confidentiality are possible, whereas for the communication data, only normal confidentiality is required.

In this section, we first describe the fundamentals of symmetric encryption as one of the foundations of the encryption and thus mechanisms for providing confidentiality. Then, we describe asymmetric encryption and discuss the differences with respect to symmetric schemes. And after that, we briefly describe how the most practical encryption scheme works, the hybrid encryption, which combines the advantages of both, symmetric and asymmetric encryption.

4.4.2 Quality of Secrecy and Fundamental Principles[81]

Encryption schemes differ in the degree of secrecy they can provide. We distinguish between schemes which (a) provide perfect (unconditional) secrecy, (b) are based on assumptions about computational difficulty, and (c) schemes which are cryptographically strong. While schemes in (a) can be proven unconditionally secure, so that they are secure even against attackers with unlimited computing power and storage. Schemes in class (b) are just as long considered secure as the corresponding assumptions are not violated and the attacker has only limited resources.

Encryption schemes in (c), in contrast, are not proven secure at all. They are rather considered secure, since they have been subject of serious research for years and so far no one has come up with a successful attack which could (close to universally) break one of them.

[81] See e.g. Schneier, B. and Kelsey, J.,"Security audit logs to support computer forensics", *ACM TISSEC*, vol. 2, issue 2, 1999, pp. 159-176.

Even though the schemes in (a) provide the most desirable secrecy, they are often impractical in a way that they require unreasonable much effort, for instance, for the key exchange. In practice, the selection of a particular scheme is, therefore, a trade off between effort and secrecy. In most cases, encryption schemes in classes (b) or (c) are applied in current identity management systems.

How much resources are needed to break a given scheme of class (b) or (c) depends on the parameterization of that scheme. Note that choosing more secure parameters often means to increase the effort of encryption and decryption as well. Fortunately the relation between the increase of effort for regular use and attackers is non-linear, meaning that making a given scheme more secure (and thus increasing the effort for encryption and decryption) typically implies much more effort for attacks.

Given this and considering the family of Moore's Laws, which estimate that computational resources (like computing power, storage bandwidth etc.) are growing exponentially, it is necessary to adjust the parameters of cryptographic schemes used on an annual or at least biannual base. This might introduce all kinds of organizational problems as update procedures have to be specified accordingly, e.g. how to proceed if the parameters used for encryption of a document are not secure anymore? Either such documents have to be re-encrypted with secure parameters. That would have to be done before the previous parameters are actually considered insecure and requires the destruction of all outdated copies of such documents as well. Or the document has to be considered as leaked to the public when the parameters become insecure.

Talking about secrecy of cryptographic keys in relation to communicating parties and their knowledge one can distinguish between cryptographic algorithms which use *symmetric keys* and algorithms which use *asymmetric* ones. The term symmetric refers to the fact that the same key is used for encryption and decryption. This is quite convenient, once the key is distributed to everyone who should be able to encrypt and decrypt. A drawback of symmetric encryption schemes is that everyone who obtains the key gains the capability to encrypt *and* decrypt. In particular, a separation of duties is impossible in the purely symmetric case. For asymmetric encryption schemes, there are always two keys belonging to each other, one for encryption and the other one for the decryption. This allows to publish the encryption key without giving the secrecy of encrypted documents away and thus significantly simplifies the key exchange problem.

One of the most fundamental principles of modern cryptography was defined by Auguste Kerckhoffs 1883[82] and is now known as Kerckhoffs' principle: the security provided by a given cryptographic algorithm should not depend on the secrecy of the algorithm itself but on the secrecy of cryptographic keys. This implies that any measure which depends on "security by obscurity" can be seen as untrustworthy by definition—moreover history teaches that such attempts, in fact, lead always to insecure solutions. Note that the level of protection offered by a cryptographic algorithm depends on the size of the keys used. In general a large key implies better protection against attacks on the secrecy.

An important point when implementing cryptographic schemes and protocols is the fact that security needs some kind of "trusted anchor", i.e. one cannot achieve protection within a

[82] Kerckhoffs, A., 'La cryptographie militaire*' [*The military cryptography*]*, *Journal des sciences militaires*. vol. 9, 1883, pp. 5–38 and pp. 161–191.

completely untrusted environment. Although the Trusted Computing Group (TCG)[83] strives to minimize the root of trust with the help of specialized hardware (so-called Trusted Platform Modules) and related software, today's reality is, that the root of trust comprises the whole end-system used (e.g. the client personal computer or the server machine). As the names "trusted anchor" or "root of trust" imply, one has to trust these components. If this assumption does not hold the whole security chain will be broken. This fact is especially relevant if one factors in the well known problems of standard PCs with standard operating systems, where faulty implementations open up all kinds of security holes, which are actively exploited by malicious software like viruses or Trojan horses.

4.4.3 Symmetric Encryption Schemes

In order to establish a secure channel by encryption, all involved parties have to exchange keys. The very case is that the key for encryption and decryption is the same. We call such encryption schemes *symmetric* (see figure 4).

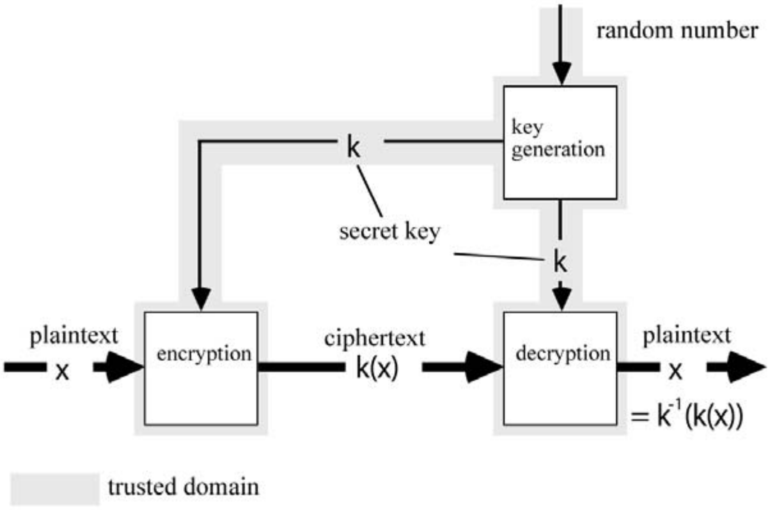

Figure 4: Symmetric encryption scheme

For instance, all schemes which provide perfect secrecy are symmetric. In fact, there is just one scheme, which provides perfect secrecy and it is known as *one time pad*. It is a substitution cipher, that is, for encryption, a cipher text sign is substituted for each plaintext. The selection of each cipher text sign depends on the key and the plaintext. In order to achieve unconditional security, the key must not be iterated and, thus, needs to be as long as the plaintext. Furthermore, for each pair of plaintext and key, there must not be any other pair which is mapped to the same cipher text sign. Additionally, the key must not be used more than one time, it must be perfectly secret against adversaries before, during and after the usage, and needs to be shared between the sender and the recipient, at the same time.

Besides substitutions, permutations and transpositions are building blocks of symmetric encryption schemes. Apart, each of these can be broken with reasonable effort by means of identifying key iterations or by means of supplementary knowledge about the plain text.

[83] http://www.trustedcomputinggroup.org/. This also contains the specifications or standards of this group which are frequently updated. https://www.trustedcomputinggroup.org/specs/TPM/

Therefore a combination and iteration of substitutions, permutations, and transpositions is necessary to construct a cryptographically strong cipher.

Historically the Data Encryption Standard (DES) was one of the first widely adopted symmetric encryption algorithms because it was standardized in 1997 by the American National Bureau of Standards (NBS)—today known as National Institute of Standards and Technology (NIST)—making its use an obligation for certain types of classified communication in the public and private sector. Today DES can be seen as insecure as it can be completely broken with reasonable effort[84].

The successor in the line of cryptographic strong symmetric encryption schemes is the Advanced Encryption Standard (AES). It was standardized by NIST in October 2000. AES has been designed to overcome the shortcomings of DES and to adapt to newer approaches in cryptanalysis. The algorithm supports key sizes of 128, 192 and 256 bits. Given that a key size of 128 bit can be seen as secure today, it can be assumed that AES will be secure for the next decade(s).

The design of AES allows very efficient and cost-effective implementations in hardware as well as software. Particularly this hardware option is attractive for industry, since hardware such as cryptographic co-processors can significantly speed up calculations.

4.4.4 Asymmetric Encryption Schemes

Encryption schemes, which require different keys for encryption and decryption are called *asymmetric*. They are widely used, since the separation in encryption and decryption key allows making the encryption key public, while keeping the decryption key secret. This, indeed, only works as long as the secret decryption key cannot be obtained from the public encryption key. It is, therefore, not necessary anymore to *share* a secret key with each communication partner, which significantly reduces the number of necessary key exchanges. A person may distribute one and the same encryption key to all communication partners without risking that anyone of these communication partners can eavesdrop the communication contents, which were encrypted by another partner with the same encryption key. Asymmetric schemes are, therefore, also known as public-key cryptography. Figure 5 depicts the basic functionality of an asymmetric encryption scheme.

[84] A specialised hardware called Copacobana (www.copacobana.org) costs around 10.000$ and can break DES within 6.4 days on average.

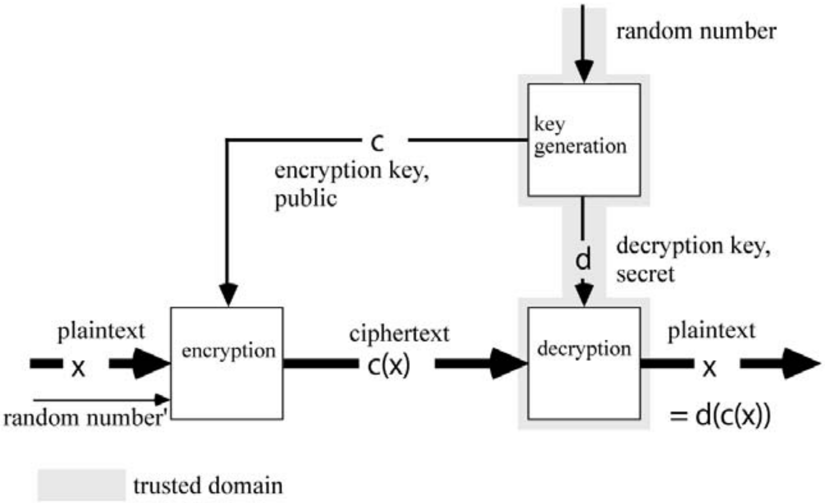

Figure 5: Asymmetric encryption scheme

Unfortunately, these rather organizational advantages of asymmetric schemes go along with the inconvenience of far more complex computations compared to symmetric schemes. Additionally, there is no asymmetric scheme which provides perfect secrecy, since the secret decryption key is, in fact, related to the public encryption key. It is rather the question how complex (and therefore hard) it is for an adversary to obtain the secret decryption key. In common asymmetric schemes, the complexity of a successful attack on the secret key is proven to be as complex as, for instance, the factorization of a large number or the discrete logarithm problem. Such schemes, like for instance ElGamal, are therefore in class (b) and provide secrecy as long as the underlying mathematical problem cannot be resolved by any adversary in reasonable time.

Other asymmetric schemes, like the well known RSA for instance, are based on mathematical problems which are not proven computationally difficult. RSA, however, has been analyzed for long and so far no universal attack has been published. Thus, RSA is considered to be in class (c).

Note that the level of protection offered by symmetric and asymmetric algorithms cannot be easily compared by just comparing the sizes of the keys used, e.g., a 128 bit key of AES can be considered to be secure whereas a 1024 bit key of RSA might not constitute a reasonable level of protection given the computing power available today. By rule of thumb a symmetric key should have at least 128 bits and an asymmetric one should have at least 2048 bits.[85]

4.4.5 Hybrid Encryption Schemes

In order to combine the organizational advantages of asymmetric encryption schemes with the speed-up of symmetric schemes, it is common to use asymmetric schemes and symmetric schemes in combination. This combination is called a hybrid encryption scheme. In hybrid

[85] This rule of thumb also is supported by information security agencies, e.g. the German Federal Office for Information Security (recommendations: 100 bit symmetric keys which practically means 128 bit key length and 2048 bit key length for the asymmetric RSA-algorithm). See http://www.bsi.de/gshb/deutsch/m/m02164.htm and http://www.bundesnetzagentur.de/enid/8a5a69b5d3a2d6480990667c8e7b5ffd,0/Veroeffentlichungen/Algorithmen_sw.html

schemes, the asymmetric scheme is used to exchange a relatively small symmetric session key, which is then used as the encryption key in the fast symmetric scheme.

4.5 Summary

In this chapter basic technical approaches to ensure authenticity and authorisation of users in administrative procedures and IMS were presented. The following instruments play a major role in the context of eGovernment:

- Three factors of authentication, namely knowledge, possession and biometrics are used in governmental IMS to implement various levels of authentication needed in different administrative procedures.
- Centralised repositories of reference data for users, typically stored in LDAP directories; in the context of cryptographically based methods such as electronic signing and electronic signatures Public Key Infrastructure (PKI) plays an important role as centrally accessible reference.
- For ensuring confidentiality of data e.g. when transferring authentication and authorisation information via (insecure) networks or storing them on media cryptographic techniques are used. They are also important in the context of protocols such as Kerberos, and Credentials.
- Important cryptographic techniques include symmetric and asymmetric cryptographic algorithms as well as hash algorithms. Fundamental properties including the quality of secrecy of broadly used cryptographic algorithms were discussed in this chapter.
- Where federated IMS are used, markup languages such as the eXtensible Access Control Markup Language (XACML) play an important role.
- Also of importance in the context of governmental IMS are considerations concerning trust. Who is going to trust who based on what? Taking this aspect into consideration in an appropriate way may improve the acceptance of governmental IMS and services relying on these IMS significantly.

5 Advanced technical Approaches

In this chapter two aspects of identity management going beyond compliance with data protection legislation will be introduced and discussed, namely options for identity management in networking infrastructure and advanced approaches for secure audit logging. Audit logging is a relevant task in the context of (governmental) IMS. This chapter deals with two questions:

1. How can identity management be optimised, taking into consideration that identifiers used in the networking infrastructure principally enable a certain level of identifiability? This question may be relevant in the context of governmental services that require anonymity, such as e-voting or e-petitions. Technical approaches to improve anonymity on different OSI layers are presented in section 5.1 and discussed. However, they will be directly applicable in special governmental application scenarios only.
2. How can audit logging be implemented in a way that avoids the shortcomings of today's syslogging services and products? Examples for these shortcomings are: (a) lack of audit acceptability of log data, as IT administrators, whose activities are logged, are able to change log data without leaving any evidence and (b) vulnerabilities in protocols used to transfer audit data to a logging service, so that these data can be manipulated during the transport via the network. An advanced approach for secure logging is presented in section 5.2.

5.1 Management of Identities in Networking Infrastructure

Authentication and identification are almost omnipresent requirements in eGovernment. In most cases it is necessary to make sure that the communication partner is a particular citizen or belongs to a particular group of authorized citizens. However, there are also applications where the acceptance of (governmental) services depends on assurance of anonymity. For instance, the access to a database might be granted to citizens without authentication, if the access log alone might be used for discrimination (or might have negative consequences for citizens). Another popular example is the access of discussion forums where sensitive topics are discussed. These might not be the standard examples for services in eGovernment, since they are not describing organisational issues of governments in the first place. Though, many governments have shown responsibility even for those issues when they put data protection acts in place. The benefit of the technologies presented in this section is the transfer of control over the identity from the data controller (that might be the government) to the data provider (that might be the citizen). The technologies support the construction of a trustful environment where the data controller can just acquire as much personal data as the citizens are willing to disclose. In consequence, these technologies are merely a foundation of every (governmental) service where authentication or identification is not required from the start.

A detailed overview about identity obfuscation techniques can be found in Cvrcek and Matyas[86] and Alkassar[87]. In this section, we outline the most important identity obfuscation techniques and discuss their use in privacy-enhancing IDM.

[86] Cvrcek, D., Matyas, V. (eds.), *D13.1: Identity and impact of privacy enhancing technologies*, FIDIS Deliverable, 2007, available at <www.fidis.net>, last consulted 15 February 2009.

5.1.1 Private Information Retrieval

We start with an example where the goal is to access a database, but to conceal what data is actually requested. In this setting the access to the database might be a public service whereas citizens make requests. The trust and thus the use of the service may depend on the fact that the database provider is not able to trace which citizen (or user of the database) requested a particular record. A simple solution for building up that kind of trust is to allow just a single request which delivers the entire database as a result. Since every citizen would (have to) request the entire database and select the record of interest locally, i.e. apart from the control of the database provider, there is no way for the data controller to decide which citizen accessed which record. This solution is not efficient, though. A discussion of more efficient approaches can be found in Cvrcek and Matyas.[88]

5.1.2 Broadcast Networks

Apart from protecting particular requests in database services, it would be also an asset to protect the identity of the requester as long as the service does not require authentication. This protection is subject of this and the following sections.

A simple example for networks that protect the anonymity of network users is a broadcast network. It can be compared with radio broadcast where there are several stations sending on different frequencies and potentially anyone in the closer area around the sender is able to receive the broadcast. Besides anyone in the closer area may decide whether to switch on the receiver or not. Thus, the sender may publish information, but it is impossible for the sender to decide which information has been received by which receiver.

This rather old scheme exists in modern networks as well. These implementations are usually comparable to the Citizens' Band, i.e. anyone may become a sender. The use in governmental services might be bounded to cases where subscribing a particular information channel might become a reason for discrimination, though distributing the information is essential, for instance for proper health care. The technological details are described in Cvrcek and Matyas.[89]

5.1.3 DC Networks

Plain DC Networks, also known as superimposed sending, preserve sender anonymity. The technological details can be found in Cvrcek and Matyas.[90] A major application of this technology is whistle-blowing, but it is also well suited for any other application where the sender would have to expect punishment, if his or her identity is revealed. The major drawback is that replies to a previous sender are not possible without additional means. The DC network is merely working like a huge billboard where everyone can publish and it is not possible to find out who published a particular message.

[87] Alkassar, A. and Hansen, M. (eds.), D3.8: Study on protocols with respect to identity and identification – an insight on network protocols and privacy-aware communication, FIDIS Deliverable, 2008, available at <www.fidis.net>, last consulted 15 February 2009.

[88] Cvrcek, D., Matyas, V. (eds.), *D13.1: Identity and impact of privacy enhancing technologies*, FIDIS Deliverable, 2007, available at <www.fidis.net>, last consulted 15 February 2009.

[89] Cvrcek, D., Matyas, V. (eds.), *D13.1: Identity and impact of privacy enhancing technologies*, FIDIS Deliverable, 2007, available at <www.fidis.net>, last consulted 15 February 2009.

[90] Cvrcek, D., Matyas, V. (eds.), *D13.1: Identity and impact of privacy enhancing technologies*, FIDIS Deliverable, 2007, available at <www.fidis.net>, last consulted 15 February 2009.

5.1.4 MIX Networks

Mix networks can be used to achieve receiver anonymity, similarly to broadcast networks, sender anonymity, similarly to DC networks, and a combination of both, i.e. sender and receiver anonymity. A major difference to the previous schemes is that addressing is an inherent part of the scheme. Thus it is possible to establish a connection between two communication partners whereas no one needs to know more than a anonymous address of the other one. In eGovernment applications, however, mixes are merely applied as practical solution for achieving sender anonymity. There are several mix implementations which are working well in the Internet, such as TOR, AN.ON, MIXMASTER, etc. The technological details are summarised in Cvrcek and Matyas.[91]

5.2 Secure logging

Logging is a building block for *a posteriori* validation. First, it allows to check whether security properties that cannot be enforced by an execution monitor have been violated.[92] For example, obligations, i.e. actions that are triggered once access is granted to an object, are not enforceable by an execution monitor. In consequence, the fulfilment of obligations has to be verified a posteriori by audit. Moreover, reliable and secure logs are necessary to control the behaviour of all kind of privileged users.

Further, a reliable identification of violation is important in settings where users give personal data “away” to communication partners, such as in eGovernment applications. Users are not longer able to technically control how their data is used. In this case, a reliable audit log may lead towards more transparency for users. Thus, mechanisms to reliably identify violations are necessary. Audits, i.e. the analysis of logs, can identify such violations. However, an audit has no validity, if the log it is based on, could have been changed or otherwise corrupted, either by an external attacker or users with privileged rights, such as system administrators. In other words: if the log might be “incorrect”, the audit result has no significance.

Therefore, authentic and reliable logging mechanisms are needed. This subsection presents the BBox, a component for secure logging developed at the University of Freiburg.[93] The BBox is a component of an audit architecture that acts as a black box of the system. It provides a secure logging mechanism that ensures that the activity of the system, represented in terms of log data, is recorded in a tamper-evident, confidential manner.[94] The idea behind the BBox is the realisation of a system’s digital black box[95] mimicking in a number of aspects its counterpart in aircrafts, i.e. the flight recorders. Without interfering with aircrafts’ operation, a flight recorder records transcripts of specific events as they happen in the aircraft.

[91] Cvrcek, D., Matyas, V. (eds.), *D13.1: Identity and impact of privacy enhancing technologies*, FIDIS Deliverable, 2007, available at <www.fidis.net>, last consulted 15 February 2009.

[92] Halmen, K. W., Morrisett, G. and Schneider, F. B., *Computability Classes for Enforcement Mechanisms*, 2006, available at <http://www.cs.cornell.edu/fbs/publications/EnfClassesTR2003-1908.pdf>, last consulted 10 March 2009.

[93] Accorsi, A., *Automated Counterexample-Driven Audits of Authentic System Records*, PhD Thesis, 2008, available at <http://www.freidok.uni freiburg.de/volltexte/6048/pdf/Diss_RafaelAccorsi.pdf>, last consulted 10 March 2009.

[94] Accorsi, A., *Automated Counterexample-Driven Audits of Authentic System Records*, PhD Thesis, 2008, available at <http://www.freidok.uni freiburg.de/volltexte/6048/pdf/Diss_RafaelAccorsi.pdf>, last consulted 10 March 2009.

[95] Oppliger, R., and Ritz, R., ‘Digital evidence: Dream and reality’, *IEEE Security and Privacy* , vol. 1, issue 5, 2003, pp.44-48.

Like the logging protocols of Schneier and Kelsey,[96] the BBox uses hash chains and evolving keys to prevent the tampering of the logfile. The BBox extends the existing protocols by providing in addition security during the transmission phase from the collector to the logging entity. A detailed comparison of the BBox, the protocols based on Schneier and Kelsey and those of the "syslog-family" is given in Accorsi.[97]

We first analyze the requirements to "secure" logging and describe the threats to authenticity and security of a log file. Further, the components of the BBox and the basic cryptographic functions for the secure logging mechanism - evolving keys and hash chains – are described. An overview and a comparison on existing logging protocols can be found in Accorsi[98] and details concerning the logging mechanism are given in Accorsi.[99]

5.2.1 Requirements for Secure Logging

Irrespective of its envisaged application, it is imperative for any service wishing to employ log data to ensure that it is reliable in the first place. In particular, if log data is to be used as a parameter to determine compliance with policies or needs to be submitted as legal evidence, two distinct, albeit closely related issues must be addressed, namely *admissibility* and *credibility.*[100] Admissibility prescribes a set of criteria to judge the acceptance of log data to the court or to an auditor on its behalf. Being of a strictly legal nature, the following does not focus on the admissibility properties of log data. Credibility requires log data to be authentic and, hence, reliable enough to influence the outcome of a proceeding, decision or audit. Put another way, credibility states that log data is a faithful representation of the events communicated by the devices. Therefore, irrespective of how admissibility is defined, the higher the credibility of log data is, the greater its admissibility. From a technical viewpoint, unlike admissibility, credibility of log data – expressed in terms of its authenticity – is a property that depends on how log data is stored in a system and the kinds of access that are performed upon it. Thus, credibility can be to some extent technically enforced in computer systems. To this end, one first defines what authenticity means in the context of logging and then analyses the extent to which it can be provided in computing systems.

Records in general and log data in particular are labelled *authentic* if the data they support have not been altered or otherwise corrupted over time.[101] Considering the standard protection

[96] Schneier, B. and Kelsey, J.,"Security audit logs to support computer forensics", *ACM TISSEC*, vol. 2, issue 2, 1999, pp. 159-176.
[97] Accorsi, A., Digital Evidences based on Log Data: What Secure Logging Protocols Have to Offer?, submitted to COMPSAC 2009.
[98] Accorsi, A., Digital Evidences based on Log Data: What Secure Logging Protocols Have to Offer?, submitted to COMPSAC 2009.
[99] Accorsi, A., *Automated Counterexample-Driven Audits of Authentic System Records*, PhD Thesis, 2008, available at <http://www.freidok.uni freiburg.de/volltexte/6048/pdf/Diss_RafaelAccorsi.pdf>, last consulted 10 March 2009.
[100] Kenneally, E., 'Digital logs – Proof matters', *Digital Investigation*, vol.1, issue 2, 2004, pp. 94–101.
[101] Sanett, S., and Park, E., 'Authenticity as a requirement of preserving digital data and records', *IASSIST Quarterly*, vol. 24, issue 1, 2000, pp.15-18.

goals for computer security,[102] authenticity for log data boils down to stating that it fulfils *data integrity* properties. Unlike other kinds of data, there is no published definition of, nor consensus regarding the meaning of integrity for log data. In this work, log data stored in a log file fulfils integrity if it is *accurate*, i.e. entries in the log file have not been modified, *complete*, i.e. entries have not been deleted from the log file and *compact*, i.e. entries have not been illegally appended to the log file. Data integrity is central to the authenticity of log data, but not sufficient. Given a communicated event, one also needs to ensure *origin integrity*, i.e. that the device communicating the event is known and authorised to notify events to the collector. Moreover, unauthorised subjects should not read log data, otherwise these subjects would be able to compose messages and send them to the collector, thereby posing a threat to the authenticity of log messages. Also, in the eGovernment setting where individuals' privacy is relevant too, the confidentiality of events regarding personal data is essential as well. Hence, the *confidentiality* of log data in transit and at rest must be provided.[103] Taking stock, the authenticity of log data is characterised in terms of its *data integrity, origin integrity* and the *confidentiality* of log data during the transmission and at rest. Notice that the definition of authenticity does *not* entail the correctness of log data, i.e. that data sent by the devices faithfully correspond to the events happening in the system. This work assumes correctness, while the BBox ensures that the audit considers authentic, unaltered log data.

5.2.2 Threats to the authenticity of a log file

The authenticity of log data can be attacked using possibly a mixture of the following attack strategies. This list merely reports on the most relevant attack strategies directly related to the security requirements; a more detailed account is given in Gallegos et al., Maier and Mercuri.[104]

- **Replay of log messages**. The attacker records a set of log messages sent by a device to a collector and, at a later time point, resends these messages to the collector, possibly in a refreshed form. This attack strategy violates the requirement of compactness and that of accuracy.
- **Manipulation of log data**. This attack can be upon the log data being sent from a device to the collector or log data already stored at the collector. In the first case, the attacker has access to the communication medium and can modify data during the transmission. In the second case, in gaining access to the log file, the attacker can

[102] See Bishop, M., *Introduction to Computer Security*, Addison-Wesley, Boston, 2005 and Rannenberg, K., Pfitzmann, A., and Müller, G., 'Sicherheit, insbesondere mehrseitige Sicherheit', in Müller, G., and Pfitzmann, A. (eds.), *Mehrseitige Sicherheit in der Kommunikationstechnik*, pp. 21-30, Addison-Wesley, New York, 1997 and Pfitzmann, A. and Hansen, M., *Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management - A Consolidated Proposal for Terminology, v0.31*, available at <http://dud.inf.tu-dresden.de/Anon_Terminology.shtml>, last consulted 15 February 2009.

[103] When personal data are involved, not only does confidentiality have to be ensured, ate least, when considering Directive 95/46.

[104] Gallegos, F. et al., *Information Technology Control and Audit*, Boca Raton: Auerbach, Florida, 2004 and Maier, P., *Audit and Trace Log Management*, Boca Raton: Auerbach, Florida, 2006 and Mercuri, R., 'On auditing audit trails', *Communications of the ACM* , vol. 46, issue 1, 2003, pp.17-20.

modify entries containing log data. Attacks of this type violate the accuracy of log data.

- **Deletion of log data**. The attacker captures log messages before their receipt at the collector. Similarly, the attacker may delete log entries of the log file after breaking into the system. Attacks of this type violate the compactness of log data.

- **Read access to log messages.** During the transmission, the attacker intercepts and reads messages sent in clear-text over the network, or decrypts them using the appropriate encryption keys. Similarly, log data stored in a collector may become accessible to an attacker. This violates the confidentiality requirement for log data.

The threats above are directly related to the authenticity of log data. More elaborated attacks that subsequently violate authenticity could encompass:

- **Device impersonation.** The attacker takes over a device and starts to send (forged) log messages to a collector or at least observe the log messages being sent to the collector. Hence, this attack also affects the confidentiality of log messages. Note that this is not a violation of origin integrity, as the sensing device is still the one authorised to send log messages. Impersonation and its exploitation primarily corrupt the correctness of log data.
- **Availability of log service**. The attacker, possibly but not necessarily impersonating one or more devices, floods the collector with void log messages. In consequence, legitimate messages sent by the devices are not logged. The attacker can then use this setting to attack the computing environment, since due to the denial of service, messages delivering evidence about these attacks are not protocolised by the collector.

These attacks can be carried out irrespective of the underlying logging mechanism used to guarantee the authenticity of log data, but will not be considered further. Guidelines for circumventing such attacks can be found.[105]

5.2.3 The BBox: An architecture for secure logging

The architecture of the BBox, its components and the information flows happening therein are depicted in Figure 6. Its functionality can be distinguished in two logical units: first, the “recording unit” in charge of logging communicated events in a secure manner consists of the log message handler and the entry append handler; second, the “retrieval unit” responsible for the generation of log views, which encompasses the log view handler and the entry retrieval handler. Put another way, these logical units constitute the (sole!) input and output interfaces of the secure log file.

[105] See for example Glynos, D., Kotzanikolaou, P., and Douligeris, C.,’Preventing impersonation attacks in MANET with multi-factor authentication’, *Proceedings of the 3rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks,* IEEE Computer Society Press, 2005, pp. 59-60 and Howard, M., and Leblanc, D., *Writing Secure Code*, Microsoft Press, Redmond, 2001 and Kent, K., and Souppaya, M., *Guide to Computer Security Log Management*, 2006, available at <http://csrc.nist.gov/publications/nistpubs/800-92/SP800-92.pdf>, last consulted 10 March 2009.

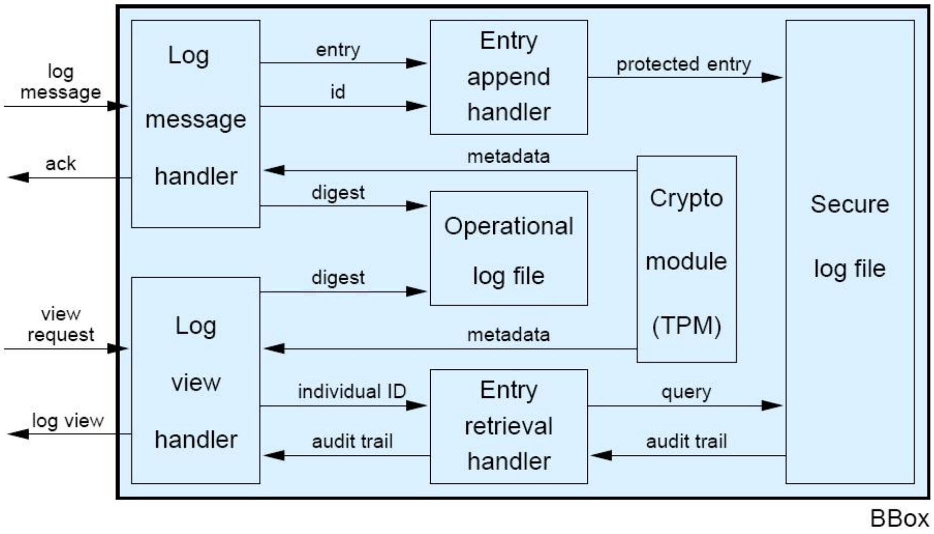

Figure 6: Components of the BBox

The BBox architecture consists of the following components:

- **Log message handler** (LMH). The LMH receives incoming log messages sent by the device and carries out an integrity check to determine whether the contents of the message are eligible to be appended to the secure log file. This integrity check is necessary, for an attacker may try to replay old messages or alter them during the transmission between the device and the BBox. To carry out the integrity test, the LMH requires metadata, such as current clock time, which is obtained from the crypto module. A digest recording this activity is appended to the operational log file.

- **Entry append handler** (EAH). If a log message passes the integrity test carried out in the LMH, its payload (event) and individual ID are given to the EAH, which transforms these components in a protected entry for inclusion in the log file. To this end, a secure logging mechanism is employed to produce an entry that accounts for integrity properties.

- **Secure log file**. This is the container where events are securely recorded after being prepared by the EAH.

- **Log view handler** (LVH). Only authorised individuals may have access to log views, i.e. audit trails encoding the activity of the system related to individuals' data items. The LVH controls the disclosure of collected data by receiving view requests, authenticating individuals and passing on the necessary information to the entry retrieval handler. Eventually, it also prepares the resultant audit trails and sends it to the requesting individual. A digest recording this activity is appended to the operational log file.

- **Entry retrieval handler** (ERH). The ERH receives the ID of the requesting individual and produces a corresponding query over the secure log file. To this end, information in the crypto module is needed in order to allow the decryption of the corresponding log entries. The resultant audit trail is then handed over to the LVH.

- **Crypto module**. This is a trusted computing module (TPM) responsible for, among others, storing the cryptographic keys, providing metadata and a basis for remote attestation.

- **Operational log file**. The functioning of the BBox is recorded in an "event sink", write-once, read-many log file. Events recorded in this file include the decision whether a log message has passed the integrity test, information as to whom a log view has been issued and service disruptions, such as (re-)initialisations and shutdowns. This kind of information is important for administrative, as well as legal reasons to assert the appropriate functioning of the BBox.

5.2.4 Building blocks for an authentic log file

The operational mode of the BBox encompasses three main phases: the *initialisation* phase, at which an offline BBox is prepared to be put online; the *online* phase, at which incoming log messages and view requests are processed; and the *shut down* phase, at which the BBox is taken offline. In this deliverable, the details of the operations and the phases will not be described further. For further reading on how entries can be appended to an existing log file ensuring the authenticity of log entries, see e.g., Fidis D14.6., and a detailed description of the BBox can be found in Accorsi.[106]

We present here the techniques for authentic logging: *Evolving cryptographic keys*[107] are used to encrypt the payloads of log messages and *hash chains*[108] are used to intertwine the entries in the log file.

Evolving keys. In contrast to usual cryptography, where keys are kept the same over time, in evolving key cryptosystems keys change, or evolve, from time to time, thereby limiting the damage that can result if an attacker learns the current cryptographic key. In the BBox, each payload is encrypted with a unique key K_i derived from an evolving *entry authentication key* G_i and the entry identifier I by hashing these two values. The entry authentication key G evolves for each entry. Hence, the keys K are independent from each other, so that if the attacker obtains a particular K_i, he can neither obtain K_{i-1} nor K_{i+1}. To make it harder for

[106] Accorsi, A., *Automated Counterexample-Driven Audits of Authentic System Records*, PhD Thesis, 2008, available at <http://www.freidok.uni freiburg.de/volltexte/6048/pdf/Diss_RafaelAccorsi.pdf>, last consulted 10 March 2009.

[107] Franklin, M., 'A survey of key evolving cryptosystems', *International Journal of Security and Networks*, vol.1, issue 1-2, 2006, pp.46-53.

[108] Lamport, L.,'Password authentication with insecure communication', *Communications of the ACM* , vol. 24, issue 11, 1981, pp.770-772.

attackers to decrypt the payload of messages, the entry index I is stored as a hash value, so that even if the attacker obtains some G, he still has to obtain I to gain access to the payload.

To provide for such guarantees, the following must be considered in employing evolving cryptographic keys. First, the function producing new entry authentication keys must be one way and, thus, provide for the independence of key values: given the key G_j, for $j \neq 0$, it is infeasible to compute any of the previous keys G_i with $0 \leq i < j$. Second, the initial value G_0 upon which all the other keys are based must be kept secret. Third, in evolving the key, the computation of the next values G must irretrievably overwrite the previous key. These aspects are considered when developing the secure logging mechanism.

Hash Chains. A hash chain is a successive application of a cryptographic hash function to a string. By knowing the initial value and the parameters with which the chain is generated, the integrity of an existing hash chain can be checked for broken links by recomputing each element of the chain. (Alternatively, it is also possible to check the integrity of contiguous regions of the chain instead of its whole.) The BBox uses hash chains to create an interdependency between an entry i and its predecessor $i - 1$, thereby linking entries to each other. Moreover, since elements of the hash chain can as well be seen as a checksum of the involved parameters, in computing an element of the chain and comparing it with the existing link, the BBox can also assert whether the corresponding entry has been modified or not. Hence, tamper evidence for integrity properties are accounted for. Together, these cryptographic techniques lay the foundation for our secure logging mechanism. The resultant entries are triples $\mathrm{E_i} = (HI_i, \{\mathrm{P_i}\}_{\mathrm{K}_i}, \{HC_i\}_{\mathrm{K}^{-1}_{BBox}})$, where HI_j is the hashed index of the entry, $\{P_i\}_{K_i}$ is payload P_i encrypted with the unique key K_i and HC_i is the ith link of the hash chain, which is signed with the private key of the BBox.

5.3 Summary

In the context of governmental IMS management of identities based on information collected from the networking infrastructure is not a traditional application scenario. However, in the context of e-voting and e-petitions, and in the context of the work of secret services options for the management of identities established this way may become or is already important.[109]

There exists a lot of possibilities and techniques for identity obfuscation when transferring data over networks. All of them have their specific advantages and disadvantages. Broadcast, Private Information Retrieval and DC-Networks offer a very high level of protection even against powerful adversaries. However, they need many resources in terms of computational power and bandwidth. They also do not scale very well. Simple proxies on the other hand require only little effort, but cannot provide much protection, especially because one has to trust entirely the operator of a given proxy. Today Mix networks seem to be the only approach that offers reasonable protection at reasonable costs and may find their application in certain governmental services. Nevertheless, Mix-based systems like AN.ON or Tor are still under development and have not achieved a quality of service level appropriate for the masses. Moreover, the legal situation in Europe, namely different implementations of the data

[109] The development of TOR was supported for many years by U.S. governmental research institutions, see http://de.wikipedia.org/wiki/Tor_(Netzwerk)

retention directive, hinders the successful implementation of identity obfuscation technologies for the mass market.

In the context of governmental IMS audit logging is an important instrument for an a posteriori validation of authentications carried out through and administrative activities carried out within the IMS. An authentic and integer audit log is the necessary precondition for valid audit results. In this context syslogging based on various audit logging protocols and products is an established approach. However, these approaches show vulnerabilities to certain man-in-the-middle-attacks.

The BBox[110] provides a secure logging mechanism that ensures that the activity of the system, represented in terms of log data, is recorded in a tamper-evident, confidential manner. In this chapter the architecture and the components of the BBox are described together with the two basic cryptographic techniques the BBox is built on. In case the presented architecture proves to be mature and enters the market of ICT products, so that interfaces for other products and systems become available, the BBox approach may soon be valuable in the context of all eGovernmental services and related infrastructures.

[110] Accorsi, A., *Automated Counterexample-Driven Audits of Authentic System Records*, PhD Thesis, 2008, available at <http://www.freidok.uni freiburg.de/volltexte/6048/pdf/Diss_RafaelAccorsi.pdf>, last consulted 10 March 2009.

6 Combined technical and organisational Approaches: Privacy Policy Handling

In Europe, a privacy or data protection policy is understood as the declaration about which organisation processes which personal data of its clients, how, in the context of which government or business process in accordance with the corresponding (national) data protection act (summarised based on [111]). They need to be publicly accessible documents. From the compliance with data protection legislation a number of tasks in the context the enforcement of the privacy policy emerge. They include the implementation of data protection principles in all phases of the life cycle of administrative procedures, namely the planning phase, implementation or building and the operations phase. In addition, external influencing factors need to be taken into consideration, e.g. changes in technical infrastructure, changes in operative requirements in the context of administrative procedures and changes in legislation.

Analysing these requirements, one can observe that they are very similar to those to be found in the context of information security.[112] In both cases a catalogue of (technically oriented) measures, embedded in an organisational framework, seems to be a good solution.

This chapter describes the outline of an organisational framework for the handling of privacy policies and presents technical approaches to support this handling. As already outlined, in the enforcement of privacy policies concerning administrative procedures, IMS play an important role. Parts of the policy, especially the data handling and access control policy, may partially or completely be implemented in the IMS (see section 6.3).

6.1 Privacy Policy Handling

Generally speaking, in the context of privacy policy handling a Data Protection Management System (DPMS) similar to an Information Security Management System (ISMS) as defined in ISO/IEC 27001 is an effective approach to support all activities in the life cycle of the policy (Meints 2007). A DPMS contains the following components:

- Function bearer(s) that are qualified, able to enforce the privacy policy within the organisation and carry out their role not in conflict with other roles that are assigned to them; typically a central function bearer is the so called Data Protection Officer (DPO).
- A process framework[113]
- Documentation, among them the data protection policy, a data protection / security concept and operative documentation (e.g. an inventory of procedures, process handbooks, operational advice for employees, guidelines, documentation of the implementation of technical and organisational (security) measures, data protection and security management reports).

In the following section a focus will be put on the process part of the DPMS.

[111] See http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2004/wp100_en.pdf.

[112] Meints, M., 'Datenschutz durch Prozesse' [Data protection through processes], *Datenschutz und Datensicherheit,* vol. 31, issue 2, 2007.

[113] See e.g. Müller, G. and Wohlgemuth, S. (eds.), *D14.2: Study on Privacy in Business Processes by Identity Management, FIDIS Deliverable,* 2007, available at <www.fidis.net>, last consulted 15 February 2009, pp. 42-47.

6.2 Organisational measures for the enforcement of privacy policies

Concerning the data protection policy the process framework needs to take into consideration:

1. Policy initialisation: Policy definition, setting into force and set up of a DPMS
2. Policy implementation, covering all life cycles of administrative procedures and supporting applications / ICT infrastructure; relevant cycles are at least[114]:
    a. Planning phase
    b. Implementation phase
    c. Operations phase
3. Revision of the existing policy implementation: Is the existing policy implementation in the context of administrative procedures adequate? In case the implementation is not adequate it needs to be adjusted.
4. Regular policy checking: Do changes in the environment, e.g. the legal grounds, require a revision of the policy? In case the policy was modified, the implantation needs to be revised and possibly adapted as well (see also previous task).

Concerning the second task in the process framework (policy implementation) a number of requirements need to be fulfilled. In this context the planning phase is very important, as mistakes in the design of a administrative procedure and the relating ICT infrastructure cannot easily be mitigated in the operations phase. Relevant tasks in the planning phase are:

- Development of a concept (also called data protection concept) how to implement the data protection requirements from the privacy policy concretely in the administrative procedure. This includes:
    - Development of a concept of roles and rights (probably transferred to a formalised access control policy) including avoidance of role conflicts and the use of a “second set of eyes” where needed
    - Development of a concept for audit logging (which data is going to be stored how long, which reports are going to be generated how and who gets access to audit data?)
    - Development of processes and technical support (e.g. automated reports or access to log data) to handle data subjects rights (such as information, correction of personal data and erasure) during operations of administrative procedures

In the implementation phase the following tasks are relevant:

- Testing of the fulfilment of data protection requirements (taken from the data protection concept), documentation of the results of the testing
- Formal releasing of the positively tested procedure (including technical solutions)

[114] In practice also more fine grained models for the description of the life cycles are used. For example in cases where procedures are ceased, an additional phase is needed.

During operations the following tasks are relevant:

- Management of incidents during operations from a data protection point of view
- Regular audits to check whether the policy is still implemented and procedure to handle deviations
- Handling of data subjects requests based on their rights, e.g. containing information about personal data processed, data correction and erasure

As many of these tasks and naturally the following processes are similar to tasks carried out in other domains such as the IT service management or information security management, process integration from the point of view of efficiency and effectiveness makes sense (see also FIDIS D14.8).

6.3 Description of existing technologies to enhance privacy in eGovernment

Technical enforcement of data protection and privacy in government closely work together with organisational structures. Most important instruments in this context are (a) defined responsibilities of governmental organisational units for administrative procedures and related personal data (mainly defined by law), (b) software applications supporting these procedures, e.g. by workflow support, (c) role-based access control in the context of these applications and (d) logging and audits (see Fidis D16.1).

Administrative procedures need to be documented together with the legal grounds for processing personal data. This can be done using data protection policies (also called privacy policies) or registers of procedures by the responsible governmental offices. For data protection policies the Article 29 Data Protection Working Party has made an important proposal in the Working Paper 100, suggesting a three-layer model for privacy policies.[111] So far a few examples are known only where privacy policies are carried out in machine readable, standardised formats so that techniques for automated policy handling can be used (see chapter 6.4).

Relating to IMS the implementation of policies is an important task. This relates to general privacy policies (c.f. chapter 6.4) as well as to access control policies. The implementation of access control policies in governmental IMS typically is implemented via roles in IMS or mutually agreed access control policies in federated environments. In the context of federated IMS federation frameworks such as Higgins[115], the Identity Meta-Framework from Microsoft[116], Liberty Alliance[117] and access control policy mark-up languages such as the eXtensible Access Control Markup Language (XACML) developed by OASIS[118] may be used to support access control. However, in the context of governmental IMS this approach seems to be used rarely today, as the country reports in Fidis D16.1 indicate. Hierarchical structures of IMS, e.g. via root CAs, CAs and sub-CAs, and trust relationships in management domains of IMS seem to be used more commonly. Examples for the latter approaches of

[115] See http://www.eclipse.org/higgins/

[116] See http://www.identityblog.com/stories/2005/07/05/IdentityMetasystem.htm

[117] See http://www.projectliberty.org/

[118] See http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml

establishing federated IMS are elaborated on a European level in the GUIDE[119] and the STORK eID[120] project. On a national level e.g. in Germany the establishment of a hierarchical PKI for the public administration in the federal states also is an example.

Role based access control is used either in centralised IMS or in an access control module directly in the applications. Depending on the complexity of administrative procedures and related processes, in some cases, very fine-grained access rights are used. However, the effectiveness of these technical measures depends greatly on the internal organisational structure (e.g. hierarchy) of governmental offices. Different roles carried out by the same person can conflict or lead to undesired linking of personal data.

In addition to these presented instruments privacy enhancing mechanisms and in some cases Privacy Enhancing Technologies (PETs) are used. Examples are:

- Anonymisation of personal data, e.g. in the context of the compilation of statistical data, e-petitions and e-voting [121]
- Pseudonymisation of personal data, e.g. in the health sector (for example by the use of cryptographic techniques and proper key management in cancer registers in Switzerland[122] and Germany[123] and databases of physiological samples,[124] where pseudonymous identifiers are used for statistical and research purposes).
- Definition and technical enforcement of governmental sectors by authenticating citizens based on sectoral, pseudonymous identifiers. Such a concept was implemented within government via the Austrian citizen's card (Bürgerkarte, see e.g. Meints, Hansen 2006). For the use of the national eIDs in the private sector such a concept was implemented in Austria; for Germany it is planned in the context of the national eID, the so-called ePersonalausweis (ePA).[125]
- The application of procedure-specific PETs relating to access handling and access control. In the health sector examples are known where Privacy Preserving Data Mining (PPDM) techniques are used as PETs.[126] Other examples are the use of techniques to avoid the collection of (otherwise identifying) personal data not needed in the context of video surveillance in the public space.[127]

In addition, internal data protection assuring measures are technically supported. Most important in this context are audit logs together with automated reporting. This typically

[119] See http://istrg.som.surrey.ac.uk/projects/guide/overview_index.html

[120] http://www.eid-stork.eu/

[121] See Alkassar, A. and Hansen, M. (eds.), *D3.8: Study on protocols with respect to identity and identification – an insight on network protocols and privacy-aware communication, FIDIS Deliverable,* 2008, available at <www.fidis.net>, last consulted 15 February 2009, pp. 69-89. and Alkassar, A. and Volkamer, M. (eds.), *E-Voting and Identity*, Springer-, Heidelberg 2007.

[122] See e.g. http://www.lustat.ch/ms_Datenschutzkonzept_2001.pdf

[123] See e.g. http://www.krebsregister.nrw.de/index.php?id=8

[124] See for e.g. Meints, M. and Hansen, M. (eds.), *D3.6: Study on ID Documents, FIDIS Deliverable,* 2006, available at <www.fidis.net>, last consulted 15 February 2009.

[125] See e.g. the unofficial version of the draft 2.0 for the concept of the German national eID (ePA) at http://netzpolitik.org/wp-upload/bmi_epa-grobkonzept-2-0_2008-07-02.pdf

[126] See for example http://www.lustat.ch/ms_Datenschutzkonzept_2001.pdf and http://e-hrc.net/media/ExtHealthNetworksMuscle02Feb2005.htm

[127] E.g. the privacy audit for the CCTV system of the Parliament of the Federal State of Schleswig-Holstein in Germany, see https://www.datenschutzzentrum.de/audit/kurzgutachten/a0613/.

covers management and execution of roles and access rights to personal data. Audit logs and reporting also may cover, in addition to security targets (such as confidentiality, integrity and availability), specific data protection aspects, e.g. minimisation of personal data by secure deletion. Logs and reports are important for internal (carried out by the internal Data Protection Officer or Advisor) and external data protection audits (carried out by the responsible Data Protection Commission). For the Belgian citizen card a concept was implemented allowing an audit of citizen's card related access of government officials by the citizen themselves. Citizens are able to analyse the corresponding log files via a web portal.[128]

In this context, also well established security measures need to be mentioned, as they also support the confidentiality and integrity of personal data in eGovernment. These measures are typically applied according to state-of-the-art standards in information security and frequently refer directly or indirectly to international standards, such as the ISO/IEC 27000 series, Common Criteria (ISO/IEC 15408)[129] or CobiT[130]. In the context of this deliverable, the application of cryptographic techniques was already introduced and discussed.

6.4 Semi-automated support for privacy policy handling

Privacy and trust policy negotiation can be defined as a set of messages exchanged between a user and a service provider in which both parties agree on data to be released by the user in exchange for a service (e.g. providing proof of age with a credential for accessing a service) and data to be released by the service provider to give the user assurance (evidence that it is sufficiently "trustworthy" as required by the user's preferences). During policy negotiation also, the data handling policy for the user-released data, typically including obligations to be enforced by the services side, is agreed on.

In practice, users usually only have the choice to accept, in exchange for a service, to release data for the services side's legal terms and conditions. If they do not accept, the service is usually not provided[131].

The Platform for Privacy Preferences (P3P) protocol can be seen as a simple policy negotiation protocol which follows this "take-it-or-leave it" approach. It allows web sites to express their privacy policy (data handling policy) in machine-readable (XML-based) format, which indicates which data are to be collected, for what purposes, for how long they will be retained and with whom they will be shared. Users in turn can define their privacy preferences (data release policy) in machine-readable format. P3P user agents are fetching a services site's privacy policy, informing the users about the site's privacy practices, matching the privacy policy against the user's preferences and taking appropriate actions, such as alerting the user in case of a mismatch. Early drafts of P3P included a protocol for multi-round negotiation, which was, however, dropped from the specification, as it was believed that it made P3P too

[128] The access to this service is available via https://www.mijndossier.rrn.fgov.be/, but requires a client certificate which is provided from the Belgian citizen card.

[129] See http://www.commoncriteriaportal.org/

[130] Available free of costs via http://www.isaca.org/

[131] Even though at least according to Paragraph 3 (3) of the German Teleservices Data Protection Act of the German Federal Information and Communication Services Act (Informations- und Kommunikationsdienste-Gesetz) and Paragraph 12 (4) of the German Interstate Agreement on Media Services, a service provider shall not make the use of tele- and multimedia services conditional upon the consent of the user to the effect that his/her personal data may be used for other (secondary) purposes. According to the Teleservices Data Protection Act, this obligation is only applicable if other accesses to these services are not or not reasonably provided to the user.

complicated[132]. While P3P can help to standardise privacy notices and to advance transparency, P3P alone does not provide mechanism for enforcing P3P policies (ensuring that companies follow privacy policies); it does not ensure compliance of privacy policies with privacy laws and does not guarantee a minimum, non-negotiable level of privacy protection for individuals. However, in the context of eGovernment it is important to mention that many US government web sites have posted P3P policies to comply with the privacy requirements of section 208 (c) of the E-Government Act of 2002, which requires government agencies to post privacy notices on all federal government websites in a standardized machine-readable format. As P3P is the leading standard for privacy policies in machine-readable formats, it has been used by many government agencies for implementing P3P policies on their web sites, including the Federal Trade Commission, Department of Commerce and US Postal Service.

The PRIME Architecture which has been developed within the EU FP6 project PRIME[133] on Privacy and Identity Management for Europe takes a more advanced approach than P3P does to negotiation as illustrated in figure 1 and described in figure 7:

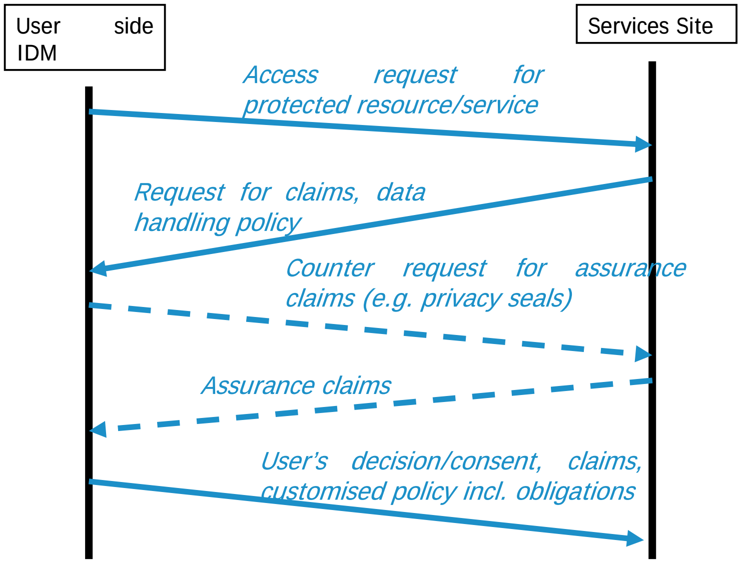
User side IDM
Services Site
Access request for protected resource/service
Request for claims, data handling policy
Counter request for assurance claims (e.g. privacy seals)
Assurance claims
User's decision/consent, claims, customised policy incl. obligations

Figure 7: Data and Policy Exchange in PRIME (the dashed line stands for optional message flows)

The negotiation process starts when the user-side identity management (IDM) application requests access to a resource or service protected by a service's access control system. The server in this case returns a *request for claims* and a corresponding *data handling policy* in a formalised, machine-readable form.

A *claim* is a statement made by an entity about another entity or set of entities. A claim can be endorsed by a third party, which certifies the claim in an integrity-protected manner. An example of a claim is "The requester is of age greater than 18 years, claimed by the requester, endorsed by an EU-member-state-issued passport"). A claim request (or: request for claims)

[132] Preibusch, Sören, "Privacy Negotiation with P3P", W3C Privacy Workshop 17.10.2006, http://www.w3.org/2006/07/privacy-ws/papers/24-preibusch-negotiation-p3p/.

[133] https://www.prime-project.eu/

is issued in order to obtain claims that satisfy the access control policy for a requested resource.

The data handling policy specifies how the recipient of claims should handle the data once released by the data subject. It consists of two parts. The first part is related to access control and is enforced by the access control system of the recipient; the second part is privacy obligations, which are conditions to be met by the service provider after the data has been released (e.g. conditions to notify users or delete data under certain circumstances), and is enforced by the life-cycle data management system of the recipient. The enforcement of privacy obligations can be related to access control decisions or be completely orthogonal to access control, as is the case with time-based data management obligations.

The data request (in form of the claim request) is analysed by the user's IDM application, which makes a recommendation on what claims to make under which policy on the basis of the user's privacy preferences, the history metadata, and the user's real-time input. Note that the claim request is already crafted by the server to satisfy the server's authorization policy and data handling policy requirements.

Optional step: The user's analysis can yield a *counter-request* for assurance claims (e.g. privacy seals issued to the server by parties trusted by the user, such as consumer protection agencies). Assurance claims can help the service provider to "prove" to the user that they are going to handle user data in the agreed way or that they are using appropriate data processing machinery to manage the user data. The counter-request for claims is sent to the service provider and the service provider returns appropriate claims (such as a privacy seal). When the user receives the appropriate claims, they are analysed by her IdM system. If the analysis has an appropriate outcome, the user will release the claims initially requested by the service provider. Note that the user-side analysis is governed to a large extent by the user's *release policy*, a policy that dictates under which conditions data may be released to other parties.

An easy-to-understand representation of the data handling policy, the claims to release, the overall transaction context, the result of the assurance check, and compliance of the server's data handling with the user's own preferences is displayed to the user. A customisation of the proposed data handling policy through the user may be possible if options have been proposed by the service provider within the data handling policy. The customisation allows the user to bring their own data handling requirements into the data handling policy and can involve things like defining notification obligations in case data are transferred to third parties by the service provider, or notifications on policy enforcement, or the opt in to direct marketing. Finally, user consent is solicited for the overall transaction to be carried out. The user's consent, claims and the customised data handling policy are sent back to the services site, which will associate the personal data provided by the user (by the use of claims) with the negotiated (customised) policy. The negotiated data handling policy will be enforced through the services site's authorization engine and the life-cycle data management component.

The agreed data handling policy applies transitively for the user's data, that is, it applies whenever the user's data held by a service provider are disclosed to a third party or from the third party to another party, and so on. In this case, the entity that wants to disclose more data, regards the data handling policy agreed with the user as preferences to be enforced in the agreement on a data handling policy with the receiving entity. Such transitive disclosures can be performed to further entities with the effect that the newly agreed data handling policy is at least as restrictive as the user's originally customized data handling policy, also reflecting the user's privacy preferences.

6.5 Summary

The processing of personal data by governmental offices needs to be transparent. In eGovernment privacy policies are one important instrument to state relevant information about administrative procedures and the processing of personal data in this context to the citizens. Within governmental offices IMS are an important technical infrastructure to enforce the implementation of stated policies. In this chapter the following aspects were investigated:

- technical and organisational measures needed to run a Data Protection Management System (DPMS) that ensures the enforcement of data protection policies
- the role of IMS in DPMS and in the enforcement of data protection policies and
- technical platforms for the support of semi-automated handling of privacy policies, especially P3P and the data and policy exchange architecture, developed in the PRIME project.

The presented approaches are not commonly used by governmental administrations today and need to be classified as best practice. While organisational structures are needed to ensure the enforcement of legal requirements concerning data protection, good practice models for integrated DPMS were presented only recently. While P3P is implemented broadly by U.S. government offices on public websites, technical platforms for the support of semi-automated handling of privacy policies do not seem to play an important role for websites of European administrations yet. The data and policy exchange architecture developed in the PRIME project still is subject to further research in the EC funded PrimeLife project[134].

[134] See http://www.primelife.eu/

7 Summary and conclusions

After the FIDIS deliverable D16.1 provided a broad description of the concepts that together provide the building blocks for a framework for privacy-friendly Identity management in eGovernment, this deliverable set out to describe requirements for such an identity management system by building further on the results of deliverable 16.1. Three separate but intertwining approaches have been taken to do this, namely: a technical approach, an organisational approach and a legal approach. The number of requirements offered here have been relatively limited, seeing as we chose to focus on those that were particularly relevant to the eGovernment context. In the following paragraphs we shall provide a brief recapitulation of the most important findings contained in this deliverable.

Circles of Trust

In the eGovernment setting a large number of disparate entities cooperate. For such collaboration to be successful, agreements need to be made regarding the exchange of identity information among the communicating entities. Such a form of co-operation resembles to a large extent what has been described in Identity Management literature as a “Circle of Trust” (CoT), whereby a group of service providers and identity providers share linked (partial) identities and have pertinent business agreements in place regarding how to do business and interact with identities.

The foundation of a CoT is the reaching of an agreement on how identification and authentication will be organized. Most eGovernment identity management systems have put mechanisms in place for identifying and authenticating their users, most notably by the provisioning of identity documents. However, there are many additional issues concerning information use and governance which need to be addressed in order to create both compliant and successful applications. The European Commission also considers the main purpose of electronic identification for public services as the easing of access and offering personalised and smarter services. Since such electronic identification management is one of the key enablers of eGovernment, approaches for the requirements for successful application of such identity management are described in this deliverable.

Technical approaches

In the technical sections of this deliverable, on the one hand requirements all IMS need to meet are described, and on the other hand advanced features are described that might be useful in certain cases but especially for eGovernment.

A description is given of the fundamentals of symmetric encryption as one of the foundations of encryption and thus one of the mechanisms for providing confidentiality. These mechanisms provide secrecy of data and of digital signatures which are used to ensure authenticity. Then, asymmetric encryption is described and the differences discussed with respect to symmetric schemes. After that, the way in which the most practical encryption scheme works, the hybrid encryption, is described. This combines the advantages of both, symmetric and asymmetric encryption. Authentication techniques tackle, in contrast to encryption schemes, the integrity and accountability of a user of a system or a message. They can be combined with encryption in order to achieve both integrity and secrecy.

Even though cryptography shows many prospects, the burden of key distribution is not fully resolved. If one wants to use symmetric algorithms, this is more obvious, as a trustworthy (i. e. secrecy-protecting) channel is needed for the transportation of the secret keys. But even in the case of asymmetric cryptography, where public keys are used and therefore no secrecy-protecting channel is necessary, one still faces the problem of integrity and accountability when distributing keys.

Public Key infrastructures (PKI's) are a possible approach to solve these problems. Using PKI's public keys are reliably assigned to persons by means of digital signatures and a certification authority (CA). The deliverable shows that the large tree of implicit trust which is an essential component of PKI's, and also taking into account different policies CA may use, does not really ensure that PKI's are an adequate solution especially in eGovernment. Not even the common X.509 standard appears to be up to the task because of its inherent flexibility and extensibility.

In the advanced technical section, several aspects of advanced techniques, that might be succesfully applied in eGovernment, are considered. From the point of view, that the citizen in his use of public services wants to achieve maximum protection of his personal data, much effort has been put into developing so-called identity obfuscation techniques. It might even be said that the citizen wishes to achieve anonymity in communication networks. Perfect anonymity cannot be achieved, which is why obfuscation techniques have been developed to meet the possible requirements for the states in between anonymity and identification. The deliverable shows, that there exist a lot of possibilities and techniques for identity obfuscation. All of them have their specific advantages and disadvantages. Broadcast, Private Information Retrieval and DC-Networks offer a very high level of protection even against powerful adversaries. However, they require much resources in terms of computational power and bandwidth. They also do not scale very well. Simple proxies on the other hand require only little effort, but cannot provide much protection, especially because one has to trust entirely the operator of a given proxy. Today Mix networks seem to be the only approach, that offers reasonable protection at reasonable costs. Nevertheless, Mix Networks are still under development and have not achieved a quality of service level appropriate for large numbers of users like users of public services.

Another advanced technique, described in the deliverable, is the Bbox, a component for secure logging. The BBox is a component of an audit architecture that acts as a black box of the system. It provides a secure logging mechanism that ensures that the activity of the system, represented in terms of log data, is recorded in a tamper-evident, confidential manner. It comes especially into play when enforcement, i.e. the prevention of "bad events", cannot be fully ensured. In this case, mechanisms are sought to reliably identify violations. This applies especially to settings where users surrender personal data to communication partners, as can be said to be the case in typical eGovernment applications. Audits, i.e. the analysis of logs, can identify such violations. However, an audit has no validity, if the log it is based on could have been changed or otherwise corrupted. In other words: if the log might be "incorrect", the audit result has no significance. While secure logging (or revision and tamper-proof logging) is a basic and essential service that could be implemented using various technical solutions, BBox with its black box approach, in fact, is an advanced implementation architecture and prototype. The BBox extends the existing protocols by providing in addition security during the transmission phase from the collector to the logging entity. The architecture and the components of the BBox and the two cryptographic techniques the BBox is built on are described.

Organisational approaches

Trust between the collaborative entities in eGovernment can also be achieved by defining organisational and legal requirements in addition to the technical requirements. Every participant must adhere to certain procedures and policies. In the privacy friendly environment which is sought after for eGovernment, privacy policy handling turns out to be valuable. It may help to significantly limit the operational risk of each participant when it seeks to initiate its own application or data exchange. Finally, it allows privacy considerations to be taken into account during the design of the system.

A much applied organisational measure in the quest for privacy friendly IdM in eGovernment is the adoption of a data protection policy. This can be understood as the declaration of which organisation processes which personal data of its citizens, how, in the context of which government or business process, in accordance with the corresponding (national) data protection act. These policies need to be publicly accessible documents. From the compliance with data protection legislation a number of tasks in the context of the enforcement of the privacy policy, emerge. They include the implementation of data protection principles in all phases of the life cycle of administrative procedures, namely the planning phase, implementation or building phase and operations phase. In addition, external influencing factors need to be taken into consideration, e.g. changes in technical infrastructure, changes in operative requirements in the context of administrative procedures and changes in legislation. A catalogue of (technically oriented) measures embedded in an organisational framework seems to be a good solution.

Technical enforcement of data protection and privacy in government closely work together with organisational structures. Essential instruments in this context are (a) defined responsibilities of governmental organisational units for administrative procedures and related personal data (mainly defined by law), (b) software applications supporting these procedures, e.g. by workflow support, (c) role- and attribute- based access control in the context of these applications and (d) logging and audits. Role based access control is carried out either in centralised IMS or in an access control module directly in the applications. The effectiveness of these technical measures depends to a significant extent on the internal organisational structure (e.g. hierarchy) of governmental offices. Different roles carried out by the same person can be conflicting or lead to undesired linking of personal data. Privacy enhancing mechanisms and in some cases Privacy Enhancing Technologies (PETs) can be used. Examples of these are anonymisation of personal data, pseudonymisation of personal data for example by the use of cryptographic techniques as well as authenticating citizens based on sectoral, pseudonymous identifiers. Finally procedure-specific PETs can be applied relating to access handling and access control. Internal data protection assuring measures are technically supported. Most important in this context are audit logs together with automated reporting. State-of-the-art standards in information security are also a natural part of the privacy policies. They support the confidentiality and integrity of personal data in eGovernment.

Legal approaches

Finally the organisational aspects of the requirements for privacy-friendly IdM are tied up with the legal ones. This becomes clear by the application of the basic principle of data protection, that data access should be restricted to authorized entities and at the same time be limited to the data needed, so that an authorized entity can execute its task adequately. Every processing operation in eGovernment requires as a rule a legal basis to legitimize the processing. First off, this requires allocating responsibilities to acting entities. Most importantly the role of controller, processor and third party must be assigned. This should be specified clearly in legislation, and (at the very least) in a written agreement.

We have developed a case for the role of authoritative sources. These can be useful instruments in assuring the accuracy of the data. As such they can play an important role in the advancement of privacy principles such as data minimisation. In establishing the privacy policy consideration should also be given to the designation of authoritative sources. Furthermore, authoritative sources can also be instrumental in user- and access management. In light of the fact that the profiles of the attributes of the requesting entities may vary over time, if the privacy policy lays down the users' privileges, the authentic source permits efficient user management. Emphasis is laid on the possible role of TTP's who should be charged with maintaining and managing operating data registries and reference directories. The national DPA's should closely supervise these TTP's. Technical measures should be deployed to detect and prevent unauthorized manipulation. Mechanisms for these were described in the technical sections.

Authorization management is also a standard legal requirement. This provides an overview of valid recipients for each object that qualifies as personal data, as well as a list of the actions they are allowed to perform upon these resources. Authorization policies should specify in which capacity a resource or service is accessible to users, as well as the situation (i.e. for what purpose) and the time-frame. Where intermediaries (e.g. mediators, service integrators) are used, these policies should in first instance be managed and enforced at that level. The technical authorization mechanism used must of course also allow for sufficient granularity as to the permissions of every possible requesting/ asserting entity. In order to mitigate one major flaw of this approach, viz the lack of a mechanism that verifies whether the data is processed for a legitimate purpose, as long as the reqested data falls within the scope of the authorization profile, the deliverable suggests considering the use of technical 'purpose specification' mechanisms. These mechanisms may be useful from both for the enforcement of privacy policies in real time, as well as after the fact. Careful consideration must however be given to the question of which entities shall be trusted with registration of these purposes.

In order to summarise the main findings of this deliverable, in the following table the requirements the study has delineated, which are necessary to obtain privacy-friendly IdM in eGovernment, are summed up. In the first column the relevant data protection principle is listed. Secondly, the technical and organisational approaches that can be followed to achieve compliance with these requirements in the eGovernment setting, are given, with reference to the section of the report, where they are described. Finally, in the third column, best practices that were identified within the study, which enable similar compliance but with an additional value, are listed.

Legal requirement	***Technical and organisational approaches recommended to achieve compliance***	***Technical and organisational approaches: best practices***
Legitimacy of processing	*Organisational:* *- Legal basis (specific, clear and precise, foreseeability) and/or informed consent (3.3, FIDIS 16.1)* *Technical:* *- Registration of legal bases, prior authorizations and/or consent in repositories; adapting technical policy rules to include these elements as policy conditions (3.5), document and audit regularly* *- User Interface for obtaining informed consent*	*Organisational and technical:* *- Obtain consent in addition legal basis whenever possible*
Data minimization	*Organisational:* *- Privacy impact assessment (inter alia verification that only the information which is absolutely needed for a specific administrative procedure is disclose; explicating data life cycle, incl. intended storage duration for each data element) (FIDIS 16.1);* *- Use of Authoritative Sources (avoid unnecessary duplication) (3.4)* *- Document and audit all technical and organisational measures on a regular basis* *Technical:* *- Implement access and processing limitations supporting sufficient level of granularity (3.5, 4.1);* *- Mechanisms to respond to data requests with only that information that the requesting entity is authorized to receive (3.5, 4.1);* *- No propagation and/or verification of more attributes*	*Technical:* *- Purpose specification (3.5);* *- Annonymous communication (5.1)* *-Additional measures to avoid unnecessary linkability (pseudonym management, identity obfuscation techniques) (5.1, 6.3);* *- Authentication based on anonymous credentials (4.1, 5.1)* *- Additional measures to avoid unauthorized or unnecessary monitoring (5.1)*

Legal requirement	***Technical and organisational approaches recommended to achieve compliance***	***Technical and organisational approaches: best practices***
	than needed (3.6); *- Destruction or anonymization of personal data once the purpose for which it was collected / processed has been completed (taking into account need for accountability at later time);*	
Data accuracy	*Organisational:* *- Designation and use of Authoritative Sources (3.4)* *- Establish procedures for verification of each attribute with a level of assurance proportionate to the interests at stake;* *- Review and update procedures for personal data which is being kept for a long period of time;* *- Establish procedures on how to communicate and deal with suspected inaccuracies (3.4);* *- Indication of "level of confidence" by the data provider where appropriate;* *- In the event of indirect collection, verify data with data subject where possible prior to further processing* *- Document and audit all technical and organisational measures on a regular basis* *Technical:* *- Restrict modification rights to authorized entities (3.5, 4.1);* *- Implement appropriate security policies (e.g. use of cryptography to ensure authenticity and integrity) (3.7, 4.1, 4.2, 4.3)*	
Finality	*Organisational:* *-Privacy Impact Assessment (see FIDIS 16.1);*	*Organizational:* *-Prior authorization by national DPA (3.5)*

Legal requirement	***Technical and organisational approaches recommended to achieve compliance***	***Technical and organisational approaches: best practices***
	Technical: *- Registration of legal bases, prior authorizations and/or consent in repositories; adapting technical policy rules to include such bases and/or authorizations as policy conditions (3.5)* *- Document and audit all technical and organisational measures on a regular basis*	*Technical:* *-Purpose specification (3.5)* *-Additional measures to avoid unnecessary linkability (pseudonym management, identity obfusciation techniques) (5.1, 6.3)* *-Privacy-enhanced access control technically enforcing purpose binding based on policy languages (EPAL, XACML) (4.1).*
Security of processing	*- Establishing of an organisational framework for information security management e.g. an (Information Security Management System, ISMS, see FIDIS D16.1) and data protection (e.g. a Data Protection Management System, DPMS (6.1, 7.2)); this framework serves as an anchor for the technical and organisational measures listed below* *- Appropriate identification, authentication and authorisation of entities, which shall typically involve (both organisational and technical):* *- Establish Circles of Trust (defining roles and responsibilities, division of tasks wrt authoritative sources, verification of identity and other attributes, trusted third parties, contractual specification of liabilities, ...) (3.1, 6.2)* *- Manage identity life cycle in a way which provides an assurance level proportionate to the interests at stake (see FIDIS 16.1);* *- Establish procedures for verification of each attribute requesting/asserting entity with a*	*Technical:* *- Purpose specification (3.5)* *- Additional measures to avoid unnecessary linkability (pseudonym management, identity obfusciation techniques) (3.2, 5.1, 6.3)* *- Additional measures to avoid unauthorized or unnecessary monitoring (5.1)* *- Privacy-enhanced access control based on policy languages (EPAL, XACML) (4.1).* *- Certification of an ISMS (ISO/IEC 27001 certificate) or DPMS* *- Use of an IMS or components of an IMS that are successfully security certified (e.g. according to ITSec, ISO/IEC 15408 (Common Criteria), FIPS-140 etc.)*

Legal requirement	***Technical and organisational approaches recommended to achieve compliance***	***Technical and organisational approaches: best practices***
	level of assurance proportionate to the interests at stake (e.g. use of multi-factor authentication mechanisms) (3.6, 4.1, 4.2, 4.3, 3.2); *- Implement access and processing limitations supporting sufficient level of granularity (3.5, 4.1);* *- Implement appropriate security policies (e.g. regarding use of cryptography to ensure confidentiality, authenticity, integrity) (3.7, 4.1, 4.2, 4.3, 3.2, 4.4);* *- Use of Authoritative sources in user- and access management (3.4);* *- Restrict physical access to terminals which enable personal data processing where appropriate;* *- Associate restrictions and obligations with each data processing operation (3.8);* *- Adopt internal privacy policies (documenting all security measures, specifying inter alia persons responsible, what to do in the event of a breach, ...), provide education and awareness training for all persons who come in contact with personal data;* *- Confidentiality agreements;* *- Designate security officers;* *- Document and audit all technical and organisational measures on a regular basis*	
Accountability	*Organisational:* *- Establish Circles of Trust (defining roles and responsibilities, division of tasks wrt authoritative sources, verification of identity and other*	*Technical:* *- Purpose specification (3.5)* *- Secure logging and enhanced transparency mechanisms allowing direct data subject access to view the processing*

Legal requirement	***Technical and organisational approaches recommended to achieve compliance***	***Technical and organisational approaches: best practices***
	attributes, trusted third parties, contractual specification of liabilities, …) (3.1, 6.2) *- Adopt internal responsibility and accountability mechanisms (e.g. designating 'owners' for both equipment and processing operations involving personal data)* *Technical and organizational:* *- Authentication, authorisation and access control (4.1)* *- Use of non-repudiation mechanisms (4.1, 4.2, 4.3)* *- Log data processing operations (displaying which entity performed which action at which time, and in which context), agreed task division as to which entity will log which actions (3.9, 5.2);* *- Implement restrictions and obligations (e.g. notification services) (3.8)* *- Document and audit all technical and organisational measures on a regular basis*	*operations performed upon his personal data (5.2)*
Transparency and data subject rights (notification, access, rectification, object, deletion)	*Organisational:* *- Designation of controller in relevant legislation, widely communicating to whom and how data subject may direct requests regarding data subject rights, internal procedures to reply to these requests in a timely manner;* *- Indicate source of personal data and logic of processing when notifying data subject of decision based on such data;* *- Document and audit regularly* *- Posting of privacy policies on web sites*	*Organisational:* *- Provide notification to the data subject and/or to the public in the event of security breach* *Technical:* *- Enhanced transparency mechanisms allowing direct data subject access to view the processing operations performed upon his personal data (5.2) (accompanied by published procedure on how to report suspected inaccuracies and exercise right to object or demand deletion)* *- Standardised machine-*

Legal requirement	***Technical and organisational approaches recommended to achieve compliance***	***Technical and organisational approaches: best practices***
	- Post privacy notices on all federal government websites in a standardized machine-readable format (e.g., in P3P) (6.1).	*readable privacy policies (e.g., P3P) (6.1)(note however that in the USA, these are even required for compliance of E-Government web sites).* *- Policy negotiation (6.1)*

Further research

It can be concluded that the deliverable provides a three-pronged approach towards developing requirements for privacy friendly IdM in eGovernment. These three strands, the technical, the organisational and the legal, reinforce each other. They are best joined together in an integrated privacy policy. Solutions and recommendations have been provided where possible, conundra are shown where encountered. It has become clear that several questions are still open for research.

To summarise the most important questions, as they are indicated in the deliverable:

- Are PKIs an adequate solution especially in eGovernment in view of the unsolved question of the burden of key distribution?
- How can techniques for identity obfuscation (such as MIX Networks) be made sufficiently helpful and achieve a quality of service level appropriate for a large number of users? Is it possible to harmonise the European legal situation to make large scale implementation of these techniques possible?
- How can an audit achieve effectiveness if the log it is based on could have been changed or otherwise corrupted? Can the validity of audits be increased by using new logging services such as the BBox prototype?
- How can allocation of responsibilities be guaranteed in case the legislative measures warranting the processing do not provide sufficient clarity? The mere obligation of art 8 ECHR to provide a legal basis which is sufficiently clear and precise in authorising the processing, is – without effective enforcement and follow-up - no more than that.
- How can it be arranged that the national DPA's closely supervise intermediaries, who are considered to be acting as Trusted Parties or Trusted Third Parties, as the case may be?
- Which are appropriate technical solutions that can realistically be implemented to ensure that only the data needed for the processing are disclosed, even if the requesting entity's authorization profile would in principle otherwise permit access to greater amounts of data?

8 Bibliography

Accorsi, A., *Automated Counterexample-Driven Audits of Authentic System Records*, PhD Thesis, 2008, available at <http://www.freidok.uni freiburg.de/volltexte/6048/pdf/Diss_RafaelAccorsi.pdf>, last consulted 10 March 2009.

Accorsi, A., *Digital Evidences based on Log Data: What Secure Logging Protocols Have to Offer?*, submitted to COMPSAC 2009.

Alkassar, A. and Hansen, M. (eds.), *D3.8: Study on protocols with respect to identity and identification – an insight on network protocols and privacy-aware communication, FIDIS Deliverable,* 2008, available at <www.fidis.net>, last consulted 15 February 2009.

Alkassar, A. and Husseiki, R., *D3.9: Study on the Impact of Trusted Computing on Identity and Identity Management*, FIDIS Deliverable, 2008, available at <www.fidis.net>, last consulted 15 February 2009.

Alkassar, A. and Volkamer, M. (eds.), *E-Voting and Identity*, Springer-, Heidelberg 2007.

American Library Association (ALA), *E-Government Act of 2002- Details and Background*, available at <http://lita.org/ala/washoff/woissues/governmentinfo/egovernment/backgroundab/background.cfm>, last consulted 15 February 2009.

Anderson, A., A Comparison of Two Privacy Policy Languages: EPAL and XACML, Proceedings of the 3rd ACM workshop on Secure web services, ACM New York, 2006.

Ardagna, C. A. et al., Exploiting cryptography for privacy-enhanced access control. To appear in Journal of Computer Security, 2009.

Article 29 Data Protection Working Party, Working Document on E-Government, *WP* 73, 8 May 2003.

Bauer, M., Meints, M., and Hansen, M. (eds.), *D3.1 Structured Overview on Prototypes and Concepts of Identity Management Systems, FIDIS Deliverable,* 2005, available at <www.fidis.net>, last consulted 15 February 2009.

Belgian Privacy Commission, Recommendation nr. 01/2008 of 24 September 2008 concerning user- and acces management in the governmental sector, 24 September 2008, 3, available at <http://www.privacycommission.be/nl/docs/Commission/2008/aanbeveling_01_2008.pdf>, last consulted 19 December 2008.

Bishop, M., *Introduction to Computer Security*, Addison-Wesley, Boston, 2005.

Böhme, R. and Pfitzmann, A., 'Digital Rights Management zum Schutz personenbezogener Daten?'[Digital Rights Management for the protection of personal data?], Datenschutz und Datensicherheit, vol. 32, issue 5, 2008, pp. 342-347.

Bot, D. de, *Privacybescherming bij e-government in België. Een kritische analyse van het Rijksregister, de Kruispuntbank van Ondernemingen en de elektronische identiteitskaart* [Protection of privacy in the e-government of Belgium. An critical analysis of the National Register, Crossroadsbank for Enterprises and the electronic ID card], Vandenbroele, Brugge, 2005.

Buitelaar, H., Meints, M. and Alsenoy, B. van (eds.), *D16.1: Conceptual Framework for Identity Management in eGoverment, FIDIS deliverable*, 2008, available at <www.fidis.net>, last consulted 15 February 2009.

Camp, L. van, *'Designing for Trust'*, available at: http://www.ljean.com/files/whatIsTrust.pdf, November 2004, last visited: 10 April 2009

Council of Europe and Commissioner for Human Rights, *Protecting the right to privacy in the fight against terrorism,* December 2008, CommDH/IssuePaper(2008)3, 6.

Cvrcek, D. and Matyas, V. (eds.), *D13.1: Identity and impact of privacy enhancing technologies*, FIDIS Deliverable, 2007, available at <www.fidis.net>, last consulted 15 February 2009.

Deadman, S. (ed.), *Circles of Trust: The Implications of EU Data Protection and Privacy Law for Establishing a Legal Framework for Identity Federation*, 2005, available at <www.projectliberty.org>, last consulted 15 February 2009.

Deprest, J. and Robben, F., *eGovernment: the approach of the Belgian federal administration*, available at <http://www.ksz.fgov.be/En/Como/2003%20%20EGovernment%20paper%20v%201.0.pdf>, last consulted 15 February 2009.

eGovernment Unit of New Zeeland, *Authentication for e-government. Best Practice Framework for Authentication*, available at <http://www.e.govt.nz/services/authentication/authentication-bpf/bpf.pdf>, last consuled 15 February 2009.

European Commission Communication of 25 April 2006, i2010 eGovernment Action Plan - Accelerating eGovernment in Europe for the Benefit of All [COM(2006) 173 final, (i2010 eGovernment Action Plan), of which a summary is available at <http://europa.eu/scadplus/leg/en/lvb/l24226j.htm>, and full text (12 p.) at <http://ec.europa.eu/information_society/activities/egovernment/docs/highlights/comm_pdf_com_2006_0173_f_en_acte.pdf>, last consulted 15 February 2009.

Franklin, M., 'A survey of key evolving cryptosystems', *International Journal of Security and Networks*, vol.1, issue 1-2, 2006, pp.46-53.

Gallegos, F. et al., *Information Technology Control and Audit*, Boca Raton: Auerbach, Florida, 2004.

Glynos, D., Kotzanikolaou, P., and Douligeris, C.,'Preventing impersonation attacks in MANET with multi-factor authentication', *Proceedings of the 3rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks,* IEEE Computer Society Press, 2005.

Halmen, K. W., Morrisett, G. and Schneider, F. B., *Computability Classes for Enforcement Mechanisms*, 2006, available at <http://www.cs.cornell.edu/fbs/publications/EnfClassesTR2003-1908.pdf>, last consulted 10 March 2009.

Howard, M., and Leblanc, D., *Writing Secure Code*, Microsoft Press, Redmond, 2001.

Huysmans, X., and Alsenoy, B. van, '*Conceptual Framework for Identity Management in eGovernment and Requirements Study*', Deliverables 1.1 and 1.3 of the IBBT project 'IDEM' (Identity Management for eGovernment), 2007. (not made public)

IBM, Enterprise Privacy Authorization Language (EPAL), W3C Submission Request, Nov. 2003, available at <http://www.w3.org/Submission/2003/07/>, last consulted 3 February 2008.

IDA Authentication Policy. *Basic policy for establsihing the appropriate authentication mechanisms in sectoral networks and projects*, available at <http://ec.europa.eu/idabc/servlets/Doc?id=19281>, last consulted 15 February 2009.

IDEM glossary, available at <https://projects.ibbt.be/idem/uploads/media/2007-12-27.idem.glossary.v1.07.pdf>, last consulted 15 February 2009.

ITU-T SG17 Focus Group for Identity Management, *"Report on Requirements for Global Interoperable Identity Management"*, September 2007, www.itu.int/ITU-T/studygroups/com17/fgidm, accessed 4 December 2007.

Kahn Consulting, *Computer security log files as evidence*, available at <http://www.kahnconsultinginc.com/images/pdfs/KCI_ArcSight_ESM_Evaluation.pdf>, last consulted 10 March 2009.

Kenneally, E., 'Digital logs – Proof matters', *Digital Investigation*, vol.1, issue 2, 2004, pp. 94–101.

Kent, K., and Souppaya, M., *Guide to Computer Security Log Management*, 2006, available at <http://csrc.nist.gov/publications/nistpubs/800-92/SP800-92.pdf>, last consulted 10 March 2009.

Kerckhoffs, A., 'La cryptographie militaire*'* [The military cryptography], *Journal des sciences militaires*. vol. 9, 1883, pp. 5–38 and pp. 161–191.

Korba, L. and Kenny, S., *Towards meeting the privacy challenge: adapting DRM*, available at < http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.8617>, last consulted 10 March 2009.

Koorn, R.(ed.), *Privacy Enhancing Technologies – White Paper for Decision-Makers*, written for the Dutch Ministry of Interior and Kingdom relations, available at <http://www.dutchdpa.nl>, last consulted 22 May 2007.

Lamport, L.,'Password authentication with insecure communication', *Communications of the ACM* , vol. 24, issue 11, 1981, pp.770-772.

Langheinrich, M. and Roussopoulos, M. (eds.), , *Technology-Induced challenges in Privacy & Data Protection in Europe*, ,available at <http://www.enisa.europa.eu/doc/pdf/deliverables/enisa_privacy_wg_report.pdf>, last consulted 15 February 2009.

Leenes, R. (ed.), *D5.2b: ID-related Crime: Towards a Common Ground for Interdisciplinary Research, FIDIS Deliverable, 2006,* available at <www.fidis.net>, last consulted 15 February 2009.
Maier, P., *Audit and Trace Log Management*, Boca Raton: Auerbach, Florida, 2006.

Meints, M., 'Datenschutz durch Prozesse' [Data Protection through processes], *Datenschutz und Datensicherheit,* vol. 31, issue 2, 2007, pp. 91-95.

Meints, M. and Hansen, M. (eds.), *D3.6: Study on ID Documents, FIDIS Deliverable,* 2006, available at <www.fidis.net>, last consulted 15 February 2009.

Menezes, A.J., Van Oorschot, P.C. and Vanstone, S.A., *Handbook of Applied Cryptography*, CRC Press, Boca Raton, 1997.

Mercuri, R., 'On auditing audit trails', *Communications of the ACM* , vol. 46, issue 1, 2003, pp.17-20.

Moses, T., eXtensible Access Control Markup Language (XACML) Version 2.0, OASIS Standard, Feb. 2005, available at <http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-core-spec-os.pdf>, last consulted 3 February 2008.

Moses, T., (ed.). Privacy policy profile of XACML v2.0, OASIS Standard, Feb. 2005, available at <http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-privacy_profile-spec-os.pdf>, last consulted 3 February 2008.

Müller, G. and Wohlgemuth, S. (eds.), *D14.6: From Regulating Access Control on Personal Data to Transparency by Secure Logging*, *FIDIS Deliverable*, to appear in 2009, available at <www.fidis.net>, last consulted 15 February 2009.
Müller, G. and Wohlgemuth, S. (eds.), *D14.2: Study on Privacy in Business Processes by Identity Management, FIDIS Deliverable,* 2007, available at <www.fidis.net>, last consulted 15 February 2009.

Müller, G. and Wohlgemuth, S. (eds.), *D14.3 Study on the Suitability of Trusted Computing to support Privacy in Business Processes*, *FIDIS Deliverable*, 2008, available at <www.fidis.net>, last consulted 15 February 2009.

Oppliger, R., and Ritz, R., 'Digital evidence: Dream and reality', *IEEE Security and Privacy* , vol. 1, issue 5, 2003, pp. 44-48.

Pfitzmann, A. and Hansen, M., *Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management - A Consolidated Proposal for Terminology, v0.31*, available at <http://dud.inf.tu-dresden.de/Anon_Terminology.shtml>, last consulted 15 February 2009.

Preibusch, Sören, "Privacy Negotiation with P3P", W3C Privacy Workshop 17.10.2006, http://www.w3.org/2006/07/privacy-ws/papers/24-preibusch-negotiation-p3p/.

Rannenberg, K., Pfitzmann, A., and Müller, G., 'Sicherheit, insbesondere mehrseitige Sicherheit', in Müller, G., and Pfitzmann, A. (eds.), *Mehrseitige Sicherheit in der Kommunikationstechnik*, pp. 21-30, Addison-Wesley, New York, 1997.

Robben, F., *Een voorstel van informatiebeveiligingsbeleid bij de uitbouw van E-government door de federale overheidsdiensten,* available at <http://www.law.kuleuven.ac.be/icri/frobben/publications/2005%20-%20Voorstel%20van%20informatieveiligheidsbeleid%20bij%20de%20uitbouw%20van%20E-government.pdf>, last consulted 10 March 2009.

Robben, F., *Gebruikers- en toegangsbeheer: beschikbare diensten*, available at <http://www.law.kuleuven.be/icri/frobben/presentations/20061108.ppt>, last consulted 15 June 2007.

Rössler, T., Identification and Authentication in Networks enabling Single Sign-On, available at <http://www.iaik.tugraz.ac.at/teaching/11_diplomarbeiten/archive/roessler.pdf>, last consulted 15 February 2009.

Sanett, S., and Park, E., 'Authenticity as a requirement of preserving digital data and records', *IASSIST Quarterly*, vol. 24, issue 1, 2000, pp.15-18.

Schneier, B. and Kelsey, J.,"Security audit logs to support computer forensics", *ACM TISSEC*, vol. 2, issue 2, 1999, pp. 159-176.

Schneier, B., *Applied Cryptography,* Addison-Wesley, New York, 1996.

Slone, S. (ed.), *Identity Management. A white paper*, 2004, available at < http://www.opengroup.org/onlinepubs/7699959899/toc.pdf>, last consulted 15 February 2009.

The Platform for Privacy Preferences 1.0 (P3P1.0) Specification, W3C Recommendation 16 April 2002, available at <http://www.w3.org/TR/P3P/>, last consulted 15 February 2009.

Wouters, K. et al., *Secure and Privacy-Friendly Logging for eGovernment services*, Ares 2008 - Proceedings of the Third International Conference on Availability, Security and Reliability, March 2008, IEEE Computer Society

ZUCKER, L.G. ‘Production of trust: Institutional sources of economic structure, 1840-1920’, In, B.M. STAW and L.L. CUMMINGS (ed.), ‘Research of organizational behavior’, JAI Press Inc., Londen, 1986, p. 53-111.

9 Annex 1: Glossary

9.1 Acronyms

AES	Advanced Encryption Standard
CA	Certification Authority
CRL	Certificate Revocation List
CSP	Credential Service Provider
DES	Data Encryption Standard
DPA	Data Protection Authority
DPMS	Data Protection Management System
EPAL	Enterprise Privacy Authorization Language
IETF	Internet Engineering Task Force
IMS	Identity Management System
ITU-T	International Telecommunication Union
LDAP	Lightweight Directory Access Protocol
MAC	Message Authentication Codes
MITM	Man In The Middle
OASIS	Organization for the Advancement of Structured Information Standards
OSCI Transport	Online Services Computer Interface Protocol
P3P	Platform for Privacy Preferences
PET	Privacy Enhancing Technology
PIN	Personal Identification Number
PKI	Public-Key Infrastructure
PoP	Proof of Possession
PRIME	Privacy and Identity Management for Europe (EU FP6 project)
RA	Registration Authority
RFC	Request for Comments
SA	Sectoral Application
SAML	Security Assertion Markup Language
SSL	Secure Sockets Layer
TTP	Trusted Third Party
X.509	ITU-T: The Directory: Public-Key and attribute certificate frameworks
XACML	Extensible Access Control Markup Language

9.2 Definitions

Access control

Access control determines who can view and/or access what resources, under which conditions, what they can do with it, and the type of device they can use it on.

Access control model

An access control model is a generic access control policy that is defined at a high level outside specific systems, then through translation, is applied on the appropriate level, as suitable to each individual system.

Access control policy

An access control policy is a set of rules to administer, manage and control the access to network resources.

Account

An account is a formal agreement between a principal and a service provider for regular dealings and services that defines user's or system's access to a resource or service. Each service may define a unique set of attributes to define an account.

Accounting

Accounting means monitoring the resource usage according to agreed criteria and processing the information into values that are suitable for use of a charging system.

Accuracy principle

The accuracy principle means that personal data should be accurate and, where necessary, kept up to date.

Anonymity

Anonymity refers to the quality or state of being not sufficiently identifiable to an attacker, within the set of all possible subjects that could cause an action or might be acted upon (the anonymity set).

Assertion

A statement from a verifier to a relying party that contains identity or attribute information about the claimant. Assertions support the claimnat's identity but are not bound to the token possessed by the claimant. A relying party trusts an assertion based on the source, the time of creation, and attributes associated with the claimant. Examples of assertions include SAML assertions and cookies.

Attribute

An attribute is a physical or abstract named property belonging to an entity.

An attribute typically has a value. An entity's identity is characterized through the values of its attributes.

Audit

A (security) audit is an independent review and examination of system records and operations in order to test for adequacy of system controls, ensure compliance with established policy and operational procedures, detect breaches in security, and to recommend any indicated changes in control, policy, and procedures.

Authentic data registry

An authentic data registry (or authentic registry) is a data registry about authentic data or an authentic copy of that data.

Authentic data repository

An authentic data repository is a data repository that contains either authentic data or an authentic copy of this data.

Authentic Source

An authentic source is a data repository of authentic data (i.e. a database where authentic information is maintained), maintained by one or more data managers.

Authentication

Authentication is the process of corroborating a claimed set of attributes or facts with a specified, or understood, level of confidence. Authentication serves to demonstrate the integrity and origin of what is being pretended.

Authorisation

Authorisation refers to (1) the permission of an authenticated entity to perform a defined action or to use a defined service/resource; (2) the process of determining, by evaluation of applicable permissions, whether an authenticated entity is allowed to have access to a particular resource.

Certificate

A certificate is an affidavit whereby an accredited certification body attests to the truth of certain stated facts. In identity management, the term is often used to refer to a public-key digital certificate. X.509 Digital Certificates are a prominent example thereof.

Certificate revocation list

A certificate revocation list (CRL) is a signed list of certificates that are not longer considered valid by the certificate issuer.

Certification authority

A Certification Authority or CA is an entity that certifies public keys. This means that it guarantees the relationship between the identified entity and the public verification key. This association is achieved in a digital certificate that binds the public key to a partial identity of an entity. It is a trusted party or trusted third party that accepts the responsibility of managing the certificate process by issuing, distributing and verifying certificates.

Circle of trust

A circle of trust is a group of service providers that trust each others identity management system and therefore share linked (partial) identities of identifiable entities and have pertinent business agreements in place regarding how to do business and interact with these identities.

Claim

A claim is a statement made by an entity (the claimant) about an entity (the claims object) to an entity or set of entities (the claims audience) or an undefined audience.

An entity may claim an attribute, authentication or authorization and will generally present a credential as proof, for the purpose of enabling authentication, to validate the user's identity, or to identify what the user is authorized to do.

Client

In computing, a client is a computer system or process that requests a service of another computer system or process.

In eGovernment, a client is typically a citizen, a company or a government entity that engages the professional advice or services of a provider.

Confidentiality

Confidentiality is keeping the information secret from all but those authorized to have access to it.

Consent

Consent is a free, specific, informed, unambiguous and sometimes even explicit or even written indication of the wishes of the data subject, by which he signifies his agreement to personal data relating to him being processed.

Context

A context is one setting of an entity's environment.

It can, for example, be a sphere of activity, a geographical region, a communicational setting, an application, a logical or a physical (security) domain. In identity management we typically refer to the meaning of the term 'context' as a (1) communicational setting or (2) a security domain.

Corroboration

Corroboration is the confirmation by provision of sufficient evidence and examination thereof that specified requirements have been fulfilled. This evidence typically but not necessarily takes the form of credentials.

Credential

A credential is a piece of information, mainly an attribute or a set of attributes attached to the entity that makes use of it, attesting to the integrity of certain stated facts. The production of adequate credentials need not involve disclosure of identity.

Cryptographic protocol

A cryptographic protocol is a distributed algorithm defined by a sequence of steps precisely specifying the actions required of two or more entities to achieve a specific security objective.

In identity management, the term "protocol" is often used to refer to a *cryptographic* protocol.

Data controller

The data controller is the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.

Data handling policy

A data handling policy is a policy that defines restrictions on secondary use of personal data.

Data processor

The data processor is a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the data controller, without coming under the direct authority of the data processor.

Data release policy

A data release policy is a policy that governs release of personal data (properties/credentials).

Data repository

A (data) repository is a central place in which an aggregation of data is kept and maintained in an organized way, usually in computer storage.

Data subject

A data subject is an identified or identifiable natural person to whom information relates.

Digital Identity

A digital identity is a partial identity in an electronic form.

Digital signature

A digital signature is data appended to, or a cryptographic transformation of, a data unit that allows a recipient of the data unit to prove the origin and integrity of the data unit and protects against manipulation (e.g. forgery by the recipient).

Domain

A domain is a set of entities, their information objects and a common policy.

Electronic signature

An electronic signature is data in electronic form which is attached to or logically associated to other electronic data and serves as a method of authentication. The term is a generic term which, by its use, incorporates almost all methods that achieve some level of data or entity authentication.

Encryption

Encryption is a means of transforming data from a readable form (known as plain text or clear text) to one that is unintelligible (referred to as cipher text).

Entity

An entity is an item of interest, inside or outside a system, such as an automated process, a subsystem, a device, a person or group of persons that incorporates a specific set of attributes.

Federated Identity

A federated identity is a credential of an entity that links an entity's partial identity from one context to a partial identity from another context.

Finality principle

Personal data must be collected for specified, explicit and legitimate purposes. The purpose of the processing should be defined the latest at the moment of the collection of the data.

Government entity

A government entity is any service, institute, natural person or legal person who fulfils tasks related to a public service, or tasks from public interest.

Government entities can act in different functions or capacities, such as a service provider, identity provider, data managers, data initiators etc.

Hash function

A hash function is an algorithm that computes a value based on a data object (such as a message or file, usually of variable-length and possibly very large). The data object is mapped to a smaller data object, called the 'hash result' or 'hash value'.

Identification

Identification is the process of using claimed or observed attributes of an entity to establish a partial identity of that entity.

Identifier

An identifier is an attribute or a set of attributes of an entity which uniquely identifies the entity within a certain context.

Identity

The identity of an entity is the dynamic collection of all of the entity's attributes. An entity has only one identity.

Identity management

Identity management (IDM) is the definition, designation and administration of identity attributes as well as the administration of the choice of the partial identity to be (re-) used in a specific context, to manage the access to and the usage of applications, services and resources.

It includes the management of identity attributes by their owners (user-side IDM) and/or by those parties with whom the owners interact (services-side IDM).

Identity Management System (IMS)

An identity management system is the organisational and technical infrastructure used for the definition, designation and administration of identity attributes.

Identity provider

An identity provider is a service provider that creates, maintains, and manages identity information for principals and provides entity authentication to other service providers, e.g. within a federation.

Integrator

An integrator is a mediator that integrates, orchestrates and/or aggregates services from different authoritative sources and delivers the result to the authorized requesting entity.

Integrity

Integrity is a quality that implies that the items of interest (facts, data, attributes etc.) have not been subject to manipulation by unauthorized entities. Establishing the integrity of a claim refers to the service that corroborates the integrity of the items of interest.

Interoperability

Interoperability is the communication, using standards, between several information technology systems held by various institutions or institutions.

Log

A log is a repository for records, which contain the information that is logged.

Logging

Logging is a process that records the linkage between an action and the identity of the entity or role that has invoked the action.

Man-in-the-Middle Attack (MITM)

A technique used by Internet hackers. It results in the hacker ‘positioning’ themselves between the user and the system they are transacting with. This allows them to monitor communications and obtain information transferred between the parties.

Mediator

A mediator is an entity that manages data traffic from and to authoritative sources. Mediator services include (1) routing, (2) transporting, (3) transforming or (4) granting access to the authentic data to authorized users. The latter implies prior authentication.

Non-repudiation

Non-repudiation refers to the concept of ensuring that a commitment or action cannot later be denied by one of the entities involved.

Oasis

OASIS was founded in 1993 under the name SGML Open, as a consortium of vendors and users devoted to developing guidelines for interoperability among products that support the Standard Generalized Markup Language (SGML). OASIS changed its name in 1998 to reflect an expanded scope of technical work, including the Extensible Markup Language (XML) and other related standards, in particular SAML. Refer to website for further information: http://www.oasis-open.org/home/index.ph.

Object

An object is a non-acting entity that contains or receives data or information, to which access is controlled.

Permission

Permission describes the privileges granted to an authenticated entity with respect to low-level operations that may be performed on some resource (e.g., read, write, delete, execute, create).

Personal Data

Personal data means any information relating to an identified or identifiable natural person (the data subject).

Personal Identification Number

A Personal Identification Number (acronym "PIN") is a factor used for entity authentication, whereby an entity enters a usually four digit number, which is only known to this entity.

Policy

A policy is one or more definite goals, courses or methods of action to govern present and future decisions. Policies are implemented or executed within a particular context (such as policies defined within a business unit). Common examples of these policies are security policies, access control policies, privacy policies, registration policies etc.

Policy negotiation

Policy negotiation is exposing the desired or appropriate part of a policy to another sector and/or context, which is necessary to support partial interconnection between different sectors and/or contexts.

Privacy

Privacy is the fundamental right of a natural person to respect for his private and family life, his home and his correspondence. Privacy is a *relative* fundamental human right. A limitation of the right to privacy is possible if (1) it is in accordance with the law, (2) necessary in a democratic society, (3) in a number of cases.

Privacy Policy

A privacy policy is a policy that contains statements on the basis of data protection regulation, inter alia on:

- what personal data is being collected,
- for what purpose the collected personal data is being used,
- how long the data is being retained,
- how a natural person can access and correct his/her own collected data,
- how the natural person can opt-out; and
- what security measures are being taken by the entities that process and/or control the data.

Processing Personal Data

Processing of personal data means any operation or set of operations which is performed upon personal data, whether or not by automatic means. Examples of personal data processing that are included in the Data Protection Directive are collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.

Profile

A profile of an entity or a group of entities is an organized set of attributes that characterizes and/or represents the specific properties of that entity or entities within a given context for a specific purpose.

Proportionality principle

According to the data protection and privacy rules, the proportionality principle has to be understood:

- in terms of storage duration: The processed data may not be kept in a form permitting identification of data subjects for longer than is necessary for the purposes for which the data were collected or for which they are further processed.
- in terms of necessity of the data: The processed data should be adequate, relevant and not excessive in relation to the purposes for which they are collected and further processed. This is also called the principle of *data minimization*.
- in terms of further processing of the data: The purposes of further processing should not be incompatible with the purposes initially defined the data were collected.
- in terms of privacy: to evaluate whether a *limitation to the right of privacy is legitimate*, one should verify whether there is a reasonable relation between the limitation to the right at the one hand, and the legitimate goals that are being strived for at the other hand.

Protocol

A protocol is a set of rules (i.e., formats and procedures) to implement and control some type of association (e.g., communication) between systems (e.g. an internet protocol).

In particular, a protocol is a series of ordered steps involving computing and communication that are performed by two or more system entities to achieve a joint objective.

Provider

A provider is an entity that performs the role of a service provider or an identity provider in an IDM architecture.

Pseudonym

A pseudonym is an identifier that is either self-chosen or assigned, to identify that entity to a relying party within a context. Under certain circumstances, it can be used for improving privacy features of an IDM system, in which case it is typically being used to demarcate the linkability of that identity of the entity to other identities of that entity.

Public-key digital certificate

A public-key digital certificate is a certificate consisting of two parts, namely:

A data part, which consists of the name of an entity, the public key corresponding to that entity, possibly additional relevant information (e.g., the entity's street or network address, a validity period for the public key, and various other attributes). A signature part, which consists of the signature of the certification authority over the data part.

Public Key Infrastructure

The system of certification authorities (CA) (and, optionally, registration authorities (RA) and other supporting servers and agents) that perform some set of certificate management, archive management, key management, and token management functions for a community of users in an application of asymmetric cryptography is called a 'Public Key Infrastructure' (PKI).

Registration

Registration is the process of collecting and corroborating a specific set of attributes of an entity, which typically relate to the partial identity (e.g., the age), a characteristic or a mandate of that entity, with sufficient certainty, before putting at the disposal means by which the entity can be authenticated, or the characteristic or mandate can be verified.

Registration authority

The registration authority (RA) is the entity entitled and trusted to perform the registration service, i.e., the service of identifying entities and registering them in a way that allows the secure assignment of credentials to these entities.

In a digital signature context, registration authorities (RA's) are entities separate from the CA's that, unlike the CA's do not sign either digital certificates or certificate revocation lists, but have responsibility for recording or verifying some or all of the information (particularly the identities of subjects) needed by a CA to issue certificates and certificate revocation lists, and to perform other certificate management.

Relying party

A relying party relies on the results of the authentication to establish the identity or an attribute for the purpose of a transaction (e.g., an eGovernment service).

Resource

Resources can be classified as computing and non-computing systems and services.

Computing systems and services are for example offering disk space on a file server, electronic mailboxes, the system software, applications, services, data repositories, data objects and so on.

Risk

Risk is the level of impact on possible operations of the entity (including mission, functions, image, or reputation), assets of the entity, or individuals resulting from the operation of an information system given the potential impact of a threat and the likelihood of that threat occurring.

Risk assessment

Risk assessment is the process of identifying risks and evaluating them.

Risk management

Risk management is the process of identifying risk, assessing risk, and taking steps to reduce risk to an acceptable level. The objective of performing risk management is to enable the organization to accomplish its mission(s):

- by better securing the IT systems that store, process, or transmit organizational information;
- by enabling management to make well-informed risk management decisions to justify the expenditures that are part of an IT budget; and
- by assisting management in authorizing (or accrediting) the IT systems on the basis of the supporting documentation resulting from the performance of risk management.

Role

A role is a set of one or more authorisations related to a specific application or service.

(SAML) assertion

In identity management the term "assertion" is often used in reference to a "SAML assertion", i.e., an XML-based standard defining a means for making *assertions* about events, attributes, and policy evaluations concerning subjects.

Sector

A sector is a sociological, economic, legal or political subdivision of society. It is a synonym for an administrative domain. In organizational systems sectors defined by the organization in most cases are part of exactly one communicational context. Nevertheless, a sector can also map with several contexts.

Security Policy

A security policy is a set of rules and practices that specify or manage how a system or organization provides security services to protect sensitive and critical system resources and govern the use and provision of security services and facilities.

Service Level Agreement

A service-level agreement (SLA) is a contract between one or more service providers and one or more customers that define the services provided, the metrics associated with these services, acceptable and unacceptable service levels, liabilities on the part of the service provider and the customer, and actions to be taken in specific circumstances.

Service provider

A service provider is an entity that acts in its function to provide services to principals or other entities of an IDM architecture.

Subject

A subject is a possibly acting entity, such as a natural person, a legal person or a computer.

Third Party

A third party is an entity other than the two or more entities initially communicating, which is alien to their internal relationship.

Trust

Operational definitions of trust require a party to make a rational decision based on knowledge of possible rewards for trusting and not trusting. Trust enables higher gains while distrust avoids potential loss. In identity management "Trust" is typically understood in its operational sense.

An entity can be said to trust a second entity or a system when it makes the assumption that the second entity or system will behave exactly as it expects. This assumption is shared by all those in an exchange. Trust may apply only for some specific actions.

Trusted third party

A trusted third party (TTP) is a third party trusted by other entities to perform one or more specific actions within a specific context.

Unlinkability

Unlinkability is the state in which two items of interest (typically sets of personal data) in an identity management system are not more related after the observation by an attacker than they were related taking into account the a priori knowledge.

Unobservability

Unobservability is the state of one or more items of interest being undetectable from any item of interest of the same type at all. Undetectable means that the existence of an item of interest cannot sufficiently be distinguished by an attacker.

User

A user is an entity that interacts with another entity in a specific context. It can be either external to or a member of that entity.