FIDIS

Future of Identity in the Information Society

Title:	“D5.4: Anonymity in electronic government: a case-study analysis of governments’ identity knowledge”
Author:	WP5
Editors:	Bert-Jaap Koops, Hans Buitelaar, Miriam Lips (TILT, the Netherlands)
Reviewers:	Ruth Halperin (LSE, UK) Hans Hedbom (KU, Sweden)
Identifier:	D5.4
Type:	Deliverable
Version:	1.0
Date:	29 May 2007
Status:	[Final]
Class:	[Public]
File:	fidis-wp5.del5.4-anonymity-egov.doc

Summary

The objective of this deliverable is to provide a first attempt to answer the question whether citizens becomes more anonymous or more known by the government when digital identification and authentication technologies are applied in the process of public service provision. From a historical-philosophical and a sociological angle, arguments are broached that may contribute to finding a foundation for answering this question. The issue is then addressed through case studies which illustrate different aspects of the matter at hand. The case studies allow an assessment of the state of anonymity of the citizen in the electronic as opposed to the paper-based relationship with the government. The cases vary from an organisational through a more technical (e.g., PET techniques) to a legal, data-protection perspective.

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

Versions

Version	*Date*	*Description (Editor)*
0.1	27.04.2007	• Initial release of complete version; editing (HB)
0.2	04.05.2007	• Editing; Conclusion inserted; completed for internal review (BJK)
0.3	28.05.2007	• Reviewers and authors' remarks and comments inserted (HB)
0.4	29.05.2007	• Final editing (BJK)
1.0	31.05.2007	• Final version (BJK)

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)*
Executive Summary	Bert-Jaap Koops (TILT)
1 Introduction	Hans Buitelaar (TILT)
2 Historical-philosophical background	Paul de Hert (VUB)
3 Sociological background	Isabelle Oomen (TILT)
4 Case studies permit and border control Netherlands	Isabelle Oomen (TILT)
5 Case study e-signatures Germany	Martin Meints (ICPP)
6 Case study IDM Belgium	Xavier Huysmans (K.U. Leuven, ICRI)
7 Case study medical data Switzerland	Bernard Anrig, Emmanuel Benoist, David-Olivier Jaquet-Chiffelle (VIP)
8 Conclusion	Bert-Jaap Koops (TILT)

Table of Contents

Executive Summary

Government needs information to govern. Particularly in direct relations with citizens, they need information from citizens in order to do their job. In many of these situations, governments want to know exactly which citizen they are dealing with, and hence, they require identification. This identification desire is strengthened by the development of personalisation and customised public service provision. Technology is a prime facilitator of this process, allowing for increasing possibilities, in an ever wider variety and depth, of government-citizen interaction.

The research question addressed in this report is: *Does the identity knowledge of the government grow through the development of e-government?* 'Identity knowledge' is the collection of descriptive information that is connectable to an individual; it is a bipolar continuum with identifiability at one end and anonymity at the other, which can be assessed by Marx' five concentric circles of identity information (core, unique, sensitive, private, and individual information). The identity knowledge capacity of governments (i.e., the amount of data, the centralisation of those data, the speed of information flows, and the number of points of contact between government and citizen) is another concept that can be used in answering this research question.

In this report, we have tried to give a first, tentative answer by analysing various case studies. These have been selected to illustrate the broad range of e-government and public services in Europe: different sectors, small-scale to large-scale, across a range of technological measures, in four European countries. Some cases are descriptive of current, concrete processes, while others focus more on analysing more abstract problems in public service provision and how technology can be used to overcome these. The scope of this study does not allow for a systematic, comparative approach; rather, a heuristic combination of cases is used to provide a caleidoscopic view of anonymity and identification in eGovernment. Future research with a more systematic approach is needed in order to offer a more complete and firmer answer.

The conclusions in the case studies are that, first, new technologies change the identification process in the Netherlands; in one case, this led to a small increase in the identity knowledge capacity of the government, in another case not. Second, in Germany, the transition from paper-based to electronic signatures leads to a small increase in identity knowledge, since signature certificates include data like the citizen registration number and date of birth, and pseudonymous certificates are not allowed for citizens to use in official government procedures. Third, in Belgium, the identity management infrastructure for e-government under construction focuses on interoperability and efficient information management, with a single global identifier to identify all citizens across several contexts; although use of this number is regulated by law and and the Privacy Commission, significant opportunities are created for increasing the knowledge capacity of the government in future. Fourth, in Switzerland, anonymisation of medical statistical data shows that technology can facilitate non-identifiability of patients while still allowing the same patient to be followed through different treatments in time and space. Linkability can be effected without identifiability, through cryptography-generated pseudonyms.

The tentative answer to the research question that can be given on the basis of this limited collection of case studies, is that citizens indeed become more known by the government when digital identification technologies are applied in the process of public service provision, if only moderately so. The knowledge capacity of governments grows in the transition from paper-based to electronic communications, although currently only to a minor extent. The identity knowledge increases with some more personal data of citizens, such as birth date or a photograph. There may be some cause for concern in this from the piont of view of data protection, but we should not exaggerate the threat to privacy that this poses. This is a tentative conclusion, which more systematic future research in this area could try and refine or adapt.

This seems only part of the answer, however. What also emerges from the case studies, is that identification infrastructures are slowly being built, often centring on single and global identifying numbers, rather than sector-specific numbers, which facilitate data mining and profiling, even if this occurs infrequently today. Given the tendency of governments to call citizens ‘customers’ and to stress personalisation, which encompasses a ‘natural’ desire for identification and increased identity knowledge in order to improve ‘customisation’, it should be researched to what extent these technological possibilities will be exploited by governments for public service provision in the near future. Part of this future research should also be the study of technological and organisational means to counterbalance the increased identity knowledge of governments, such as anonymisation techniques, credentials, and smart pseudonym systems. Privacy-friendly identity management in e-government is not likely to happen by itself, but is definitely worth exploring and stimulating by further research.

1 Introduction

With the introduction of e-government, the importance of digital identification assumes ever greater proportions. Whereas in the past the citizen and the public service provider got into contact with each other in the front office, e-government concepts attribute an increasingly crucial role to the back office. ICT is an essential part of this service. Standards like efficiency and effectiveness are introduced. Also, the citizen is more and more seen as a client of public organisations.[1] To resolve the question of the lack of identifiability of the citizen in the backoffice, organisations increasingly gather as much personal information as possible to ensure they are dealing with the right person. This development suggests the assumption that citizens are more and more identified by the government through electronic provision of public services. But is the difficult balancing act between being anonymous[2] and being identified as a citizen in the government-citizen relationship really changing, when compared to the traditional physical world? And if so, to what extent?

This leads us to the following question which we hope to answer in this deliverable:

> *Does the citizen become more known by the government when digital identification and authentication technologies are applied in the process of public service provisions? In other words, is the identity knowledge of the government growing through the development of e-government?*

The first two chapters provide a tentative theoretical foundation for the discussion that follows. First, in Chapter 2, a historical-philosophical approach to the question of the relationship between liberty and identifiability is explored in the utilitarian philosophical school which Jeremy Bentham developed. He provided a philosophical backing to a tradition of identification and control. Bentham tries to argue that the increased control that the state got over its citizens by using more and more identification techniques (the tattoo on the wrist!) actually enhanced his personal liberty: early or preventive identification of crimes and criminals would make law enforcement less intrusive and therefore make society more liberal. Second, in order to provide a basis for answering the question of this deliverable, and zooming in on the public-service delivery part of government, in Chapter 3, the sociological concept of identity knowledge is introduced. Identity knowledge refers to the descriptive information that is connectable to an individual. Furthermore, this chapter describes a methodology which allows some form of measuring of the identity knowledge capacity of an organisation.

The second part of the deliverable presents several case studies, in Chapters 4 to 7. These have been selected to illustrate the broad range of e-government and public services in Europe: different sectors (e.g., environmental planning, border control, health care), from small-scale (tree-felling permit) to large-scale (identity management in general), and across a range of technological mechanisms (password protection, biometrics, identity cards, cryptography), in different countries (Netherlands, Germany, Belgium, and Switzerland). Some cases are more descriptive of current, concrete processes (such as the Dutch cases), while others (such as the Belgian case) focus more on an analysis of more abstract problems

[1] See, for example, Lips, A.M.B., Hof, S. van der, Prins, J.E.J., & Schudelaro, A.A.P. (2005), *Issues of Online Personalisation un Commercial and Public Service Delivery*, Nijmegen: Wolf Legal Publishers.

[2] On the meaning of anonymity, see Pfitzmann, A., Hansen, M., Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology, v0.28, May 29, 2006, p.6. http://dud.inf.tu-dresden.de/Anon_Terminology.shtml.

in public service provision and how technology could be used to overcome these. The scope of this study does not allow for a systematic, comparative approach; rather, we have used a heuristic combination of cases. Together, they provide a caleidoscopic – fragmented but illuminating – view of anonymity and identification in e-government that needs to be extended in future research.

Two case studies from the Netherlands use the sociological methodology of Marx and Rule to show how arguments can be drawn from recent Dutch e-government developments to determine whether the citizen becomes more or less anonymous during the process of identification in an e-government setting versus a traditional paper-based setting. Another approach is a description of the transition from paper-based signatures to electronic signatures and the resulting identifiability and digital linkability of (trans-)actions through or via signature certificates, illustrated in the German context. It turns out that privacy enhancing measures are not easily at hand to prevent possible security leaks in the case of unauthorised linkability. In a more extensive discussion of the Belgian identity and data-management building blocks of federal e-government, a contribution is made to the documentation of the idea of what is possible in IDM for e-government. It turns out that a ready and easy answer is not yet available, but that the choice for a single global identifier for all citizens is not without risk. Finally, the use of statistical information in the Swiss health sector is discussed, especially data on treatments of patients at hospitals. If data are collected and stored in their original form by a central governmental agency, privacy problems arise due to the sensitivity of these data that typically carry uniquely identifying content. This chapter shows that identification data need not necessarily be processed even when exploiting new opportunities of electronic data processing to generate statistical data throughout patients' lifetimes. It is argued that data must be anonymized before being used for statistical procedures either by the agency itself or external entities (like statistical offices, research labs, etc. which are authorized to use these data).

As the scope of this study does not allow for a systematic, comparative approach, our limited combination of cases does not allow for definitive answers, but the objective of this report is to give a tentative first answer to the research question, which is done in the concluding chapter. It is hoped that subsequent research can refine and adapt this answer.

Part I. Theoretical backgrounds

2 Jeremy Bentham on the need for identification by governments

"Personal liberty would be enhanced by rendering it possible to relax the rigor of criminal proceedings. Such imprisonment as was directed merely to securing the presence of the prisoner during the pendency of process would become rare, when the man was known to be detained, so to speak, by an invisible chain." (BENTHAM, J., 'Principles of the Penal Code', Part IV (Chapter LIV), 259)

In this chapter, an introduction is provided to a philosophical, utilitarian foundation for the concept of identification by governments. Jeremy Bentham has argued that identification by governments actually adds to individuals' liberty. Even though this may seem a *contradictio in terminis*, Bentham's arguments for this view on identification are refreshing and remarkably modern, which make it worthwhile to present his views as a general background for this report.

2.1 Bentham on human rights and liberty

Jeremy Bentham (1748-1832) is well-known for his refutation of the idea of human rights, such as those used in the French Declaration of rights. Natural rights, he writes, are 'simple nonsense' and 'natural and imprescriptible rights' were 'nonsense upon stilts'.[3] This will be recalled in *Theory of Legislation*, a work that is central in this chapter.[4]

In this book, Bentham holds that natural laws are the product of the imagination of those that invoke them and anyone may lay down what he pleases (Principles of Legislation, Ch. XIII). Liberty can be secured only where 'real' rights are established through a legal system. To work out such a system, Bentham proposes to use the utility principle rather than 'natural rights' to resolve conflicts.

'The utility or interest of an individual' are the basic ingredients of Bentham's societal calculus. 'The science of legislation consists in determining *what makes for the good of the particular community whose interests are stake*, while its art consists in contriving some means of realization' (Principles of Legislation, Ch. I).

According to Bentham, the 'common interest' to be furthered by his schemes for reform 'corresponds to the immediately subordinate right and proper ends of government'.[5] These ends or objects which the legislator should seek to attain are *security*, *subsistence*, *abundance* and *equality*, the first being the most important (Principles of the Civil Code, Part I, Ch. 2).[6]

[3] Bentham, J., 'Anarchistic Fallacies', in *The Works of Jeremy Bentham*, Bowring, J. (ed.), Edingburg, William Tait, Vol. II, 501. See equally: Bentham, J., 'Principles of Legislation' in *Theory of Legislation*, (1802) Oxford, Oxford University Press, Vol. I, 1914, Ch. III, 11. See also: Waldron, 1987; Rosen, 1987; Freeden, 1991.

[4] *Theory of Legislation* consists of two volumes. Cf. Bentham, J., 'Principles of Legislation' and 'Principles of the Civil Code', in *Theory of Legislation*, (based on manuscripts written around 1788, collected by E. Dumont and first published in 1802), Oxford, Oxford University Press, Vol. I, 1914, 310p.; Bentham, J., 'Principles of the Penal Code', in *Theory of Legislation*, (based on manuscripts written around 1788, collected by E. Dumont and first published in 1802), Oxford, Oxford University Press, Vol. II, 1914, 362p.

[5] 'Subordinate' in this context meant subordinate to the greatest happiness of the greatest number. Cf. Gunn, 1986.

[6] Each of these objects must, more or less, make sacrifices to the others, and the adjustment of their conflicting claims presents a problem extremely difficult of solution. Bentham recognizes 'security' as fundamental. It is the fount of life, of subsistence, of abundance, of happiness. When, for example, *security* and *equality* are in opposition, there should be no hesitation -*equality* must give way. Like liberty it is no more than a chimera. Cf. Principles of the Civil Code, Part I (Ch. 2); (Ch. 3) and (Ch. 11). See Gunn (1986) for reference to other works of Bentham containing the same message on the 'four ends of government'.

Other 'objects', such as equality, justice, and liberty, command respect and ought to enter into the views of the legislator, but they must be subordinate to the happiness and security of the community (Principles of Legislation, Ch. IV).

'Some persons', Bentham says, 'may be surprised to find that 'Liberty' is not ranked among the principal objects of the law', but 'we must regard it as a branch of "Security" to avoid confusion'. '*Personal liberty* is security against a certain class of wrongs which affects the person; while what is called *political liberty* is also a branch of security -security against injustice at the hands of the persons entrusted with government' (Principles of the Civil Code, Part I, Ch. 2). Liberty or individual liberty are not part of Bentham's priorities. Liberty is a 'chimera' in the world of politics, a passion building up to fanaticism, that blinds men at a point where they do not trouble 'whether a state is well administered, whether its laws afford protection to persons and property, whether, in a word, its people are happy' (Principles of Legislation, Ch. IV).

In these paragraphs, Bentham comes very close to Hobbes who considered liberty to halt once the social contract was signed. Both discuss civil society wherein the idea of liberty is absent or entirely subjected to the general interest or happiness.[7] Where there is law, there is no freedom, Bentham assumes, and where there is freedom, there can always come law, since freedom is one of the goals that can be subjected to the happiness and security of the community. More so than Burke, merely opposed to 'abstract freedom', Bentham is the first post revolutionary thinker to make liberty completely disappear from the legal domain. What remains of it should be legally considered as a branch of 'Security' to avoid confusion.

The sum total of the Benthamite liberty is very different from the idea of complex freedom. In his conception of liberty there is only the question of liberty understood as a privacy right, viz. as a shield against intrusion and interference by other people or by holders of authority. Liberty understood as the ideal of the free or autonomous person, giving full weight to the individual's willingness and tendency for self-development, is left out of Bentham's legal picture. This approach to liberty necessitates two or three remarks.

Firstly, it is important to recognize the influence of Montesquieu's and Beccaria's work, with which Bentham was familiar.[8] Bentham's use of their narrow liberty concept can help to explain the poor treatment of liberty in most western nineteenth century constitutions. Affirming liberty was not felt to be a necessity anymore. The fathers of the Belgian 1831 Constitution considered the legal job concerning liberty done, by drafting an article protecting the citizen against unlawful arrest.[9]

Secondly, Bentham's 'proper ends of government' do not correspond wholly with the 'ends' of the social contract enumerated in the French and American basic constitutional

[7] For a friendly comment on Hobbes by Bentham, see Principles of Legislation, Ch. XIII. There is however a difference of degree. Bentham recognizes that every law curtails in one way or another liberty and is therefore likely to be followed by a feeling of pain. Hence, law should only be enacted when it is necessary and when the motives are more weighty than the general reason against all coercive legislation. Cf. Principles of the Civil Code, Part I, Ch. 1.

[8] We recall that for Montesquieu, liberty is primarily a right to be safeguarded against others and is therefore (nothing more then) 'a right of doing whatever the law permits' or a right to 'safety or tranquility of mind'. Although Beccaria elaborated on a larger concept of liberty, he did not lay the emphasis on it and concluded his work with a political axiom in very much the same terms as Montesquieu.

[9] Cf. art. 7 the Belgian 1831 Constitution. See Vanderlinden, 1992.

documents.[10] The expressions *subsistence* and *abundance* are absent and *security*, Bentham's primary end, is equally lacking. To 'effect' 'safety of the people' is clearly one of the tasks of government in the American Declaration, but this has to be done while 'securing the said unalienable rights'. The French *Déclaration* is even less ambiguous. There is no other 'end' of government than this of preserving the said natural and inalienable rights. 'Security' is clearly not on the list and it cannot be equated with the notion of *sûreté* used in article 2 of the French *Déclaration*.[11] If the said article includes 'sûreté' in the list of natural rights, the term was not to be given the actual meaning of the term 'security'. Rather, the National Assembly meant by it a guarantee against arbitrary state interventions in the life of the citizen. The intention was Beccarian, viz. to provide for protection against the so-called *lettres de cachet*, not to target delinquency (Bonnemaison, 1987). Hence, Article 2 is essentially saying no more than that the citizen should be protected against actions of the government. Even when interpreted very broadly[12], the said Article can under no circumstances be read as stating that security is a primary goal of government, implying far going subjection of all individual rights and implying the cooperation of the citizen to fulfil this task. Bentham, on the contrary, seems to assume that without government there will be no security and hence no individual rights.

Thirdly, and more technically, one should understand Bentham's approach to liberty by taking into account his narrow theory of rights. We already discussed his famous attack upon natural and inprescriptible rights (*supra*). Correlative to this is his adherence to what is later termed 'the benefit theory of rights' (Freeden, 1987). A right is the legal expectation of the discharge of a legal duty, intended to benefit the right-bearer. A legislator can only distribute rights or obligations, Bentham holds, and there cannot be question of a right, without there being question of a duty imposed on somebody else.[13] The second part of the proposition can be questioned. In our opinion, nothing stands in the way of the recognition of legal goods that do not impose duties on particular persons.

2.2 The principle of utility

We wrote that Bentham proposes to use the utility principle instead of 'rights' to resolve conflicts. 'The right end of all human action is', says Bentham, 'the creation of the largest possible balance of happiness', and this tendency to produce happiness is what he meant by *utility*. For Bentham, the morally right action in any circumstances is the one that will tend to maximize total happiness. Each individual counts equally in the calculation of how much pleasure is produced by an action, and the total of pleasurable states is summed to determine how we should act. Bentham's utilitarianism aims at achieving the greatest aggregate

[10] In both declarations, the 'ends' of the contract consist in preserving the natural and inalienable rights of man. These rights are in the American Declaration: life, liberty and the pursuit of happiness' and in the French *Déclaration*: liberty, property, 'sûreté' and resistance to oppression.

[11] Article 2 *Déclaration*: 'Le but de toute association politique est la conservation des droits naturels et imprescriptibles de l'homme. Ces droits sont la liberté, la propriété, la sûreté et la résistance à l'oppression'.

[12] Namely, in the sense that governments should protect people's safety against intrusion and the actions of other citizens.

[13] 'Objects which the legislator is called upon to distribute among the members of a community may all be reduced to two classes: (1) Rights; (2) Obligations. (...) Rights and obligations, although distinct and, indeed, opposite in character, nevertheless arise at the same moment, and throughout their common existence remain inseparable. In the nature of things, the law cannot confer a benefit upon anyone without at the same time imposing a burden upon someone else' (Principles of the Civil Code, Part I, Ch. 1).

happiness, that is the largest total sum of happiness irrespective of how that happiness was distributed.

This utility principle is the foundation of all Bentham's schemes of legislation. 'The end and aim of a legislator should be the happiness of the people. In matters of legislation, general utility should be his guiding principle' (Principles of Legislation, Ch. I). Nature has placed mankind under the governance of two sovereign masters, *pleasure* and *pain*. To these two motives the *principle of utility* subjects everything: the principle is respected on an individual or collective level when 'anything' tends to augment pleasure.[14] Hence, (criminal) law becomes 'a simple piece of practical business'. Measuring pleasure and pain allows one for to measure the good or evil of any action;[15] manipulating pleasure and pain allows for controlling the behaviour of man.

This brings us to Bentham's theory of punishment. Bentham conceived certain pains and pleasures so annexed to actions as to form bonds, constraining a man, as it were, to the observance of some particular rule of life or conduct. Hence he believed that the whole duty of man might be enforced by the operation of 'physical', 'political or legal', 'moral or popular', and 'religious' *sanctions*. 'Many men', says Bentham, 'fear the wrath of Heaven; many men fear loss of character; but all men are acted upon, more or less, by the fear of the gaol, the scourge, the gallows, the pillory, and so forth" (Principles of Legislation, Ch. VII). This quote gives us the best possible illustration of Bentham's theory of punishment formulated as a genuine theory of social control.

2.3 Bentham on social control

In his *Theory of Legislation*, Bentham proves to be a master of social control thinking, fully mastering contemporary terminology such as 'surveillance', 'observation', etc., and the diversity of the actors involved. He is a prime advocate of full and extensive use of all possible sanctions,[16] *and* of Hood's four key resources of government power: nodality or centrality in an information network, treasure, authority and organization (Hood, 1983). In the following, we select some of Bentham's abundant suggestions to improve the mechanisms of social control.

A nice starting point is his discussion of postal delivery. How can a government speed up this service? In order to speed up postal delivery and laxity of carriers, Bentham suggests, government should avoid using authority and organization. By relying on nodality and combining carriage of mails and the conveyance of passengers, the travellers who then accompany the letter-carrier will become inspectors of his conduct. They will impose constant and unpaid surveillance on them and are ready to inform government on negligence of duty.

[14] 'When we say that anything is in harmony with the utility or the interest of an individual, we mean that it tends to augment the sum total of his well-being. When we say that anything is in harmony with the utility or the interest of a community, we mean that it tends to augment the sum total of the well-being of the individuals of which the community is composed' (Principles of Legislation, Ch. I).

[15] 'The *Principle of Utility*, accordingly, consists in taking as our starting-point, in every process of ordered reasoning, the calculus or comparative estimate of pains and pleasures, and in not allowing any other idea to intervene' (Principles of Legislation, Ch. I).

[16] 'The Legislator should, however, never lose sight of the fact that he has under his direct command the *political* sanction only. The other three must, of necessity, be his rivals or his allies, either hostile to him or subservient to him; and should he leave them out of his calculations, his results will be full of error. If, however, he can bring them to unite in support of his aims, he will wield enormous power; but his only chance of joining forces is under the standard of Utility' (Principles of Legislation, Ch. VII).

General happiness will increase. The state saves on information and legal proceedings and the services will be rendered quickly and cheaply (Principles of the Penal Code, Part IV, Chapter LIII).

Bentham is very keen on nodality and state use of information. It allows him, without making him too illiberal, to propose some very illiberal measures to improve citizens or enforce virtue. Public virtue, belief in the general Principle of Utility and patriotism are for Bentham important prerequisites to keep people away from over privatization, the sentiment of benevolence and bastard patriotism; states should not refrain from 'disabusing' the minds of people to relate this message. Deceit, trickery and 'secret controlling of public opinion' are by no means illegitimate. Sometimes a mere change of *name* is enough to overcome resistance. If people accept the term *Emperor*, but not the term *King*, use the former. The legislator may manipulate the deceptive power of terms such as *liberty*, *equality* and *subjects*,[17] and use the force of example to alter public opinion, or simply wait till the mind of the people changes.[18]

Authority should be used to establish all sorts of registries, titles and stamps, certificates and so on. Since everyone wishes to obtain for himself all possible security about his legal (trans)actions, most people will comply. 'The utility of authentic attestations of this character is beyond all manner of dispute.' Property owners and consumers have more security about their transactions *and* the whole will facilitate the payment of taxes and the combat against crime and smuggle (Principles of the Penal Code, Part IV, Chapter LII).

Authority should also be used to adopt coercive laws to regulate behaviour of men. Bentham proposes to 'refine' the tool of legislation. Everybody knows that offences can be combated by punishments (*direct legislation*), but the Sovereign should also consider the use of preventive laws (*indirect legislation*). Direct legislation assails the mischief by means of a frontal attack; indirect legislation has recourse to what he calls 'oblique' methods.[19]

One of Bentham's central propositions on social control is 'To prevent Offences by making it the Interest of Many Persons to prevent them' (Principles of the Penal Code, Part IV, Chapter LIII). Additionally and if needed states should make use of treasure. Bentham makes strong arguments in favour of rewards for those that assist in the prevention of crime and in delivering the guilty person into the hands of justice, and in favour of paid informers and government spies (Principles of the Penal Code, Part II, Chapter XVI). Using spies may be one of the indirect measures of a 'refined legislator'.

[17] 'Sometimes a mere change of *name* is enough to change the sentiments of a nation. The Romans abhorred the name of *King*, but they would tolerate the style of *Dictator* or *Emperor* (...) Peter the Great relinquished the use of the description 'despot' and directed that the slaves of the nobles should thenceforth be called 'subjects'. (...) What deception lurks in the words *liberty* and *equality!*' (Principles of the Penal Code, Part IV, Chapter LIX).

[18] One illustration is given by electronic monitoring of offenders. This technique was already under debate in 1967, but public opinion was against it; only now, at a time of severe prison crowding, these techniques are becoming acceptable Comp. 'That period was at the height of the Warren Court, with its growing emphasis on due process and individual liberties, and so such technologies were viewed to be unacceptably invasive of individual privacy. Today, however, at a time of severe prison crowding, they are viewed as an important means of solving the prison crowding problem' (Blumstein, 1988 at p. 9).

[19] 'In the first case the legislator openly declares war against the enemy, announces his approach, pursues the foe, fights him hand to hand, and scales his batteries in broad day. In the second case he does not make known his plans. He lays mines, sets spies to work, seeks to frustrate the designs of the enemy, and to secure an alliance with those who might otherwise have harboured hostile intentions' (Principles of the Penal Code, Part IV (Introductionary Chapter)).

Treasure may also be used to pay for public instruction in order to enlighten and 'disabuse' the minds of the people, and to shape the public opinion or 'control openly' public opinion through a system of rewards to confer additional honours on those who are adjudged fit to receive them (Principles of the Penal Code, Part IV, Chapter LIX).

Of course, when nothing works there is always organization and force. One part of organization is the physical interference in the life of the governed. Bentham sees an important role for private policing (*infra*) and identifies only practical and financial boundaries for the growth of government and police organizations.[20] One of the boundaries that the Sovereign should respect is the public opinion. He has to take care not to shock the national sentiment and not to harass the people with police measures in times of tranquillity. In the capital of Japan everyone is required to bear his name on his outer garment, Bentham observes, 'a precaution which will seem desirable, idle, or arbitrary, according to the trend of popular prejudices'. Bentham suggests that people should have their names tattooed on their wrists, since such measure would eliminate anonymity and render it possible to relax the rigor of criminal proceedings. Unfortunately, 'the state of public opinion nowadays presents an insurmountable obstacle to the adoption of this practice'. However, not all is lost. Propaganda and example can do a lot and 'opinion might undergo a change if we devoted to the advocacy of such a scheme a good deal of patience and tact' (Principles of the Penal Code, Part IV, Chapter LIV).

2.4 Bentham on identification and control

Bentham is very firm about the superiority of legislative schemes that concentrate on identification. Early or preventive identification of crimes and criminals does not only help criminal law enforcement but also renders it less necessary: 'The bulk of offences would never be committed at all but for the hope which the culprits harbour that they may remain undiscovered. Everything which tends to improve the expedients for the Identification and Discovery of men engaged in crime adds to the general security' (Principles of the Penal Code, Part IV, Chapter LIV). This is underlined by his famous proposal to organize state propaganda for measures such as names tattooed on the wrists of people to identify them (*supra*). The existing customs around the names of individuals, Bentham observes, 'stand upon such an unsatisfactory footing' in a great nation with risk of confusion in names. And he continues to fancy about a 'new nomenclature of such a character that every individual in the whole nation should have a proper name, borne only by himself' (Principles of the Penal Code, Part IV, Chapter LII).[21]

In a comparative and historical perspective, Bentham could be labelled as a thinker giving philosophical backing to a tradition of identification and control that started decades before his time. Indeed, the main ingredients of the high policing model were already defined in the seventeenth and eighteenth century (Brodeur, 1983).[22] In a 1749 memoir on the reform of the

[20] One limit to organization is laxity (Principles of the Penal Code, Part IV, Chapter LIII) . Another are the charges for the maintenance and entertainment of the police (Principles of the Penal Code, Part IV, Chapter LIII). A third one is the care not to shock the national sentiment: see *above*.

[21] 'It would be quite feasible to establish a new nomenclature of such a character that every individual in the whole nation should have a proper name, borne only by himself. In a state already organized it may be that the difficulties arising from such a change would outweigh the advantages, but in a newly-formed colony it would be desirable to avert any confusion of this kind' (Principles of the Penal Code, Part IV, Chapter LIV).

[22] High policing, defined as policing potential threats in a systematic attempt to preserve the distribution of power in a given society (*supra*), is first of all absorbent policing. It aims to control by storing intelligence. This

French police Guillaute, an officer of the *Maréchaussée* (provincial police), made plain that gathering intelligence -the core feature of high policing- is accomplished not only through the accumulation of data on potential threats; but needs also to be enhanced by exhaustively charting the physical and social space into definite coordinates in order to increase the scope and precision of surveillance.[23] Also in 1749, a decree was proclaimed by the French *Ancien Régime* government obliging workers to be in possession of a certificate of their employer, a certificate that was later on transformed into an identity card.[24] This powerful tool in the hands of the employer and the local authorities, subjected the labour force to the discretion of the employers,[25] and restricted strongly their physical liberty: without the certificate travelling workers were considered to be vagrants and could be punished. Briefly abolished by revolutionary decrees, the system was re-established in 1803 and only definitely abolished in 1890, when softer techniques of identification and control of the labour masses were available (Heymann-Doat, 1994).[26]

A closer look at the situation in England may correct the foregoing image of Bentham as a philosophical advocate of traditions of identification and control that started before his time. Seen in this perspective, Bentham is actually very much a thinker of his days addressing concrete problems or future problems. One of the problems facing England precisely concerned the question of identification. Serfdom in the middle Ages and the Elizabethan Poor Law afterwards had created a perfect system of immobility of the poor and the dangerous.[27] Legal weakening of this system and larger phenomena such as the Napoleonic wars, the industrial revolution and the building of the railways together with the refusal of Britain's eastern colonies to accept any more convicts,[28] created population mobility. This in turn created a new problem for governmental social control. Hence the system of penal servitude, the creation of criminal history records, and the acceptance by the government of new identification techniques (Hebenton & Thomas, 1993).[29]

intelligence-gathering is all-encompassing: it extends to any domain that may further the implementation of state policies.

[23] Brodeur, 1983,with references. Guillaute therefore not only outlined the first automated system of police records but also made elaborate proposals for numbering houses, apartments and stairways; for identifying all vehicles and for listing occupations travels hotel occupancy and so forth, proposals that became reality afterwards.

[24] On the '*certificat de congé*' and the '*livret ouvrier*': Heymann-Doat, 1994.

[25] Without the certificate of with a negative evaluation of the former employer in it, workers could forget about finding other work.

[26] See also Heymann-Doat (1994) on the introduction of an identity card for foreigners by a Bill passed in 1792 and on the gradual sophistication of the control and identification techniques used to surveilling foreigners in France.

[27] Hebenton & Thomas, 1993. The Elizabethan Poor Law system made individuals live the whole of their life in one parish even after serfdom was abolished. The system made each parish responsible for its paupers and in consequence the overseers of the poor took great care to ensure that no-one settled in the parish if he might become a burden to it. It was not until 1795 that an act was passed making persons irremovable from the parish until they were actually in need of poor relief. Some movement of persons did, of course, occur but they were usually either those who were unlikely to become a burden because of wealth or because they possessed a certificate from their home parish accepting poor law responsibility for them if required, or they were those who migrated to the large towns where the movement of individuals could not be effectively controlled.

[28] After the Poor Law Amendment Acts of 1834, 1844 and 1868 the Poor Law restriction to mobility cease to apply. Hebenton and Thomas rightly observe that this is to some extent merely a recognition of changes resulting from these historical phenomena (Hebenton & Thomas, 1993).

[29] In the 1860s, Britain would study the merits of the Irish convict system, using detailed criminal history records, including photographs, and take over some of the ingredients.

'Who are you? With whom am I dealing? There would be no room for evasion in answering this important question' (Principles of the Penal Code, Part IV, Chapter LIV). The result of extensive use of identification techniques, Bentham holds, is altogether liberal: 'Personal liberty would be enhanced by rendering it possible to relax the rigor of criminal proceedings. Such imprisonment as was directed merely to securing the presence of the prisoner during the pendency of process would become rare, when the man was known to be detained, so to speak, by an invisible chain' (Principles of the Penal Code, Part IV, Chapter LIV).

There is some common sense in this argument and we can identify several examples to underpin Bentham's argument about identification and humanism or 'liberalism'.[30] Fundamentally, however, Bentham's analysis is perverse. Only a very thin liberty definition allows justifying the spread of identification techniques in the name of 'liberty'. Bentham seems aware of this, acknowledging that an 'invisible chain' remains a 'chain', and he himself gives a compelling argument against his own scheme. However, his optimism takes the lead again, and he ignores possible risks that might occur in reality in favor of his theoretical analysis.

This is relevant for our purposes, since Bentham's remarkably modern language and frame of mind are mirrored in today's debates about the surveillance society and the control that modern technologies allow governments to have over their citizens. 21st-century ICT infrastructures actually enable an 'invisible chain' to an extent that Bentham might not even have dreamt of. Rather than a wrist tattoo, new identification techniques may well serve in equivalent but less visible ways to chain citizens into compliance with societal norms. Many identification systems currently being developed and applied do not overtly have a surveillance purpose but rather have 'service-delivery' and efficiency goals. From a Benthamite perspective, however, these modern identification infrastructures can equally well serve purposes of crime prevention, law enforcement, and controling citizens.

As the next chapters show, for public service delivery, governments need to identify citizens and technology may enable them to enhance the identity knowledge of citizens. What Bentham teaches us, is that the technologies used in public-service delivery for identifying citizens may at the same time serve surveillance purposes, resulting in invisible yet strong chains between government and citizens. This is worth bearing in mind, when studying the identity knowledge of governments in electronic public service delivery.

[30] Two examples. In France in 1850, the criminal record was established and hailed as an ingenious and humanitarian method since it replaced branding (Hebenton & Thomas, 1993). In the *Déclaration* there is only question of freedom to publish, not freedom of press. The Belgian 1831 Constitution contains a much stronger formulated clause concerning freedom of expression and press, compared to Art. 11 of the French 1789 *Déclaration*: all preventive measures are forbidden, no form of censorship may ever be established and no security may be required of authors, publishers or printers, but in the same time an identification technique is introduced, known as 'serial responsibility' and pushing all actors involved in the press world to identify the name of the author. Cf. Article 18.2 of the 1831 Constitution (establishing a specific system of prosecution of press misdemeanors: if the identity of the author is known and he is domiciled in Belgium, the publisher, the printer or the distributor cannot be prosecuted). We cannot help but feel somewhat cynical about this identification-based liberalism or liberalism at the price of opacity. On Art. 18 of the 1831 Constitution, Art. 25 of the 1994 Constitution: ALEN, A. & CLEMENT, J., 'Fundamental Rights and Liberties', in ALEN, A. (ed.), *Treatise on Belgian Constitutional Law*, Deventer, Kluwer Law and Taxation Publishers, 1992, 195-196. A similar 'cascade' system was introduced in France in 1881 (Heymann-Doat, 1994).

2.5 Select bibliography

BENTHAM, J., ‘Anarchistic Fallacies’, in *The Works of Jeremy Bentham*, BOWRING, J. (ed.), Edingburg, William Tait, 1838-1843, Vol. II, 501.

BENTHAM, J. (1802), *Theory of Legislation* (based on manuscripts written around 1788, collected by E. DUMONT and first published in 1802), Oxford, Oxford University Press, Vol. I, 1914, 310p. and Vol. II, 1914, 362p.

BLUMSTEIN, A. (1988), ‘Science and Technology in Support of Criminal Justice’, in LEBLANC, M., TREMBLAY, P. & BLUMSTEIN, A. (eds.), *New Technologies and Penal Justice*, Series *Les cahiers de recherche criminologiques*, Cahier No. 9, Montreal, Centre International de Criminologie Comparée-Universite de Montréal, 2-14.

BONNEMAISON, G. (1987), *La sécurité en libertés*, Paris, Syros, 155p.

BRODEUR, J.-P. (1983), ‘High Policing and Low Policing: Remarks about the Policing of Political Activities’, *Social Problems*, Vol. 30, No. 5.

GUNN, J. (1968), ‘Jeremy Bentham and the Public Interest’ reprinted in LIVELY, J. & REEVE, A. (eds.), *Modern Political Theory from Hobbes to Marx. Key Debates*, London, Routledge, 1989, 199-219.

HEYMANN-DOAT, A. (1994), *Libertés publiques et droits de l’homme*, Paris, L.G.D.J., (third edition), 246p.

HEBENTON, B. & THOMAS, B. (1993), *Criminal Records. State, citizen and the politics of protection*, Aldershot, Avebury, 194p.

HOOD, C. (1983) *The Tools of Government*, London: Macmillan, 178p.

WALDRON, J. (ed.) (1987), *Nonsense upon stilts. Bentham, Burke and Marx on the rights of man*, London, Methuen, 120p.

ROSEN, F. (1987), ‘Bentham’ in MILLER, D. (ed.), *The Blackwell Encyclopedia of Political Thought*, Oxford, Basil Blackwell, 37-40.

FREEDEN, M. (1991), *Rights*, Milton Keynes, Open University Press, 134p.

VANDERLINDEN, J. (1992), ‘Aux origines du Titre II de la Constitution Belge de 1831’, in *Présence du droit public et des droits de l’homme*. Mélanges offerts à J. Vélu, Brussels, Bruylant, Tôme 2, 1992, 1193-1208.

3 Identity knowledge in the governmental provision of services

The citizen-government relationship is partly constituted by the government providing services to its citizens. The provision of services is a standardized process, also known as bureaucracy, and citizens can apply for these services. In this relationship, citizens have to reveal personal information about themselves so that the government knows they are entitled to receive the relevant services. The identification and verification of citizens by the government results in the governernment having identity knowledge about its citizens. This chapter provides a theoretical underpinning of the role of anonymity, identification, and identity knowledge in the relationship between government and citizens in the context of public service delivery.

First, bureaucracy is addressed and the differences between public and private service providers are discussed. Then, personal information in relation to identity knowledge is illuminated. Finally, the identity knowledge capacity of the service provider is discussed.

3.1 Bureaucracy in public organizations

As Weber noticed nearly a century ago, with the rationalization of society, bureaucracy becomes inevitable (Weber, 1968/1921) and in the contemporary society, bureaucracy – whether private or public – is ubiquitous. Without it, few of the routine features of our modern society would be possible; the collection of taxes and the production and distribution of goods and services, for example, would be difficult, if not impossible. (Dandeker, 1990)

Bureaucracies control people by replacing human judgement with nonhuman technology, thus creating a formalistic impersonality of the system. Even bureaucracy itself can be seen as a huge nonhuman technology that functions more or less automatically. Rules, regulations, and institutional structures replace the adaptability of human decisions, that is, employees of bureaucratic organizations generally follow the rules and regulations in a predetermined sequence instead of evaluating each case separately. They must get their jobs done in a certain way by a certain time without mistakes, and the role of informal systems of human action is diminished by the highly formalized structures. Bureaucracy controls not only employees of an organization but her clients as well. An organization provides services and one must apply for the services on a specific form by a specific date. One will receive those services only in a certain way and under strict conditions. (Ritzer, 1998) Client categories used by organizations decide what information a client is supposed to provide, and this information will generate a denial or a grant of a specific requested service. (Snellen, 1998)

Although bureaucracy is present in both public and private organizations, there are large differences between the two types of organizations and the services they deliver. The first difference is the monopolistic character of public organizations, i.e., often a citizen does not have a choice between different public organizations (as is the case with private organizations) because there is only one public organization that provides a particular service. A second difference is that a citizen is not always a voluntary ‘client’ of public organizations because the nation state is responsible for the collective goods. Public services will thus not only increase personal benefits and rights, but will also consist of activities that address the duties of citizens (e.g., tax collection). Third, citizens have a voice in the determination of public service delivery through voting, referenda, and public hearings. (Lips, 1998) The citizen (as citoyen, carrier of democratic rights) enters into debates with political-administrative organizations. When consensus has been established, politicians instruct the public organizations who execute these instructions. The citizen (now as client) uses the services provided by public organizations. (Zuurmond, 1996a; Zuurmond, 1996b) Fourth, public services are subject to specific norms and values, like legitimacy, legal certainty and equality of rights, as a consequence of the government's responsibility for the collective good. This results in the fifth difference between public and private services: continuity in the deliverance of services to citizens and accessibility of public services to all citizens ought to be more important goals for public services than gaining profits. The services provided to the citizen (as client of public organizations) thus have a different character than the services provided to the customer (as client of private organizations).

The view of the citizen as client is not an old view. A few decades ago, the political process of determining the 'business of government' (i.e., determination of public products, services, and information provided to citizens) was perceived as the most important part of the public service delivery. The dominant focus was on the supplier-side of public service delivery; the government knew what was best for their citizens and the government decided what way and form of service provision was most appropriate to address these citizens. This focus shifted gradually to the production and delivery of public goods, services, and information. Standards like efficiency and effectiveness were introduced in public organizations and the functioning of the public organization became most important. During this period, the view of the citizen as a client of public organizations came into being. (Lips, 1998)

Recently, the focus shifted to the feedback of citizens on both wanted and received public products, services, and information. This is at least the case in the Netherlands (Lips, 1998), but it is likely to apply to other national governments as well.[31] Where the government's attitude initially was 'we know what is good for you', it has changed to 'let us know what is good for you'. This view is a result of increased attention to the spending of administrative organizations, and standards like efficiency and effectiveness have played a role as well. Also, concepts and methods with proven success in the private sector, like management, budgeting, marketing, but also service delivery itself, have been introduced in the public sector.

This requires changes in the organization of public service provision. Governments perceive information and communication technologies (ICTs) as an important means to realize these changes. ICT applications in public service delivery potentially bring about not only increased effectiveness, increased efficiency, an improved client-orientation, and a reduction of cost, but also an improved comprehensiveness of information processes of public service delivery between government and the citizen. (Lips, 1998)

The efforts of public organizations to administer to each of many citizens their precise 'due' in terms of the organizational treatments they 'deserve', results in a bulk of demands for personal information. (Rule, McAdam, Stearns, & Uglow, 1980) The large quantity of personal information does not take place solely for the benefit of the clients, but for the benefit of the organizations as well, for organizations are trying to manage risks by gathering personal information in order to establish the kind of person they are dealing with (Lyon, 2001). Today, one of the most obvious indicators of the pervasiveness of bureaucracy is the massive expansion of the personal information which is held by a range of public – and private – organizations. As Dandeker (1990) stated strikingly, "the age of bureaucracy is also the era of the information society" (p. 2).

[31] For instance, to Belgium, see note 45.

3.2 Identity knowledge

In order to obtain certain services, the citizen has to reveal personal information about himself or herself, i.e., the citizen has to identify him or herself. Identity knowledge refers to the descriptive information that is connectable to an individual. (Marx, 2005) It is a bipolar continuum with identifiability at one extreme and anonymity at the other (Marx, 2001). Neither complete identifiability, nor complete anonymity can occur. When a person is completely identified or known, the most inner self of this person is known. This is, of course, impossible, just as complete anonymity is impossible. Although the word 'anonymity' evokes a different image, anonymity requires an audience of at least one person for one cannot be anonymous if there is no form of interaction and if no one is aware of the individual. (Marx, 1999) When a person is completely anonymous, he or she is unknown by everyone i.e., this person does not exist in the social world. So, both identifiability and anonymity are ideal types.

Identity knowledge is fundamentally social; it only gains relevance in a social context. (Marx, 1999) There are degrees of identity knowledge, reflecting not only the amount of information known about an individual but also different types of information about persons for some information is considered to be more sensitive or private than other information. Marx (2005) distinguished five different types of identity information, each at a certain distance from the core identity of an individual. He portrays these five types of identity information as concentric circles around an individual (see figure 1).

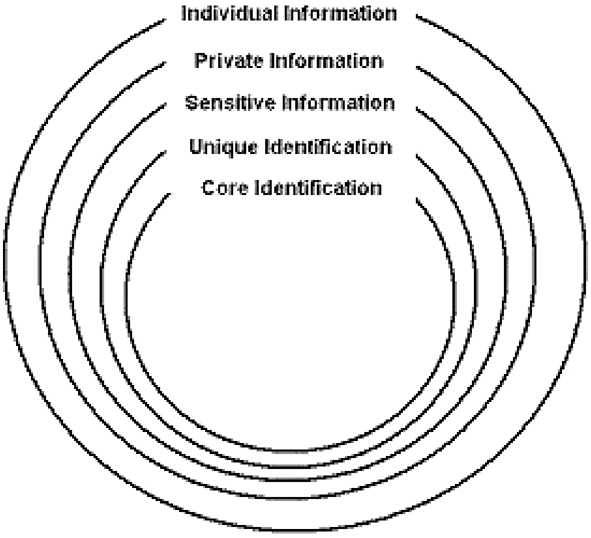

Figure 1: Concentric Circles of Information (Marx, 2005)

The outermost circle is that of *individual information*. This is automatically available information and includes any data or category which can be attached to an individual. Individual information varies from information that is relatively impersonal (e.g., gender, driving a particular type of car, living in this city, shopping in that supermarket) to that which is more personal (e.g., what is in your trolley at the supermarket or sitting in the urology waiting room at the hospital).

In contrast, *private information* is not automatically available. This type of information is unknown by others until it is communicated either by the individual himself or someone else.

Private information can be attitudes, thoughts, or preferences that are not considered to be very personal (for example what sort of food or music you like) but categorize the individual. The social category can be determined by the individual (i.e., self-definition) or by others (as a result of technical profiling for example).

The next circle is that of *sensitive information* and it includes information that is very personal and it takes its significance from the fact that it is only revealed to those we trust and feel close to. The European Union's Data Protection Directive defines sensitive data as those data involving information on race and ethnicity; political, philosophical, and religious beliefs; health; and sexual life.

Even closer to the individual is the *unique identification*. The unique identification of a person consists of various identity pegs, together forming a unique identity (i.e., only you have this particular combination) and answering the question "Who are you?" Elements that make up the individual's uniqueness are, for example, name, birth at a particular place and time, place of residence, and social security number. Apart from these more traditional identifiers, the individual's uniqueness can also incorporate biometric identifiers like fingerprinting, DNA, voice, retina, iris, or facial appearance along with a name or birth date.

Closest to the individual is the *core identification*. (Marx, 2005) The core identification identifies the core identity of an individual, i.e., the most inner person or the Self. The Self is a social being for it is shaped through the social environment, that is, social interactions and relations with others. It consists of multiple identities or multiple roles, derived from membership of various social groups together with a value and emotional significance attached to those groups. (Cooley, 1929; Goffman, 1959; Mead, 1967; Tajfel, 1981; Tajfel & Turner, 1979)

Identity knowledge thus consists of the amount and the type of personal information. When more information about an individual is available, this individual is more identifiable, hence indicating a shift towards identifiability on the identifiability-anoymity continuum. In addition, when the information of an individual is closer to the core identity, the individual is also more identifiable and this indicates a shift towards identifiability on the continuum too.

Now that the amount of identity knowledge and the different types of identity knowledge have been explored, the identity knowledge capacity of a public organization will be addressed. Rule (1973) distinguished four criteria to measure the identity knowledge capacity of an individual organization. The first criterion is the size of the files held in an identity knowledge system, that is, the number of persons and items of information about them that can be stored. The second criterion is the centralization of those files. Highly centralized files make it possible to gather information on a person at any point in the system and then use that information at any other point in the system. Third, the speed of information flows which concerns the time necessary for information on subject populations to be gathered, transmitted, processed, and then used. The fourth and last criterion is the number of points of contact between the system and its subject population. This refers to the number of points in the life of a person that are available for the collection of information. When there are many of those points in a person's life, the organization can maintain a constant and detailed connection between an individual and his or her record or file.

3.3 Bibliography

Cooley, C.H. (1929/1909). *Social Organization. A study of the larger mind.* New York: Charles Scribner's Sons.

Dandeker, C. (1990). *Surveillance, Power and Modernity. Bureaucracy and discipline from 1700 to the present day.* New York, St. Martin's Press.

Goffman, E. (1959). *The Presentation of Self in Everyday Life.* New York: Doubleday.

Lips, M. (1998). Reorganizing Public Service Delivery in an Information Age. Towards a revolutionary renewal of government? In I. Snellen & W. van de Donk (eds.), *Public Administration in an Information Age. A Handbook.* Amsterdam: IOS Press.

Lyon, D. (2001). *Surveillance society. Monitoring everyday life.* Buckingham: Open University Press.

Marx, G.T. (1999). What's in a Name? Some Reflections on the Sociology of Anonymity. *The Information Society, 15, 99-112.*

Marx, G.T. (2001). Identity and Anonymity: Some Conceptual Distinctions and Issues for Research. In J. Caplan & J. Torpey (eds.), *Documenting Individual Identity: the Development of State Practices in the Modern World.* Princeton NJ: Princeton University Press.

Marx, G.T. (2005). Varieties of Personal Information as Influences on Attitudes towards Surveillance. In K. D. Haggerty & R. V. Ericson (eds.), *The New Politics of Surveillance and Visibility.* Toronto: University of Toronto Press.

Mead, G. H. (1967/1934). *Mind, Self, and Society. From the standpoint of a social behaviorist.* Chicago: The University of Chicago Press.

Ritzer, G. (1998). The Weberian Theory of Rationalization and the McDonaldization of Contemporary Society. In P. Kivisto (ed.), *Illuminating Social Life. Classical and Contemporary Theory Revisited.* Thousand Oaks, CA: Pine Forge Press.

Rule, J.B. (1973). *Private Lives and Public Surveillance.* New York: Schocken Books.

Rule, J., McAdam, D., Stearns, L., & Ulgow, D. (2001). *The Politics of Privacy. Planning for Personal Data Systems as Powerful Technologies.* New York: Elsevier.

Snellen, I. (1998). Street Level Bureaucracy in an Information Age. In I. Snellen & W. van de Donk (eds.), *Public Administration in an Information Age. A Handbook.* Amsterdam: IOS Press.

Tajfel, H. (1981). *Human Groups and Social Categories.* Cambridge, England: Cambridge University Press.

Tajfel. H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel (eds), *The social psychology of intergroup relations.* Monterey, CA: Books/Cole Pub. Co.

Weber, M. (1968/1921). *Economy and Society. An outline of interpretive sociology.* (translated by: Fischhoff, E., Gerth, H., Henderson, A.M., Kolegar, F., Wright Mills, C., Parsons, T., Reinstein, M., Roth, G., Shills, E., & Wittich, C.) New York: Bedminster Press.

Zuurmond, A. (1996a). Informatietechnologie: democratisering of technocratisering. *Beleid & Maatschappij,* 3, 134-144.

Zuurmond, A. (1996b). De burger als citoyen. Informeren of communiceren. *Informatie en Informatiebeleid,* 14, 63-66.

Part II. Case studies

4 Identification and anonymity in public service provision: Two case studies in the Netherlands

4.1 Introduction

In chapter 3 of this report, a theoretical system was proposed which might enable measuring the degree of identifiability of citizens in their usage of public services. Building on this theory, in this chapter, two case studies conducted in the Dutch government setting are described in an attempt to discover whether whether the citizen becomes more or less identifiable in an e-government setting as compared to a paper-based government.

In public service provision, the citizen usually gets authorized access to public services on the basis of form filling, writing letters, and/or submission of official documents which are proofs of entitlement to a certain public service. At the heart of government service provision are the personal identification of the citizen and the verification of that identity through authentication processes. This standardized process of service provision to citizens is known as bureaucracy and was traditionally mainly paper based. Nowadays, however, more and more public services are also provided electronically (Lips, 2006). This raises the questions: ‘To what extent does the application of new technologies in public service provision change the identification process of the citizen by the government?’ and ‘To what extent does the citizen become more anonymous or more known by the government when new technologies are applied in the process of public service provision?’

4.2 Public service provision

In the late eighteenth and early nineteenth centuries, in the aftermath of the economic and political revolutions (i.e., the Industrial Revolution and the French Revolution), nationalism became an ideological instrument in Europe to achieve a sense of identity within a nation. It was the nation state which was represented as a unique and essential unity, the living body of its citizens. (Rietbergen, 1998) A new form of government appeared, the nation state as a political institution, with a unique combination of attributions: a written constitution, a system of law based on this constitution, and a specialized civil service committed to the rules and regulations. (Ultee, Arts, & Flap, 1996) But governability requires controlability. The civil registry, introduced by the French when they occupied large parts of Europe, was in many nations maintained after the Napoleonic Wars. Passports and other documentary controls on the movement and identification have been essential to the states’ development as nation states. (Torpey, 2000) And, as we have seen in Chapter 3, bureaucracy became the key feature in the public provision of services (Weber, 1968/1921).

Public service provision can be divided in two distinct, but related, processes: the front office and the back office. In the front office, the citizen and the public service provider get into contact with each other; this contact can be face-to-face or by mail, phone, or email. In the back office, the forms, letters and emails are handled; decisions are made and then communicated back to the citizen. In the Netherlands, municipalities were traditionally organized around a system of back offices and every public service had its own back office. Over the past decades, ICT applications were implemented in the back office systems to make the processes more effective and efficient. This led to a system of stand-alone, non-integrated back offices, i.e., back offices each had their own citizen registrations which were not connected to the other back offices or a central system.

Since the early nineties, however, central government provided the central citizen registration systems to municipalities which has not only led to connecting the different back offices of a municipality to each other (Schravendeel & Van der Drift 1993), but also to standardized, and therefore more easily accessible, information in databases. In the late nineties, a more client-oriented vision resulted in developing digital alternatives for the traditional paper-based identification and authentication means by applying new technologies in public service provision (Prins, 2001; EGEM, 2006; Lips, 2006). Electronic forms of identification are not replacing the traditional, paper-based, identification but they are an added feature prompted by a more client-oriented vision.

4.3 Case studies

Two case studies have been conducted in the Netherlands in order to evaluate the extent to which the identification process has changed, and the extent to which the citizen has become more anonymous or more identified by the government when new technologies are applied in the process of public service provision. The first case, application of a tree felling permit in the city of Dordrecht, was selected because this is a typical example of a widespread, simple service that is increasingly being delivered by electronic means. Although DigiD, a recently introduced measure for citizens to identify themselves in the realm of online public services is implemented in many municipalities, most municipalities only use DigiD to allow citizens to request information electronically (e.g., Attesta Vita or certified copy of an entry of birth). Dordrecht is a leader in electronic municipal government and one of the very few municipalities that already use DigiD for electronic personal data exchange, as is the case with the application of a tree felling permit. The second case, border passage at Schiphol Airport, was selected for being an exceptional case. Electronic border passage uses a biometric identification, an iris scan, to identify individuals, which is uncommon in public service provision.

In both case studies the traditional identification process was compared with the electronic identification process to see whether the identification processes differ from each other and to see whether a shift in the identifiability-anonymity continuum could be observed. The identification process is understood as all personal information that is asked for, gathered, stored, matched, used, and shared, i.e., the identity knowledge about an individual. The amount of personal information in the identity knowledge was considered and the personal information was classified according to the five different types of identity information, as distinguished by Marx (2005), to establish the nature of the personal information by the distance from the core identity. Finally, the identity knowledge of the provider of the public service was explored. In the first case, the traditional and electronic services are provided by the same organisation, i.e., the city of Dordrecht. In the second case, however, another party is involved in providing the electronic service.[32]

4.3.1 Case study 1: The application for a tree felling permit

In the Netherlands, each municipality has its own regulation for tree felling. In practice, however, the regulations of the different municipalities are very similar to each other and in almost all municipalities one has to apply for a tree felling permit. The application was traditionally paper-based but Dordrecht provides electronic application for a tree felling permit by means of DigiD as well. (www.egem.nl)

[32] At the end of this chapter the detailed findings of the two case studies are presented in a schematic overview.

DigiD is the abbreviation of ‘digital identity’; it is a system used to authenticate individuals in the online environment. It is controlled by *GBO.Overheid*, which is part of the Ministry of Internal Affairs & Kingdom Relations (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties). Together with Inland Revenue (Belastingdienst), *GBO.Overheid* is making DigiD available for the public at large. DigiD can be used in online relationships with many public service providers: municipalities, local governments, Inland Revenue, central student system (Informatie Beheergroep), and land registry office (Kadaster) for example. One can apply for DigiD on the website of DigiD by filling in a social security number,[33] birth date, postal code, house number, and email address. Then the self chosen username and password have to be filled in. Within a few days, an activation code is sent to the home address and after DigiD is activated online it can be used. (www.digid.nl)

The data for this case were collected from several sources: an interview with the project leader of DigiD (and Advisor Information Management) of the city of Dordrecht, information provided by the Dordrecht department that decides applications for a tree felling permit, including the forms used in this process, and Internet sources.

4.3.1.1 Findings

For the traditional, paper-based, application of a tree felling permit, the form can either be obtained at the counter at the city hall or be downloaded from the website of the city of Dordrecht (DigiD is not needed). This paper form has to be filled in and the following personal data are asked for: name (i.e., surname and initials), address, postal code, place of residence, private phone number, mobile or business phone number, place of signature, date of signature, and signature. The paper form has to be sent back by post or can be delivered at the city hall.

To obtain the digital form, the citizen has to authenticate himself or herself by means of DigiD. When the button of the web form is clicked, a redirection to the DigiD website takes place. On this website one has to log in with the DigiD username and password and the user is authenticated. This authentication is digitally sent to the City of Dordrecht who verifies this authentication. After verification, the social security number of the applicant is sent by DigiD to the city of Dordrecht; there it is matched with data from the municipal Personal Records Database (Gemeentelijke Basis Administratie, GBA). This database contains the following personal information: name, address, place of residence, gender, birth date, and social security number. The digital application form, on which the name, initials, address, postal code, and place of residence are already filled in, is sent to the applicant. The form has to be filled in further with the following personal information: private phone number, mobile or business phone number, and email address, and then it has to be electronically submitted.

Both forms will first arrive at the department ‘*Information Process Management*’, where they will be stored. The digital form can be saved directly; the paper form is scanned and then saved. In the ‘*Vaststellingsbesluit selectielijst archiefbescheiden gemeentelijke en intergemeentelijke organen vanaf 1 januari 1996*’ it is decided that a tree felling permit is subject to the general provisions according to archiving as laid down in the ‘*Archiefwet*’ 1995 (Archive Act)[34]. This decision states the minimal term an application has to be retained.

[33] The social security number is to be replaced by the citizen service number (Burger Service Nummer, BSN) in 2007.

[34] All Dutch Acts can be accessed at http://wetten.overheid.nl.

Granted permits have to be retained at least one year; denied, withdrawed, or lapsed permits have to be retained at least three years. Dordrecht, however, extended both terms of retention to ten years, so the information can be consulted at a later date (e.g., when the applicant has been granted the permit but has the obligation to replant or when one suspects that a new application for a tree felling permit has been done for the same location). Every official of the municipality is allowed to consult this information. According to the 'Wet *Openbaarheid van bestuur*' (Freedom of Information Act), every citizen can request information about tree felling permit applications. These citizens have neither access to the actual documents, nor to all personal information, but only the relevant information is communicated. In the case of a tree felling permit this relevant personal information is the location (i.e., address) where the trees are or were located. The number and species of the trees is also communicated, but this is not considered personal information. After being stored at the department '*Information Process Management*', the paper form or a print of the digital form is sent to the department '*Stadswerken*' (City Works), where the application is handled. The personal information is copied and put in the computer where it is retained for a couple of years.

If it is doubtful whether the applicant is the legal owner of the land, and hence legally justified to apply for a tree felling permit, the name, and location (i.e., address) will be matched with the information in the software program '*Flexigis*'. This program contains a Geographic Information System, e.g., information from the land registry office. The private, mobile, and business phone numbers and, when it concerns a digital application, the email address can be used to contact the applicant for more information. After a decision is made, the decision is communicated back to the applicant by mail, using the applicant's name and address. The decision is published in the local newspaper and reported are the address, number and species of the trees, and the motivation of the applicant for felling these trees. This information is also available at the City hall for a period of six weeks, so other citizens can officially object to the decision.

4.3.1.2 Interpretation and conclusion

The main differences between the traditional and electronic identification processes take place in the front office (see table 1 in the appendix). In the traditional identification process, all personal information is directly asked for and no additional information is gathered from other sources. The electronic identification process has a different sequence of the various parts of the procedure. Because information is also gathered from another source, less information is directly asked but the amount of information that is collected (i.e., directly asked and gathered from other sources) roughly equals that of the traditional identification process. In the back office, the procedures for both forms are similar.

In contrast to the traditional identification process, the electronic identification process uses the social security number to authenticate the applicant and match the number with the name, initials, and address, but this number is not stored or used in the dispatch of the application. Sometime in 2007, the social security number will be replaced by the citizen service number, which is the same number but which can be used in more relationships with public and probably private service providers as well (e.g., health organisations, health insurance companies, banks, and land registry office). Although the systems of the different service providers are not linked and DigiD is only used to authenticate individuals, it will be easier to collect information about an individual from various sources when the same unique identification number is used.

Apart from the social security number, the electronic application also requests an email address of the applicant. The email address can reveal little or much information about the applicant. One_two_three@hotmail.com for example does not reveal any information, whereas John.Anderton@PreCrime.org does not only reveal the applicant's first and last name but also the company he works for, at least on the face of it.

On the paper form, also the signature of the applicant is requested to validate the correctness of the personal information (i.e., the applicant is who he claims to be). Place and date of signature are requested too. Place and date are often requested when a signature is asked for and this originates from signing contracts but it is in this context more a habit than a necessity. The signature, although being unique information about an individual, is not used in the dispatch of the application.

According to Marx' five different types of identity information, name, and address are private information, because this information is unknown until communicated. Together, however, they form a relative unique identification of an individual. The phone numbers (i.e., private, mobile and business) are also private information. It is fairly easy to get a name and address attached to the private phone number, but this is more difficult for mobile or business phone numbers.

The traditional and the electronic applications for a tree felling permit are handled by the same organisation (i.e., the city of Dordrecht), so it is irrelevant to address the identity knowledge capacity of the organisation.

In conclusion, there are only minor differences between the traditional and electronic identification processes and these differences mainly occur in the front office. Depending on the information revealed by the email address, the electronic identification process can identify the citizen slightly more. The citizen can thus become more known by using the electronic form when applying for a tree felling permit, but this increase is only marginal, and it is information that the citizen can normally choose himself to share or not, e.g., by choosing an anonymous or non-revealing email address.

4.3.2 Case study 2: Border passage at Schiphol Airport

To establish whether an individual may enter or leave the Netherlands, the individual need to be identified. Identification takes place by passport, identity card, or another official travel document. Residents of the Netherlands can apply for a travel document at *'Afdeling Burgerzaken'* of the municipality where they are registered in the Municipal Personal Records Database (Gemeentelijke Basis Administratie, GBA). Application has to be done in person and all travel documents in possession (Dutch and foreign) and a colour passport photo that is a good likeness and complies with the standards. Anyone with a Dutch nationality can apply for a travel document, but children under 18 wishing to apply for a passport need written permission of both parents (or legal guardian). The same applies to children under 12 wishing to apply for an identity card.

Travel documents are produced and personalized at a central facility in the Netherlands and contain the following personal information: nationality, surname, given names, date of birth, place of birth, height, gender, and personal number (i.e., social security number). The same data is also stored in the machine readable zone of the travel document. The photo and the signature are also on the travel document. Travel documents issued on or after 26 August 2006 also contain a chip. This chip is invisible and it is incorporated in the data page of the travelling document. The chip contains the photograph in colour (the 'facial' image) and all the data that are printed on the data page, except for the signature. A special piece of equipment is required to read the data in the chip. The 'reader' shows the information on the screen but the reader does not store the data it has read. The reader is provided by the Ministry of Interior at various locations in the Netherlands (i.e., 27 municipalities and Schiphol Airport). (www.paspoortinformatie.nl) All data that were colleted in the application process of the travel document and the data about the travel document (i.e., document number, expiry date, and municipality where the travel document is issued) are stored in the '*reisdocumentenadministratie'* (travel document administration) and kept there for eleven years (Paspoortuitvoeringsregeling on wetten.overheid.nl).

To facilitate a faster border passage, Amsterdam Schiphol Airport developed a biometric identification system. The iris is unique for every individual, so identification and authentication of an individual can take place by iris scan. The iris scan equipment is designed by Amsterdam Schiphol Airport and is approved by the Ministry of Justice and the *'Koninklijke Marechaussee'* (KMar) (border police). The iris scan is called Privium and one can apply for a Privium card by filling in the application form. After receiving a letter of confirmation, an appointment at the Privium Service Point for a scan of the iris pattern and to pick up the Privium card has to be made. The appointment at the Privium Service Point consists of three parts: identity papers are checked, the iris patterns of the eyes are scanned and a brief explanation on how to use the Privium card is given. From the iris scan, a maximum of 256 measuring points can be recognized. These measuring points can be used to reproduce the pattern of light and dark of the iris, which is unique for every individual. The scan is only stored on the Privium card and not in a database. When an individual crosses the border, the iris is scanned and the data obtained is compared with the data stored on the card. (www.schiphol.com/privium) The Privium card is not replacing a passport, so a valid travel document is still needed to travel to other countries.

Passport control, however, is not always necessary when crossing a border. To facilitate free movement within an area without internal border controls, several countries signed the Schengen Agreement on the gradual abolition of checks at the common borders in 1985. The Schengen Convention was signed in 1990 and came into effect in 1995. The Schengen Convention abolished the checks at internal borders of the signatory countries and created a single external frontier, where the checks for all signatory countries were to be carried out in accordance with a common set of rules. Countries that are now signed up to the Schengen Convention are: Belgium, Denmark, Germany, Greece, Spain, France, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden, Iceland, and Norway. (www.ec.europa.eu)

The data for this case were collected in several ways. Due to time constraints it was not possible to hold an interview with an employee of Privium. The data from Privium were collected using Internet sources, a Privium user was interviewed, and a phone call was made to Privium to get the remaining questions answered. This user sent an email to Privium to ask which personal data were collected and for which purposes, who have access to this personal data and for which purposes and what data exchange take place and for which purposes. Privium answered these questions by email. The data for traditional border passage were mainly collected by using Internet, but also two telephone calls were made to the *Koninklijke Marechaussee* (KMar) to get some remaining questions answered.

4.3.2.1 Findings

When travelling to or from a non-Schengen country, the travel document is checked by the KMar for genuineness. Then the passport photo is matched with the individual and possibly the height and year of birth, reported on the travel document, are taken into account to establish whether this person is the same as the person on the photo. In addition, some personal information (most likely this contains the name, date of birth, and place of birth because this is the same information that is stored on the Privium card, but the KMar refused to answer questions about this) is entered in the computer, when the passport does not contain a chip, or the data on the chip is read by the 'reader', when the passport is equipped with a chip. The obtained data is compared with data in the Schengen Information System (SIS). The SIS was set up to allow police forces and consular agents from the Schengen countries to access data on specific individuals (e.g., persons wanted for arrest for extradition purposes, aliens for whom an alert has been issued for the purpose of refusing entry, missing persons or persons needing temporary police protection, witnesses and persons summoned to appear before judicial authorities, and persons submitted to discreet surveillance or specific checks for the purpose of prosecuting criminal offences or for the prevention of threats to public services) and on goods which have been lost or stolen. These data are supplied by all participating countries[35] via national sections (N-SIS) that are connected to a central function (C-SIS). The SIS is supplemented by a network known as Sirene (supplementary information request at the national entry) which allows communication between the Sirene offices in every member country. (Benyon, 1994; www.ec.europa.eu; www.politie.nl) The Dutch are, when leaving the Schengen area, also checked for unpaid fines or unserved sentences by matching name, date of birth, and personal number with data in the '*Recherchebasissysteem' (RBS)* (database of the detective force which contains all criminal offences). (www.mpbundels.mindef.nl)

The data about the border passages of individuals to a non-Schengen country are, according to the explanation of the Act 'Wijziging Paspoortuitvoeringsregeling 2001',[36] stored in a database that is managed by the KMar. The KMar, however, denies that it stores information about border passages of individuals. Which data are stored is unclear, but most likely these include name, date of birth, place of birth, and travel date. To perform their duties and to fight international terrorism, this database can be accessed by the *'Algemene Inlichtingen- en Veiligheidsdienst'* (General Intelligence and Security Service) and the '*Militaire Inlichtingen- en Veiligheidsdienst'* (Military Intelligence and Security Service) and they can match this data with the personal data from the database '*reisdocumenten-administratie'*. (Wijziging Paspoortuitvoeringsregeling on wetten.overheid.nl)

When applying for a Privium card, one has to fill in an online form at the Schiphol/Privium website. The personal data have to be filled in exactly according to the passport and they are: surname, all initials, date of birth, place of birth, nationality, type of travel document (i.e., passport or European identity card), travel document number, expiry date of the travel document, home address (i.e., street name, house number, postal code, city, and country), and

[35] Although not all EU member States have signed up to the Schengen Convention, they all participate in the SIS. Participating countries are, thus, the EU-25, Norway, and Iceland. Switzerland wants to participate in the Schengen acquis too and has started negotiations in 2002.

[36] In this explanation it is stated that the *'Algemene Inlichtingen- en Veiligheidsdienst'* (General Intelligence and Security Service) and the *'Militaire Inlichtingen- en Veiligheidsdienst'* (Military Intelligence and Security Service) gain access to information about border passages of individuals from the KMar.

email address. Private and business phone numbers are optional to fill in. A different invoice address (e.g., of the company) is also optional and when left blank the home address will be used. A payment method has to be chosen and when a direct debit from a bank account is chosen, the Dutch bank account number is also requested. These data are stored in the Privium database and some of these data are shared with the KMar, but which data is unclear.

After receiving a letter of confirmation, an appointment at the Privium Service Point for a scan of the iris pattern and to pick up the Privium card has to be made. Before a scan is made, the travel document is checked by the KMar for genuineness, after which the passport photo is matched with the individual and possibly the height and year of birth, reported on the travel document, are taken into account to establish whether this person is the same as the person on the photo. Then, the irises of both eyes are scanned and the template of the iris scan is stored on the chip of the Privium card along with the Privium card number, name, date of birth, and place of birth. The iris scan is only stored on the card and nowhere else.

When a Privium member travels to or from a Schengen country, the Privium card can be used to gain fast access into the clean area. The iris scan itself is not used, because the identity of the traveler does not need to be checked when travelling in the Schengen area. The Privium Card has to be inserted in the reader and the gate will open. In any case the Privium Service is used, not only the date and time at which the card was inserted will be processed, but also the name, date of birth, and place of birth of the user. The KMar has access to these data at all times. In contrast to the traditional border passage, the ticket is not checked as a matter of course, but the KMar and the Security Staff may check tickets at random.

Using Privium and travelling to or from a non-Schengen country, however, requires an iris scan. The card has to be inserted in the reader and the Privium member has to look in the scanner so an iris scan can be made. At least four, simple, digital, black and white photographs are made of the eye in less than a second. A code is calculated on the basis of one of the photographs and this code is then compared with the template of the iris which is stored on the Privium Card. If the iris recognition fails, the passport will be checked by the KMar. The other data on the card (i.e., name, date of birth, and place of birth) are automatically matched with the SIS and the RBS to see whether the person is allowed to leave or enter the Netherlands. When there is a match between the data on the card and the data in the SIS or RBS, the gate will not open and a passport check by the KMar will be needed. Although Privium facilitates automatic border passage without manual checks on passports and tickets, the KMar still carries out random checks on passports and tickets of Privium members.

The use of the Privium card (i.e., date and time) as well as user data (i.e., name, date of birth, and place of birth) is stored in the Privium database. The Privium database is accessible at all times by the KMar, because the KMar needs information about border passages of persons for safety reasons. (www.schiphol.com/privium)

4.3.2.2 Interpretation and conclusion

The main difference between the traditional and electronic identification processes is that the personal information of a Privium user is stored in yet another database (i.e., the Privium database) and that the data stored also include address and email address. In addition, data storage also occurs when the Privium user travels to a Schengen country and these data are accessible by the KMar, whereas with traditional border passage personal data are not stored (see table 2 in the appendix).

When travelling to a non-Schengen country, the travelling document or Privium card is asked for. Although the travelling document contains more personal information than the Privium card (see table 3 in the appendix), the KMar can easily access databases where the information on the travel document is stored (i.e., Privium database or Reisdocumenten-administratie). Because the actual identification of the individual takes place manually in the traditional border passage, the identification is less reliable than the automatic identification of comparing the template of the iris with the iris scan of the individual. Not only is there a higher chance of making mistakes, but also the passport photo is a less unique identifier than the iris template. People can change their hair style or glasses and thus look different than on the photo. In addition, identical twins can use each others passport. It is, however, impossible to change your irises and the irises are unique for every individual. This holds for identical twins too and even both irises of one individual differ from each other.

Name, date of birth, and place of birth are used to identify and match individuals. According to the five different types of identity information, this can be considered as unique information about an individual, because there is only a very small chance that there is another person with exactly the same name (i.e., surname and first names) born on the same date in the same town. However there is still a chance. From an ICT point of view it could be argued, that it would be more logical to use the personal number (social security number), which is also in the travel documents, to identify individuals. This is not permitted by law, however. A passport photo is unique information about an individual too, but a template of the iris is able to identify an individual more accurately.

The address and email address, both types of private information, are additional information, requested by Privium. These are only used to contact the Privium member and not in the actual identification process. However, the address and email address are stored in the Privium database and therefore accessible by the KMar. As we have seen in the first case, the email address can reveal little or much information, but the individual can choose what kind of email address he wants to use if he is aware of the storage of his data in the Privium database.

In conclusion, using the electronic border passage identifies an individual to a greater extent than is the case in traditional border passage. Using Privium means not only that the personal data are stored in an extra database, but also that more information is asked for, stored, and shared.

The identity knowledge capacity (i.e., the size of the files held, the centralization of those files, the speed of information flows, and the number of points of contact) (Rule, 1973) of the KMar increases with the use of the Privium card. The KMar has access to more personal information (i.e., address and email address from the Privium database) about a Privium user than a non-Privium user, so the size of the files increases. In addition, the KMar also has information when a Privium member is travelling to a Schengen country (i.e., name, date of birth, place of birth, date of passage, and time of passage), this in contrast to traditional border passage, so the number of points in the life of a person that are available for collecting information increases too.

4.4 Conclusion and discussion

The two questions of this chapter were ‘To what extent does the application of new technologies in public service provision change the identification process of the citizen by the government?’ and ‘To what extent does the citizen become more anonymous or more known by the government?' From the results of both case studies, it can be concluded that new technologies change the identification process. In the tree felling permit case, the sequence of the various parts of the identification process was changed. In the border passage case, personal information was stored in an additional database which could also be accessed by the KMar.

As to the identifiability of citizens, the cases produce mixed results. In the tree felling permit case, the citizen was equally identifiable when the identification process of the traditional application was compared with the identification process of the electronic application. So, there was no shift detected on the anonymity-identifiability continuum. In the border crossing case, however, a shift was detected on the anonymity-identifiability continuum towards identifiability. Using the electronic Privium border passage means that the personal data is stored in an extra database and more information is asked for, stored, and shared. In addition, using the electronic application (i.e., Privium) also enhances the identity knowledge capacity of the service provider (i.e., KMar).

Of course, no hard conclusions can be drawn from only two cases. To get more comprehensive answers, more cases should be studied, but this chapter can direct to future research. The identification process applied in the border passage application remains opaque and not transparent. For security reasons, information about the identification process is not easily given. The electronic identification process in the application for a tree felling permit is modelled after the traditional identification process and electronic application for public services is relatively new. So, the electronic identification process when applying for public services may change in time. When a similar study is conducted within a few years, the results can be different.

4.5 Bibliography

Benyon, J. (1994). Policing the European Union: The Changing Basis of Cooperation on Law Enforcement. *International Affairs, 70*, 497-517.

EGEM (2006). *Architectuur Model Gemeentelijke E-Dienstverlening. Project referentiemodel Midoffice ten behoeve van gemeenten.* Available at http://www.egem.nl/projecten/voorhoede gemeenten/kennisconferentie2006/documenten?searchterm=kennisconferentie%202006.

Lips, M. (2006). E-Government Under Construction: Challenging Traditional Conceptions of Citizenship. In V. Koutrakou & P. Nixon (eds.), *E-Government in Europe: Re-booting the State.* London: Routledge.

Marx, G.T. (2005). Varieties of Personal Information as Influences on Attitudes towards Surveillance. In K. D. Haggerty & R. V. Ericson (eds.), *The New Politics of Surveillance and Visibility.* Toronto: University of Toronto Press.

Prins, J.E.J. (2001). E-overheid: Evolutie of revolutie? *Nederlandsch juristenblad 76*, 515-520.

Rietbergen, P. (1998). *Europe: A cultural history.* London: Routledge.

Schravendeel, D., & Van der Drift, J. (1993). Elektronische gegevensuitwisseling wint terrein binnen de overheid. *B&G, 20*, 17-20.

Torpey, J. (2000). *The Invention of the Passport. Surveillance, Citizenship and the State.* Cambridge: Cambridge University Press.

Ultee, W., Arts, W., & Flap, H. (1996). *Sociologie: vragen, uitspraken, bevindingen*. Groningen: Wolters-Noordhoff.

Weber, M. (1968/1921). *Economy and Society. An outline of interpretive sociology* (translated by Fischhoff, E. et al.). New York: Bedminster Press.

Internet sources

www.digid.nl

www.ec.europa.eu (http://ec.europa.eu/justice_home/fsj/freetravel/frontiers/fsj_freetravel_schengen_en.htm)

www.mpbundels.mindef.nl

www.paspoortinformatie.nl

www.politie.nl

www.schiphol.com/privium

www.wetten.overheid.nl

4.6 Appendix

Table 1: Identification processes, Applying for a tree felling permit

	Traditional	Electronic
Front office		
Asked for	Name	Social security number
	Address	
	Private phone number	
	Mobile/business phone number	
	Place of signature	
	Date of signature	
	Signature	
Matched with Personal		Social security number
Records Database		
Gathered		Name
		Address
Asked for		Private phone number
		Mobile/business phone number
		Email address
Back office		
Stored	Name	Name
	Address	Address
	Private phone number	Private phone number
	Mobile/business phone number	Mobile/business phone number
	Place of signature	Email address
	Date of signature	
	Signature	
Matched with database	*Name*	*Name*
Land registry office	*Address*	*Address*
Used		
Communication	Name	Name
	Address	Address
Legally justified	*Name*	*Name*
	Address	*Address*
Contact	*Private phone number*	*Private phone number*
	Mobile/business phone number	*Mobile/business phone number*
		Email address

Personal information in italics is occasionally matched or used.

Table 2: Identification processes, Border passage at Schiphol Airport (part 1)

	Traditional	Electronic	
	KMAR	KMAR	Privium
Before border passage			
Asked for			Name Date of birth Place of birth Nationality Type of travel document Travel document number Expiry date travel document Address Email address Iris scan
Matched with Individual		Passport photo Year of birth Height	
Used			
Card			Name Date of birth Place of birth Template of the iris
Contact			Name Address

Table 3: Identification processes, Border passage at Schiphol Airport (part 2)

	Traditional	Electronic	
	KMAR	KMAR	Privium
Border passage to Schengen country			
Stored			Name Date of birth Place of birth Time and date of passage
Shared with KMAR			Name Date of birth Place of birth Time and date of passage

Border passage to non-Schengen country			
Asked for	Travel document which includes:	Privium card which includes:	
	Name	Name	
	Nationality	Date of birth	
	Date of birth	Place of birth	
	Place of birth	Template of the iris	
	Gender		
	Height		
	Personal number		
	Passport photo		
Matched	Passport photo	Template of the iris	
with	Height		
individual	Year of birth		
with	Name	Name	
SIS and RBS	Date of birth	Date of birth	
	Place of birth	Place of birth	
Stored	(Name)		Name
	(Date of birth)		Date of birth
	(Place of birth)		Place of birth
	(Date of passage)		Time of passage
			Date of passage
Used			
for	Passport photo	Template of the iris	
identification	Height		
	Year of birth		
for	Name	Name	
tracing	Date of birth	Date of birth	
	Place of birth	Place of birth	
Shared			
with AIVD	(Name)	(Name)	
and MIVD	(Date of birth)	(Date of birth)	
	(Place of birth)	(Place of birth)	
with			Name
KMAR			Date of birth
			Place of birth
			Date of passage
			Time of passage

For the personal information between brackets, it is unclear whether this information is stored.

5 Electronic signatures and pseudonymous certificates

5.1 Introduction

In the context of governmental use of signatures identification of the signer of a document traditionally plays an important role. When transiting from paper based to electronic signatures an implementation focused on identifiability may develop spillovers. Based on the German Electronic Signature Law potential spillovers are analysed and methods to avoid them by making effective use of pseudonymous electronic signatures are discussed.

5.2 Paper-based and electronic signatures

Today, documents signed traditionally on paper are stored in various places within public administrations and enterprises. Getting access to these documents and linking various corresponding transactions to a certain user needs a considerable effort and a lot of resources, as support by computers (e.g., using indices or registers) for analysing these paper documents is limited in most cases.

With a transition from a traditional to an electronic signature, this situation potentially changes for several reasons. To reduce costs for processing and storage of digital data, centralised Document Management Systems (DMS) will be used increasingly as a platform for various governmental procedures.[37] If the access control of these centralised DMS is insufficient, data may be read out and analysed by unauthorised parties. In addition, signed documents are frequently transferred via the Internet. If technical security (e.g., encryption on the transport of data via the Internet) or organisational security (e.g., in clearing houses or gateways[38]) breaks, communication can be eavesdropped and analysed.

In contrast to signed paper documents that can be analysed manually only or with limited computer support, unencrypted digital signatures can be analysed in an automated way. In most cases the certificate issued by the certificate authority issuing the electronic signature is the key for this analysis. For electronic signatures mostly X.509 V.3[39] compatible certificates are used (Meints, Hansen 2006). In addition to a unique certificate number that allows for easy linkability of cases in which electronic signature were applied, they store additional information of the legitimate user of the key(s) for the electronic signing processes. This information may include, depending on the rules of the different certificate authorities in different countries:

- First name, surname and in some countries title of the legitimate user
- Date of birth
- Address

[37] For example by the Federal State of Hessen in Germany where for the public administration a centralised DMS has been built up, see http://www.hessen-egovernment.de/irj/eGovernment_Internet?cid=fb54122426a45c32deeea574ebfd5171

[38] Implementation of an end-to-end encryption of electronically signed document, signatures and certificates seems to be difficult in cases where e.g., in a European context for compatibility reasons transformation of data in different implementations of certificates is necessary. See in this context e.g., the GUIDE project and the gateway concept developed there, see http://istrg.som.surrey.ac.uk/projects/guide/

[39] ETSI TS 102 280 V1.1.1 (2004-03), X.509 V.3 Certificate Profile for Certificates Issued to Natural Persons, available after registration from http://www.etsi.org/services_products/freestandard/home.htm.

- Citizens register number

It is safe to say that some of this information is usually not needed in cases where an electronic signature is applied. For example in most cases where a contract to buy something is signed, date of birth and citizens register number are not needed. The standardised disclosure of this information does not meet the data minimisation principle stated in the European Data Protection Directive 95/46/EC, but as the use of personal data for electronic signatures is regulated in the European Signature Directive 1999/93/EC and corresponding national laws, these certificates are clearly legal (Gasson, Meints, Warwick 2005: 25).

Traditional paper based signatures are used in governmental services in different contexts. Not all of these contexts require an ad hoc identification of the signing person, as information transferred in a signed document may belong to different circles of information (cf., section 3.2) and may have different relevancy in the context of governmental processes. Nevertheless aiming at fast and reliable identification of citizens in Germany requirements for electronic signatures are highly standardised for eGovernmental services.

5.3 Approaches for privacy enhancement

To improve the implementation of data protection principles in the context of electronic signatures, two approaches are discussed aiming at reduction of linkability of certificates:

1. Use of context-specific digital credentials instead of one "general purpose" certificate. This could be done e.g., by issuing a specific signed pseudonymous certificate or digital credential in general that includes only necessary information in the context of the signature and that uses a number that is not repeated for other certificates and thus is not linkable. In this context the certificate authority also takes over the role as identity provider, acting as a trusted third party.
2. Use of pseudonymous signatures. In this concept, which is also supported by the European Signature Directive 1999/93/EC (Gasson, Meints, Warwick 2005: 33), for one physical person, different key pairs and correspondingly different pseudonymous certificates can be used. This approach limits linkability to those signatures in which the same pseudonymous certificate has been used.

In both cases, it is possible to uncover the physical person behind the pseudonymous certificate and credential by the certificate authority when this is needed and legal (e.g., in cases of criminal investigations). There are other approaches where the physical person behind a pseudonym is not identified, but where still legitimate claims can be covered and/or misuse can be prevented (Pfitzmann, Hansen 2006).

5.4 Limitations of the described approaches

For the use of digital credentials, various technical concepts such as "Idemix"[40] and "Credentica"[41] were described and analysed in recent FIDIS deliverables, see Bauer, Meints and Hansen (2005) and Gasson, Meints and Warwick (2005). So far, digital credentials have not been used together with PKI and electronic signatures; the use of sector-specific PINs in Austria is limited to authentication and is not being used for electronic signatures. To implement this approach, two things currently not present would be needed:

[40] See http://www.zurich.ibm.com/security/idemix/

[41] See http://www.credentica.com/

1. a legal basis in the corresponding countries, and
2. technical prototypes and implementations.

Technical approaches have been discussed in the eforum in the context of the PPP project;[42] scientific publications are in preparation but not available yet.

5.5 Pseudonymous signatures in e-government

Though pseudonymous signatures are supported by the European Electronic Signature Directive, only a few European countries have included pseudonymous electronic signatures in national electronic signature legislation so far, among them Germany. Hornung has analysed the possibilities to use pseudonymous electronic signatures in the context of public services in Germany (published in Roßnagel 2006).

In Germany, pseudonymous certificates for electronic signatures are a standard service of CAs.[43] They are implemented in accordance with the law on the national ID card for using a pseudonymous name instead of the original first name and surname. In some special cases, e.g., chambers or associations for certain professions (medical doctors, etc.), the application of pseudonymous signatures together with the professional attributes might not be possible.These, however, are exceptional cases that are well understood (Roßnagel 2006: 59).

Contrary to the statement in the German Electronic Signature Law, the German government understands pseudonymous signatures as a purely voluntary service of CAs. This understanding results in the conclusion that no German citizen is entitled to make pseudonymous electronic signatures. This understanding and the corresponding implementation of the law limits the possibility of citizens to even get a pseudonymous electronic certificate issued by a CA in Germany (Roßnagel 2006: 58f).

In addition, the use of pseudonymous signatures in Germany for signing governmental tranactions is limited by law. Art. 3a of the German Law on governmental procedures (“Verwaltungsverfahrensgesetz”, VwVfG) states that “The use of electronic signatures using a pseudonym that does not allow the identification of the person using the electronic signature is prohibited”. In additional comments to the law, the German Federal Government states that pseudonyms may be used by the public administration (such as the municipal offices for social affairs), but not by the citizen, in order to avoid “abuse of governmental services […]” (Roßnagel 2006: 60).

Hornung argues that this is not proportionate, as there are governmental procedures that do not require identification. In addition, an identification of the user of a pseudonym in defined cases by the CA is possible (Roßnagel 2006: 60). On the other hand, the certificates issued together with key pairs for electronic signing in Germany are not meant and developed to identify the user of the key pair without doubt. In fact identification currently is possible via additional information issued by the CA only, as relevant information such as date of birth, location of residence etc. is not stored in the certificates (Roßnagel 2006: 61). Hornung concludes that from this perspective, the exclusion of pseudonymous signatures for governmental procedures does not make sense at all and the corresponding articles of the German Law on governmental procedures should be changed (Roßnagel 2006: 69).

[42] See http://www.eu-forum.org/article.php3?id_article=258

[43] Art. 5 par. 3 German Electronic Signature Law states that a CA shall support a pseudonymous electronic signature ‘on request’.

5.6 Conclusion

When transiting from paper documents and traditional signatures to digital documents and electronic signatures, linkability and the control thereof become an important issue. In contrast to paper-based documents, when there is a security leak, potentially a multitude of transactions in which an electronic signature was used can be observed and tracked by unauthorised parties, depending on the technical infrastructure and the security concept used. As the certificates for electronic signatures can be used as identifiers, linking different transactions can be done automatically and very fastly. In contrast, getting access to and scanning, analysing, and linking various paper documents requires much more effort.

So far, two approaches have been developed to prevent or at least to limit linkability: (a) the use of context-specific digital credentials and (b) the use of pseudonymous electronic signatures. Today, both approaches fall short in the context of eGovernmental services, as legal grounds and the technical infrastructure currently are not prepared for effective implementation.

The result is that the potential identity knowledge of governments is increasing in the transition to an electronic authentication environment. The government can link signed electronic documents much more easily than it could in a paper environment, and this easy linking of multiple signed electronic documents allows the creation of new knowledge, in the same way that data mining of digitised databases allows the discovery of knowledge that was hitherto hidden in practically unlinkable paper-based databases.

It is a missed opportunity that the German government has not fostered the privacy-enhancing use of pseudonymous signatures in eGovernment services in order to keep stable its level of identity knowledge. Moreover, the fact that the prohibition of pseudonymous certificates in the eGovernment context seems to hamper German citizens to acquire a pseudonymous certificate for non-public use implies that also the identity knowledge of private parties may be enhanced, which is contrary to the spirit of the Electronic Signature Directive.

5.7 Bibliography

Bauer, M., Meints M., Hansen, M., FIDIS Deliverable 3.1 – Structured Overview on Prototypes and Concepts of Identity Management Systems, Frankfurt a. M. 2005. See http://www.fidis.net/487.0.html

Gasson, M., Meints, M., Warwick, K. (Eds.), FIDIS Deliverable 3.2 – Study on PKI and Biometrics, Frankfurt a. M. 2005. See http://www.fidis.net/487.0.html

Meints, M., Hansen, M. (Eds.), FIDIS Deliverable 3.6 – Study on ID Documents, Frankfurt a. M. 2006. See http://www.fidis.net/487.0.html

Pfitzmann, A., Hansen, M., Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology, v0.28, May 29, 2006, http://dud.inf.tu-dresden.de/Anon_Terminology.shtml.

Roßnagel, A. (Ed.), Allgegenwärtige Identifizierung? Neue Identitätsinfrastrukturen und ihre rechtliche Gestaltung, pp. 53-69, Nomos-Verlag, Baden-Baden 2006.

6 Identification in eGovernment: the Belgian federal case

6.1 Introduction

On the Belgian federal level, eGovernment is understood as the 'continuous optimization of service delivery and governance by transforming internal and external relationships through technology, internet and new media'.[44] The added value of eGovernment lays in the belief that the usage of new technologies should not be limited to the mere transition from paper-based to electronic procedures. The electronic service delivery is thus accompanied by a thorough transformation of the back-offices.[45]

The main idea behind this transformation of back-offices is that the services should be delivered to the client in an integrated way, regardless of the administrative boundaries of the distinct government entities, through one or more front offices and oriented around 'life events' for citizens (e.g., getting married) and 'business episodes' for enterprises. This allows government entities to remain focused on their needs instead of having to deal with the specific functional organization of the public sector.[46]

From the perspective of the government's clients, the main benefits of eGovernment lay in the optimization of the service delivery in such a way that it becomes faster, more user-friendly, less-contact intensive and more transparent.[47] From a government service provider perspective, the main benefits are increased efficiency (reduced costs, more and faster services) and effectiveness (increased quality and type of services).[48]

In this chapter, we focus on a particular issue that arises in the context of eGovernment: the identity knowledge that the government has of citizens placed in the context of eGovernment in Belgium. Is this identity knowledge rising through eGovernment development?

6.2 Some important distinctions: opacity and transparency tools

To answer this question, it is good to recall the distinction we made in FIDIS deliverable D7.4 between *privacy* as an *opacity tool* and *data protection* as a *transparency tool*.

In a nutshell, in D7.4 Hildebrandt and others explained that *transparency tools* are directed towards the control and channeling of legitimate uses of power, and compel government and private actors to 'good practices' by focusing on the transparency of governmental or private decision-making and action, which is the primary condition for an accountable and responsible form of governance.[49] Data protection regulation is such a transparency tool with

[44] [Deprest and Robben 2003], p. 6; [Robben 2004], slide 2.

[45] The term "back-office" refers to the idea that government services are delivered in two phases. First the intake of the basic data for the service delivery, and a first part of the service delivery by the front-office (e.g., a government website), and then the completion of the service by the back-office. [DE BOT 2005], p. 6.
The back-office evaluates whether the *client* is entitled to get the service or not, verifies the data received from the client at other entities, and/or communicates them to other entities. As explained supra in section 3.1, the notion a citizen as a client of the 'business of government' was introduced a few decades ago. It is strongly present in current Belgian eGovernment, where citizens and enterprises are being treated as government's clients (see for example the used terminology on the federal portal www.belgium.be (in the section "over eGovernment"). This is also why other business terms, such as 'business models' are not uncommon in Belgian eGovernment.

[46] The Belgian Federal government thereby follows the recommendation by IDBC. See [IDA Interoperability 2004], p. 16 ff.

[47] See the topic 'Wat zal veranderen voor burgers en ondernemingen?' in the page 'about eGovernment' on the federal portal http://www.belgium.be.

[48] [ROBBEN 2006c], slide 3-4 and [Strickx 2004], p. 4.

[49] [HILDEBRANDT 2005b], p. 16 ff.

regard to one of the most important governmental assets: (personal) information. It sets, *inter alia*, the rules with regard to the legitimate processing of that information, creates specialized and independent bodies to control and check the actions of government, and so on.

A different set of tools are *opacity tools*, such as privacy, which protects natural persons,[50] their liberty and autonomy against state interference and also of interference of other (private) actors. They are essentially linked to the recognition of human rights and the sphere of individual autonomy and self-determination. Both doctrine and research projects that work with the concept of privacy often understand informational privacy as: 'The freedom of a natural person to sustain a "personal space", free from interference by other entities' or as 'the right of a natural person to decide for itself when and on what terms its attributes should be revealed.'[51]

In sum, tools of opacity set limits to the interference of power in relation to the individuals' autonomy and thus with the freedom to build identity and self. It can also be said that opacity tools imply the possibility and protection of the anonymity of individuals and their actions.[52] Keeping identity data out of the public sphere can, for instance, lead to avoiding undesired identification and/or profiling.

In practice, privacy as an opacity tool can be attained via identity-obfuscating techniques, based on which (some or all) user's characteristics or traces are hidden (made opaque), and thus make undesired identification more difficult. Some of the techniques used for that are anonymity, unlinkability, unobservability, pseudonymity, and encryption.[53] Various European projects research identity management system. Some of these, like the EU PRIME project,[54] develop privacy-enhanced IDM systems, where the user is empowered to decide on the release of data and on the degree of linkage to his or her personal data within the boundaries of legal regulations. A privacy-enhanced application design would then support both 'user-controlled data release' as well as 'user-controlled data linkage'.[55] This type of IDM systems, also known as FIDIS type 3 IDM systems (see FIDIS deliverable D3.1) is certainly privacy-enhancing, but it is not being implemented in Belgian eGovernment.

[50] Legal persons have different needs and wishes than natural persons. For commercial reasons, they usually have an interest in being identified and known as much as possible. As a general rule, privacy and data protection rules do not apply to them (several exceptions exist).

[51] We have found the following definitions of (informational) privacy in different Identity Management projects. In the framework of the IDEM project, we have chosen for the MODINIS IDM terminology, which is consistent with all the other definitions.

- [Westin 1970] 'Privacy is the claim of individuals, groups and institutions to determine for themselves when, how and to what extent information about them is communicated to others.'
- [NABETH 2005a] (internal portal): 'Privacy is the ability of a person to control the availability of information about and exposure of him- or herself. It is related to being able to function in society anonymously (including pseudonymous or blind credential identification).'
- [BUCHTA 2005]: 'Self-determination of what information is known about a person and how it is used'. [p. 28]
- [MODINIS IDM 2005] 'Privacy is the right of an entity – in this context usually a natural person – to decide for itself when and on what terms its attributes should be revealed. Privacy is also the freedom of a natural person to sustain a 'personal space', free from interference by other entities'. [p. 14]
- [DUMORTIER 2003c] 'the interest that individuals have in sustaining a personal space, free from interference by other people and organizations' [p. 10]
- [CLARKE 1999] 'Privacy is the interest that individuals have in sustaining a 'personal space', free from interference by other people and organisations. Information Privacy is the interest an individual has in controlling, or at least significantly influencing, the handling of data about themselves.'

[52] [HILDEBRANDT 2005b], p. 15.

[53] [NABETH 2005a], p. 26.

[54] More information on this research project can be found at https://www.prime-project.eu.

[55] See also [PFITZMANN 2006] , p. 23, footnote 68. The term 'data linkage' is further explained below.

Notwithstanding privacy's core importance, it is clear that privacy is a relatively weak fundamental right: not a single aspect of privacy takes absolute precedence over other rights and interests and never does an individual have absolute control over an aspect of his/her privacy. [56]

Privacy can be restricted when balanced against other interests (rights of others, law enforcement, public health, etc.) and under a number of conditions (such as the legality of the restriction).[57] The interference should correspond to a pressing social need and take into account the principle of proportionality.[58] Tasks government entities carry out for the public interest justify to some extent limitations of the right to privacy and exceptions of the general data protection rules.

It is self-evident that these exceptions and limitations also affect the privacy components of the identity management architecture used in eGovernment. Concretely, this means that a privacy-enhanced identity management architecture based on user control and pseudonyms – such as the one developed by PRIME – will not always be a realistic option, especially not in those contexts where privacy coexists with a number of strong other interests and exceptions, such as in eGovernment.

A relevant question that arises in this context is: when a government – such as the Belgian federal government – chooses for a Benthamite approach to identification,[59] and identity knowledge thus indeed increases, can some degree of privacy still be provided to the government’s clients?

In our point of view, this is the case. We come back to this later. First, we explain why we believe Belgian federal eGovernment indeed allows an increased identity knowledge.

6.3 Information is a strategic resource

One of the most important governmental assets is (personal) information. The vision of Mr. Deprest and Mr. Robben[60] on how information should be dealt with in Belgian federal eGovernment can be summarized as follows:[61]

- Information should be modelled in a flexible way that maximally takes into accounts the users’ needs.

[56] This is nicely illustrated by the fact that the European Court of Human Rights recognizes different sorts of human rights. It recognizes some so called 'hard core' or absolute rights that must be respected even in times of emergency when derogations to other rights are justified (article 15 § 2 European Convention on Human Rights (ECHR)). Next to this there are 'normal rights' (e.g., article 5 and 6 ECHR) which can be derogated from in times of emergency (article 15 § 1). Finally, the ECHR contains four rights which can be legitimately restricted in terms of emergency but also under some specified conditions (article 8-11 ECHR, the conditions for permissible restrictions are listed in the second paragraphs of these Articles). Privacy is one of these 'restrictable rights'. See [HILDEBRANDT 2005b], p. 19.

[57] According to the above mentioned art. 8 of the European Convention on Human Rights, exceptions to this right should be: legally embedded and necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others. The European Court of Human Rights has ruled that the law allowing such interference should be accessible, sufficiently clear, detailed, and foreseeable in order to avoid any abuse from public authorities. This rule clearly points in the direction of a concern for legal certainty and transparency.

[58] See also [BUCHTA 2005] p. 5.

[59] Cf. chapter 2 in this deliverable with an historical-philosophical background.

[60] The former is Executive Director of the Federal Public Service for ICT and the latter is Executive director of the Crossroads bank for Social Security, member of the Privacy Commission and CEO of SMALLS-MVM, a large non-profit organization that provides IT services to the Belgian Federal government.

[61] [Deprest and Robben 2003], 7-10 and 50-53.

- Information should be managed efficiently during its whole life-cycle. This means inter alia that information should be collected only once, and maximally reused. In addition, a functional task division should be agreed on, to know which government entity stores which data in authentic form. As a result, an authentic source is determined for each data within the government.[62]
- Information should be exchanged electronically where possible, based on a functional and technical interoperability framework and based on the usage of common identification keys for all relevant entities.[63]
- Information should be adequately protected, and managed in accordance with data protection regulation.

It would lead us too far to describe all these aspects here. We will limit us to the description of the last part of the third bullet point (common identification keys), which is relevant for the current case study.

First of all, a strictly regulated globally unique identifier appears to be assigned to all Belgian citizens at their first registration in the National Registry (at birth). Second, a number of identity data is being held by the National Registry. This data could also be used to uniquely identify Belgian citizens. Third, there are a number of identifiers that apply to the other entities that are relevant for the government (rights, duties). Finally, there is the roll out of the Belgian electronic identity card, which is, inter alia, being used for global identification and authentication of the holder towards the government and other entities.

We briefly discuss these topics in sections 6.4 and 6.5.

6.4 Identification in Belgian federal eGovernment

6.4.1 Identification

The identification of an entity is the process of identifying a user or a provider in a certain context or sector.[64] An entity is directly or indirectly identified or identifiable by reference to identity data, i.e., by reference to one or more identifiers and/or one or more attributes specific to the entity's physical, physiological, mental, economic, cultural or social identity.

6.4.2 Identity data

These attributes specific to the entity's identity are characteristics of that entity. In a number of recent opinions on the processing of personal data for user management in identity management,[65] the Belgian Privacy Commission explained that a (natural) person can be uniquely identified without any margin of error via the combination of:

- the identification number of the National Registry, which is a unique identifier, and

[62] This term is characterized by the (legally embedded) obligation to ask information to other administrations if the data has been collected by another administration before, and the right of the customer to refuse subsequent data collection. The concepts will be further elaborated below.

[63] [Deprest and Robben 2003], p. 20 ff.

[64] Based on [ITU-T 2005], p. 21 and [MODINIS IDM 2006]

[65] [Decision Privacy Commission 6 July 2005], [Decision Privacy Commission 24 May 2006b], [Decision Privacy Commission 6 September 2006].

- his or her family name and first names, date and place of birth, the address of the main residence.[66]

These decisions give guidance with regard to which identity data are needed to globally identify a natural person in Belgian eGovernment.

In their position paper on Belgian Federal eGovernment, Mr. Deprest and Mr. Robben explain that, besides the mentioned data, the gender, the nationality and the date and place of death are also part of the basic identification data of a natural person.[67] With regard to enterprises, they write that in Belgian Federal eGovernment, the global identification number for enterprises is linked to a set of basic identification data, i.e., the company name, address of its head office and plants, its legal status and legal situation. [68]

When we analyze some recent decisions of the Privacy Commission on the topic of identity management, it appears that the Privacy Commission systematically grants access to the above mentioned basic identification data, including the National Registry Number, when the parties involved claim this is needed for their user management.

For instance, in one of the cases, the Privacy Commission accepted the reasoning of the City of Antwerp, which argued that it shall use the National Registry number for its user management, 'because the unique identification of the users is needed and because authorization relies on authentication and authentication relies on unique identification' (our translation).[69]

The Privacy Commission accepted a similar reasoning in her decision about the user management of FEDICT (the Belgian federal public service for ICT). She accepted the point of view that FEDICT shall be authorized to use the National Registry number, for its user management which consists of:

- Identification: attributing a unique set of data that allows knowing who a person is.
- Authentication: verification whether what is claimed to be, also is correct.
- Authorization: the permission to fulfill a specified action or to use a particular service.[70]

Although the premises of this reasoning are per se correct, we believe they are incomplete. In our view, the conclusion of the Privacy Commission could have been different, if it would have made a distinction between types of identification (see below on global vs. context-specific or sector-specific identifiers).

In other words, the authorization to access services or resources is indeed typically granted on the basis of *identification of the entity*, but not necessarily on the basis of *global identification* of that entity.[71]

6.4.3 Identifiers

An *identifier* is an attribute or a set of attributes of an entity which uniquely identifies the entity in a certain context or sector.[72] They have a number of characteristics:

[66] [Advice Privacy Commission 24 May 2006], p. 3 and 5.
[67] [DEPREST AND ROBBEN 2003], p. 21.
[68] [DEPREST AND ROBBEN 2003], p. 22.
[69] [Advice Privacy Commission 24 May 2006], p. 3 and 5.
[70] [Advice Privacy Commission 6 July 2005], p. 3.
[71] Another important remark with regard to the decisions of the Privacy Commission is that in advanced IDM systems *identification* is normally not the criterion to grant authorizations.

- They can be *meaningful or meaningless*. Belgian federal eGovernment relies heavily on the usage of the global identification number for natural persons (National Registry number or RRN number). This number is meaningful, since the birth date and the gender can be deduced from it. The National Registry number consists of 11 digits. The first 6 digits stand for the date of birth (2 digits for the year, 2 digits for the month and 2 digits for the day), the 3 following digits stand for the serial number of the registration (even numbers are reserved for women) and the last 2 digits form the verification number (art. 1 et seq Royal Decree 3 April 1984).
- They can be *assigned by private organizations or by the government*. In eGovernment, identifiers are typically assigned by the government itself.
- They can be *sector-specific or context-specific, cross-sectoral or global*. The application area of sector-specific identifiers is limited to a specific (governmental) sector in one or more contexts. The application area of context-specific identifiers is limited to a specific context in one or more sectors.[73] For instance, sector or context-specific identifiers serve to identify an entity in different contexts or sectors via different identifiers, e.g., a health number, a judicial number, or a fiscal number. In contrast, cross-sectoral, cross-context or global identifiers are identifiers that are used in multiple, if not all (government) contexts and sectors. This means that an entity is identified in multiple government contexts and/or sectors via the same identifier.[74]

One crucial interoperability decision in Belgian federal eGovernment has been to choose *one common global identifier* to identify all the relevant entities across several contexts:[75]

- The National Registry number serves as unique (global) identifier for Belgian citizens and foreigners that are registered in the population registers and in the consular and diplomatic registers.[76]
- The Crossroads Bank for Social Security issues an identification number for those natural persons that are not registered in the National Registry.[77] This is the so-called ‘bis-’ and ‘ter-number’.
- Enterprises also receive a global unique identifier at their first registration in the Crossroads Bank for Enterprises.[78]

No parliamentary debate has taken place to justify this practice of using only one global identifier across several government contexts in Belgium.

The Belgian Privacy Commission explained on multiple occasions that the real problem of identification means in general, is not that they are dangerous on their own. Their real risks lay in the fact that they enable interconnections of data repositories. From her point of view, it

[72] [Modinis-IDM 2005], p. 11.

[73] See the presentation [De Cock 2006], slide 3. For more information on the difference between contexts and sectors, see [Modinis-IDM 2006], p. 10.

[74] [De Bot 2005], p. 53-57.

[75] This number is built in the same way as and is synchronized with the National Registry Number.

[76] Article 2 [National Registry Act].

[77] Article 8 §1(2) [Crossroads Bank for Social Security Act 1992].

[78] Article 5 [Crossroads Bank for Enterprises Act 2003].

is important to control these interconnections between data repositories and their purposes.[79] As we will see below, we are convinced that today's control mechanisms are not sufficient.

The Privacy Commission has asked in a number of advices to introduce sector-specific identifiers, at least for health data[80] and in the justice sector.[81] Recent developments seem to indicate that this advice will be followed:

- In its decision nr. 13/2006 of 24 May 2006 about the Phenix project (a large-scale reform of the judicial information architecture, which introduces *inter alia* electronic data exchange of court documents), the Privacy Commission explicitly requests the usage of a sector-specific identifier.
- The lines of force of the planned BE-health platform (a large-scale medium to manage patient data, the platform is currently in a test phase) with respect to identification data are as follows. An irreversible *patient identifier* is calculated from the identification number for social security (INSZ-number), based on an algorithm which is accessible at every health care practitioner. The patient identifier is (1) either unique across all the health-care practitioner / institutions, or (2) unique per patient at a health-care practitioner / institution. Only the independent organization which manages the data exchange is able to make the conversion across the different contexts (different health-care practitioners / institutions). Data which do not require the identification of the patient (anymore) via the patient identifier are encoded or anonymized.

6.4.4 Control mechanisms

It is important to note that the access to and the usage of the identity data of the National Registry, including the national registry number, is strictly limited. Besides the cases where the usage of the National Registry Number is allowed by federal law or (federal) royal decree, its usage is subject to a prior authorization by the thereto appointed sectoral committee of the Privacy Commission. Only the entities listed in the law are entitled to use and/or access this data (Article 8 National Registry Act).

These groups and purposes are Belgian public authorities, natural and legal persons acting as subcontractors of Belgian public authorities, public and private entities (Belgium) as to the information they need for fulfilling a task of public interest, notary publics and bailiffs in relation to their official tasks, pharmacists in case of delivery of dangerous medicines, and lawyers for the fulfillment of their judicial tasks.

These entities should obtain a prior authorization to use and/or to access the personal information contained in the National Registry, including the National Registry number, from an independent sectoral committee that is part of the Privacy Commission. Authorizations are granted by the Sectoral Committee of the Privacy Commission if the access conditions are

[79] See for example [Advice Privacy Commission 10 June 2002].

[80] On the introduction of 'patient identification numbers', see Advice nr. 14/2002 of 8 April 2002, Advice nr. 19/2002 of 10 June 2002, Advice nr. 30/2002 of 12 August 2002, Advice nr. 33/2002 of 22 August 2002, Advice nr.1/2005 of 10 January 2005 and Advice nr. 05 / 2006 of 1 March 2006, all available on the website of the Privacy Commission, www.privacycommission.be in the advice section.

[81] In its decision nr. 06/2006 of 1 March 2006, the Privacy Commission explains that it has been required to prepare a position on the usage of the National Registry Number in the framework of the Phenix project (informatization of Justice) (see [Decision Sectoral Committee RRN 1 March 2006c]). The fact that the Privacy Commission does not want to anticipate its answer, reveals that such a usage might not be self-evident. The advice is available on the website of the Privacy Commission www.privacycommission.be in the authorization section.

met. It mainly includes a verification if the request is compliant with the Belgian Data Protection legislation. The authorizations are subsequently made public via the website of the Privacy Commission.

The usage of the global identifier for enterprises is also (but less stringently) regulated. When the number is being used to exchange other data than the ones contained in the Crossroads Bank for Enterprises, its usage requires the prior notification to the thereto appointed sector-specific sectoral committee of the Privacy Commission (Article 18, §4 Crossroads Bank for Enterprises Act). The usage of the 'bis- and ter-number' is not restricted at all (Article 8 §2 Crossroads Bank for Social Security Act).

The divergent protection between these three types of identifiers results from the fact that the General Data Protection Directive 95/46/EC leaves the protection of unique identifiers totally up to the discretion of the Member States.[82]

6.5 Common authentication means: the Belgian electronic identity card[83]

In this section, we briefly describe some of the technical aspects of the Belgian eID, which can be seen as an example of a so-called PKI infrastructure.[84] An extensive description of this infrastructure can be found in FIDIS deliverable D3.6.

The card looks like a normal smart card (e.g., a bank card) and displays a number of personal and administrative data:

- the identity card holder's name (family name, up to two given names, and the initial of a third name),
- title,
- nationality,
- place and date of birth,
- gender,
- picture,
- two hand written signatures, i.e., the one of the card holder and the one of the civil servant who issued the card,
- validity period of the card (five years),
- the card number,
- *the national Registry Number of the holder,*
- the place of delivery of the card, and
- a machine readable ICAO (international civil aviation organization) zone.

All these visual data are also stored on the chip in a so-called identity file. The residence address of the identity card holder is stored separately, in the address file, to allow easy updating during the validity period of the card. The National Registry digitally signs the address file and the identity file to guarantee the link between both files.[85]

[82] Art. 8 Data Protection Directive states: "Member States shall determine the conditions under which a national identification number or any other identifier of general application may be processed."

[83] The section is largely based on [De Cock, Wolf, Preneel, 2006], p. 93 ff.

[84] A PKI or Public Key Infrastructure is a system of Certification Authorities (and optionally, Registration Authorities 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*.

[85] Based on [De Cock, Wolf, Preneel, 2006], p. 93.

The chip on the card can perform digital signatures and key generation. There are no concrete plans to integrate decryption functionalities in the eID. In total, a Belgian eID holds three different private signing keys: one to authenticate the citizen, one for non-repudiation signatures, and one to identify the card itself towards the Belgian government.

The eID is able to compute digital signatures with all of them. For the citizen's authentication key and non-repudiation signature key, this is only done after the card holder enters a PIN. It is relevant to note here that both certificates also contain the globally unique identifier of the certificate holder.

6.6 Analysis: a pseudonymous, user-centric infrastructure

6.6.1 Global identification and control in Belgian eGovenrment.

One of the main building blocks of Belgian federal eGovernment is to use the globally unique identifier for natural persons to achieve interoperability in general, for the exchange of data about these persons across several contexts. This is a new evolution initiated by eGovernment with important consequences. We come back to this below.

In addition, various types of additional information can now also be requested from the National Registry, for example for a number of identity management-related tasks (such as user management or mandate management), also outside the context of the National Registry. This creates globally unique identification and authentication of the concerned person (either user or subject of the exchanged data), and – at least in theory – linkability of these several contexts.

Moreover, it also appears that the content of the new Belgian electronic identity card is protected against forgery, but not kept confidential: when the card is being inserted in a card reader, the data contained in the identity file can be accessed and stored, even without the user's consent. Even though this easy storage capacity of identity data requires physical presence of the card, it goes without saying that it has important consequences which were triggered by eGovernment.

Before the eID era, someone could of course also make a paper copy of the identity card when having the card at his or her disposal. The difference lays mainly in the fact that with the eID this can be done in a much more systematic way, without the card holder even noticing it, or being able to object to it.

For the sake of clarity, in Belgium, the identity card serves in the first place as a proof of registration in the population registry. There is an obligation to carry the card with you, starting from the age of 15 ("kidscards" are being issued from the age of 12). Each identity card holder shall present the card (1) to the policy, when requested, (2) when demanding certificates (in general), (3) to bailiffs with regard to a writ of summons, and (4) (also in general) whenever he or she has to offer a proof of his identity (art. 1 Royal Decree 25 March 2003).

An open question in this context is whether any service provider (e.g., the shop that rents out videos, or your supermarket) is allowed to request a proof of identity. This is far from clear. The legal answer would be: only if the proof of identity is necessary to provide the services (based on the finality and data-minimization principles of the general Data Protection legislation).

One should admit, however, that it is common practice. As far as we know, no case-law is yet available on the topic, and it would be interesting to verify whether service providers are fully aware of the obligations they have under data protection legislation whenever they are processing eID data (e.g., notification to the privacy commission, legitimacy ground, finality principle, right to access, correct, and obligation to inform).

What we have just explained with regard to eID data also applies to the National Registry Number in general. Only a limited number of entities are allowed to process the number (see above, the entities listed in Art. 8 of the National Registry Act), but many entities are now increasingly becoming aware of it. Only the usage of the number (which means storing it separately, doing something with it) is a type of processing that falls under the terms of the law.

An interesting side-note is that before the eID era, people could object to having their National Registry number printed on the card. This is no longer the case. Moreover, the number is now also available electronically. It is contained in the X.509 certificates. The same applies to the National Registry number, which is contained in the X.509 certificates that are being used for entity authentication and non repudiation purposes. In other words, with each exchange of a digital signature created with the Belgian eID, that is, with each authentication or digitally signed document, the signatory further propagates the National Register number. It goes without saying that this increases the potential misuse of that number.

6.6.2 Global identification versus sector or context-specific identifiers

When identification is always required, and data exchange is done via a global identifier, it is possible that even though a number of data interconnections are not authorized, or illegal, they will take place anyway.

As soon as the necessary infrastructure is in place, it cannot be excluded that other decisions are made in the future, based on ad hoc arguments or on different political choices. Therefore, the decision whether or not to base the identity management infrastructure on anonymity or pseudonymity as the default position or rather on identity as the default is a fundamental one.

- The first option is to choose an *identity society*, where technology has *carte blanche*, and where the question of how fundamental rights (such as privacy) are being protected depends on the subsequent political decisions.
- The other option is to choose a *pseudonym society*, where technology is intentionally limited and regulated in advance, and where fundamental principles are incorporated in the architecture design.[86]

As explained, the most important risk connected with the usage of a globally unique identifier such as the National Registry Number lays in the fact that it makes it technically possible to link data about an entity from one context to the other, without any form of control by the entity itself. This risk could be avoided via the usage of sector or context-specific identifiers or pseudonyms.

A *pseudonym* is an identifier of a specific entity's (partial) identity, by which a certain action can be linked to this entity. It can be defined as an identifier of an identifiable entity that is

[86] [Koops 2001], p. 1560.

either self-chosen by this entity or assigned by a provider, to identify this entity to a reliable party for a period of time.[87]

A number of gradations in pseudonyms can be imagined. Depending on the frequency with which they are used, we can distinguish different types of pseudonyms:

- *Persistent pseudonyms*: pseudonyms used for an extended period of time that spans multiple sessions. Persistent pseudonyms can be used to represent an *identity federation* (we come back to this term below).[88]
- *Transient pseudonyms*: pseudonyms used for a relatively short period of time that need not span multiple sessions. [89]

Pseudonyms allow forming sets of partial identities which are not necessarily linkable to the originating entity. With respect to the degree of the resulting linkability between the applicable sector and context and other sectors and contexts, various kinds of pseudonyms (or identifiers) can be distinguished:[90]

- A *person identifier, person pseudonym,* or *global identifier* is a substitute for the holder's name, which is regarded as the representation of the holder's civil identity. It may be used in multiple, if not all, contexts and sectors.[91] Examples of this type of identifier are the unique number of an identity card, the social security number, the Belgian National Registry Number, the Belgian Crossroads bank for Enterprises number, a nickname etc.
- A *context- / sector-specific identifier* or *context- / sector-specific pseudonym* is an identifier that has only meaning to the communication partners within a specific context or a specific sector.
- An *opaque handle* is an identifier of the entity that has meaning only in the context between a specific identity provider and a specific service provider.
- A *role identifier* or *role pseudonym* is an identifier limited to specific roles, such as a customer pseudonym or an internet user name; it can be an assigned or a chosen identity.
- A *relationship identifier* or *relationship pseudonym* means that for each communication partner a different identifier of the entity is used, but the same relationship pseudonym could be used in different roles for communicating with the same partner (e.g., nicknames).
- A *role-relationship identifier* or *role-relationship pseudonym* means that for each role and for each communication partner, a different role-relationship identifier is used. The communication partner does not necessarily know whether two identifiers used in different roles belong to the same holder.
- A *transaction identifier* or *transaction pseudonym* is used for only one transaction. It is unlinkable to any other transaction identifiers and is (at least initially) unlinkable to any other item of interest (e.g., randomly generated transaction numbers for online-banking).

[87] Based on [Hodges 2006] p. 8, and [MODINIS 2005], p. 15.
[88] [Hodges 2006], p. 8.
[89] [Hodges 2006], p. 11.
[90] Based on [PFITZMANN 2006] p. 17-18. We have completed the list with opaque handles and context- / sector-specific pseudonyms.
[91] Based on [PFITZMANN 2006] p. 17-18 and [DE BOT 2005], p. 53-57.

Partial identification is the identification of an entity via a context- or sector-specific identifier or a combination of one or more characteristics, in only one context or sector.

6.6.3 Pseudonymity and trusted agents

In the literature, one often hears criticism of the usage of sector-specific identifiers, namely that it only makes data exchange more difficult also in those cases where data exchange is allowed. The point those authors want to make is that sector-specific identifiers do not prevent data linkage, but just make it more complicated, as data can still be linked via concordance tables.[92] We believe, however, that this can be solved through the usage of trusted agents.

A *trusted agent* is a trusted party, a trusted third party, or a trusted device that acts as an intermediary to forward authorized, properly authenticated service requests to remote service providers. The trusted agent can – as an (additional) privacy service – obfuscate the original requestor's identity. In other words, if the entity is authorized to request a particular service, it will forward the service request as if it was its own. In practice, if the trusted agent also offers privacy services, it will typically do the conversion between one unique identifier of the entity that accesses the conversion service and the different sector or context-specific identifiers used by different service providers.

6.7 Analysis: an identity, government-centric infrastructure

Based on our analysis of current Belgian federal eGovernment, it seems obvious that it relies heavily on globally unique identification via globally unique identifiers and other authentic identity data that applies across several contexts. In other words, Bentham's words on identification and control (*supra*, section 2.4) are still very relevant in Belgium today: "who are you? With whom am I dealing? There [is] no room for evasion in answering this important question".

What if changing the Benthamite identification scheme – apart from the contexts where the Belgian Privacy Commission agrees that this is indeed needed (namely in judicial and health matters) – is not a realistic option, for example because it would hinder the development of eGovernment? We explained earlier that privacy is not an absolute right, and there are other valuable interests, such as the public interest, which may limit the right to privacy, especially in a governmental context.

The question we asked earlier is whether some degree of privacy could still be provided in case a government chooses for a Benthamite approach on identification and identity knowledge. We believe the answer to this question is yes. As will be explained in more detail in FIDIS deliverable D16.1, besides IDM systems that focus on obfuscation of the identity, for example via user-controlled context-dependent role and pseudonym management (FIDIS type 3 IDM systems), there are other legally admissible ways to incorporate privacy and data protection requirements in an IDM architecture used in eGovernment. One could, for example, take into account the model set out above, with prior authorizations to exchange data by a sectoral committee of the Privacy Commission.

From a privacy and data protection perspective, it is important that in the Belgian federal case, information is, at least in theory, only made accessible and exchanged with thereto authorized entities, based on formal authorizations by (a subcommittee of) the Belgian

[92] See [De Bot 2005] p. 61, citing Mr. Frank Robben.

privacy commission. In sum, privacy enhancements and compliance with data protection regulation in such an alternative model is not gained via the empowerment of the user, but via control by a trusted third party.

The privacy enhancements we propose to make to this system are to make data processing of the personal data of their clients more transparent via monitoring and feedback procedures (transparency tools) and by incorporating basic data protection and privacy requirements in the basic architecture design. A potential "privacy-friendly" solution could then be to make the prior authorizations of the Belgian Privacy Commission enforceable via privacy policy enforcement. More precisely, data handling and data release policies could help enforcing authorizations of the above-mentioned trusted third party. This will be further elaborated in the deliverables of FIDIS workpackage WP16 on eGovernment.

In summary, through transparency, accountability, and enforcement mechanisms applied to the processing of "authentic" data about citizens in eGovernment, 'data owners' can regain more control, autonomy, self-determination, and thus privacy with regard to their data.

6.8 Conclusion

In this chapter, we have tried to find an answer to the question whether the identity knowledge of the government is growing through the development of eGovernment in Belgium. The answer to this question is yes. We have focused on the usage of common identification keys as a prerequisite to achieve interoperability in Belgian federal eGovernment and explained which identity data and identifiers are being provided and used in eGovernment (sections 6.4 and 6.5).

In section 6.6, we introduced the term "globally unique identifier" and analyzed what the consequences are of global identifiers for identity knowledge. We explained that we should be aware of the difficult choice between an identity society (such as defended by Bentham) and a pseudonym society:

- In the *identity society*, technology has *carte blanche*, and the question of how fundamental rights (such as privacy) are being protected depends on the subsequent political decisions.
- In a *pseudonym society*, technology is intentionally limited and regulated in advance, and fundamental principles are incorporated in the architecture design.

Apart from a few exceptions, in Belgian federal eGovernment, the first option has clearly been chosen, via the decision to build its data exchange on globally unique identifiers which allow identification and data linkage across contexts.

We thereby found out that one additional question should be asked, namely whether, when a government like the Belgian federal government chooses a Benthamite approach on identification, and identity knowledge thus indeed increases, some degree of privacy can still be provided to its citizens. We believe the answer to this question is yes as well, and we have made a number of suggestions for future research, based on transparency, accountability, and the enforcement of data protection and privacy requirements.

6.9 Bibliography

Legislation and agreements

[National Registry Act 1983] Law of 8 August 1983 on the organization of a Registry of natural persons, Belgian State Gazette 21 April 1984.

[Crossroads Bank for Social Security Act 1990] Law of 15 January 1990 on the creation and the organization of a Crossroads Bank for Social Security, Belgian Sate Gazette 22 February 1990, err. 2 June 1990, err. 2 October 1990.

[Crossroads Bank for Enterprises Act 2003] Law of 16 January 2003 on the creation of a Crossroads Bank for Enterprises, on the modernization of the trade register, on the creation of recognized companies' dockets and on diverse rules, Belgian State Gazette 5 February 2003.

[Vanvelthoven 2006a] Policy letter of the federal Minister competent for eGovernment, Peter VANVELTHOVEN, http://www.belgium.be/eportal/ShowDoc/kabinet_egov/imported_content/pdf/Beleidsnota_Peter_Vanvelthoven_2005_NL.pdf?contentHome=entapp.BEA_personalization.eGovWebCacheDocumentManager.nl, 2005.

[Vanvelthoven 2006b] Policy letter of the federal Minister competent for Employment, Peter VANVELTHOVEN, on the topic of informatics and the government: http://www.belgium.be/eportal/ShowDoc/kabinet_egov/imported_content/pdf/beleidsnotainformatisering06.pdf?contentHome=entapp.BEA_personalization.eGovWebCacheDocumentManager.nl, 2006.

[Data Protection Act 1992] Law of 8 December 1992 on Privacy Protection in relation to the Processing of Personal Data, Belgian State Gazette 18 March 1993, as modified by the law of 11 December 1998 implementing Directive 95/46/EC, Belgian State Gazette 3 February 1999, and the law of 26 February 2003, Belgian State Gazette 26 June 2003.

[Data Protection Directive 1995] 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 on the free movement of such data, Official Journal of the European Communities, L 281, 23 November 1995, 31-50.

[Reform of the Data Protection Act 1998] Act of 11 December 1998 on the transposition of the [Data Protection Directive], Belgian State Gazette 3 February 1999.

Case Law

[Decision Sectoral Committee RRN 6 July 2005] Decision of the Sectoral Committee of the Belgian National Registry on the demand of FEDICT to use the National Registry Number for its user management, number SA2/RN/2005/023/006berdef, 6 July 2005, available at: www.privacycommission.be, last visited: 26 February 2006.

[Decision Sectoral Committee RRN 6 September 2006] Decision of the Sectoral Committee of the Belgian National Registry on the demand of the Royal Federation of Belgian Notaries to receive access to the National Registry number for user management, number SA2/RN/2006/023, 6 September 2006, available at: www.privacycommission.be, last visited: 12 October 2006.

[Decision Sectoral Committee RRN 1 March 2006c] Decision of the Sectoral Committee of the Belgian National Registry on the demand the National Chamber of Bailiffs to receive access in its own name and acting as proxy for its members to the information data of the National Registry and to use the identification number of the National Registry, a.o. in the context of implementing art. 139 and 140 of the Mortgage law (Hypotheekwet), participation in the Phenix-project and the system of eBailiff, number SA2/RN/2006/003, 1 March 2006, available at: www.privacycommission.be, last visited: 29 March 2006.

[Advice Privacy Commission 10 June 2002] Advice of the Belgian Privacycommission on a law proposal to modify the law of 8 August 1983 (National Registry Act) and the law of 19 July 1991

(Population Registry Act), number 10/A/2002/003, 10 June 2002, available at: http://www.privacycommission.be, last visited: 10 May 2006.

[Advice Privacy Commission 21 October 2004] Advice of the Belgian Privacycommission on article 133 and 134 of the Program law of 8 April 2003 about identification documents, number A/2006/003, 24 May 2006, available at: www.privacycommission.be, last visited: 12 October 2006.

Doctrine

[BAUER 2005] M. BAUER, M. MEINTS, M. HANSEN (eds.), ‘D3.1, Structured Overview on Prototypes and Concepts of Identity Management Systems’, Deliverable 3.1 of the EU Funded FIDIS project available at: http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp3-del3.1.overview_on_IMS.final.pdf, 15 September 2005, last visited: 20 August 2006.

[BUCHTA 2005] A. BUCHTA, J. DUMORTIER and H. KRASEMANN, ’The Legal and Regulatory Framework for PRIME’, in S. FISHER-HUEBNER, CH. ANDERSSON and TH. HOLLEBOOM (eds.), ‘PRIME D 14.1.a: Framework V1’, available at: https://www.prime-project.eu/prime_products/reports/fmwk/pub_del_D14.1.a_ec_wp14.1_V4_final.pdf , 13 June 2005, last visited: 13 June 2005, 101-150.

[CLARKE 1999] R. CLARKE: ‘Introduction to Dataveillance and Information Privacy, and Definitions of Terms’, available at: http://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html, 16 September 1999, last visited: 7 November 2005.

[CHAMBER OF REPRESENTATIVES WORKSHOP 2004] Belgian Chamber of Representatives, ‘The topic of the role of the legislator in the information society’, Workshop report published online at: http://www.dekamer.be/kvvcr/pdf_sections/publications/ict/QT03_PR.pdf, 8 March 2004, accessed 18 April 2006.

[DE BOT 2001] D. DE BOT, ‘Verwerking van persoonsgegevens’, *Kluwer*, Antwerpen, 2001, p. 23.

[DE BOT 2005] D. DE BOT, ‘Privacybescherming bij e-government in België. Een kritische analyse van het Rijksregister, de Kruispuntbank van Ondernemingen en de elektronische identiteitskaart.’, *Vandenbroele*, Brugge, 2005, 469 p.

[DE COCK, WOLF, PRENEEL, 2006] D. DE COCK, Ch. WOLF, B. PRENEEL, ‘The Belgian Electronic Identity Card’, Chapter 5.6 of FIDIS deliverable 3.6, published online at: www.fidis.net, 31 March 2006, accessed 28 June 2006.

[DE COCK 2006] Presentation by D. DE COCK, ‘Derived structures – conceptual framework. Third MODINSI IDM workshop’, published online at: https://www.cosic.esat.kuleuven.be/modinis-idm/twiki/bin/view.cgi, 9 February 2006, accessed 5 March 2006.

[DEPREST AND ROBBEN 2003] J. DEPREST, F. ROBBEN, ‘eGovernment: the approach of the Belgian federal administration’, published online at: http://www.ksz.fgov.be/En/Como/2003%20-%20EGovernment%20paper%20v%201.0.pdf, June 2003, accessed 20 June 2006.

[DUMORTIER 2003] J. DUMORTIER, C. GOEMANS and M. LONCKE, *General report of the legal issues*, Deliverable D4 of the APES project, available at https://www.cosic.esat.kuleuven.ac.be/apes/docs/APES_d4.doc.gz, March 2003, last visited: 25 April 2005.

[FEDICT 2002] X., ‘De Universal Messaging Engine Versie 2’, published online at: http://www.belgium.be/eportal/ShowDoc/fed_ict/imported_content/pdf/FunctUME225022002.pdf?contentHome=entapp.BEA_personalization.eGovWebCacheDocumentManager.nl#search=%22UME%20directory%22,25 February 2002, accessed 20 June 2006.

[HILDEBRANDT 2005b] M. HILDEBRANDT, S. GUTWIRTH and P. DE HERT (eds.), `Fidis Deliverable 7.4, Implications of profiling practices on democracy and rule of law', available at:

http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp7-del7.4.implication_profiling_practices.pdf; 5 September 2005, last visited: 15 September 2006.

[HODGES 2006] J. HODGES (ed.), 'Liberty Technical Glossary', published online at: http://www.projectliberty.org/specs/draft-liberty-glossary-v2.0-05.pdf, 9 June 2006, accessed 15 June 2006.

[IDA 2004] X. 'IDA Authentication Policy. Basic policy for establishing the appropriate authentication mechanisms in sectoral networks and projects.', published online at: http://ec.europa.eu/idabc/servlets/Doc?id=19281, 8 July 2004, accessed 20 October 2006.

[ITU-T 2005] X, 'Security Compendium Part 2 - Approved ITU-T Security Definitions', available at: http://www.itu.int/ITU-T/studygroups/com17/def005.doc, May 2005, last visited: 20 April 2006.

[KOOPS 2001] B-J. KOOPS, 'Een nieuwe GBA, digitale kluisjes en identificatiedrang', *Nederlands Juristenblad*, 2001, 1555-1561.

[KSZ 2005] X., http://ksz-bcss.fgov.be/En/CBSS.htm#Results, accessed 24 October 2005.

[MACDONALD 2003] M. MACDONALD, 'Data Registries Comparison Report', http://www.itsregistry.org.uk/Comparison%20Report%20v3.doc, March 2003, accessed 20 August 2006.

[MEINTS 2006] M. MEINTS and M. HANSEN (eds.), 'FIDIS D3.6 Study on ID Documents', Deliverable 3.6 of the EU funded FIDIS project, available at: http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp3-del3.6.study_on_id_documents.pdf, December 2006, last visited: 16 January 2007.

[MODINIS-IDM 2005] X., 'Modinis IDM Terminology Paper', published online at: https://www.cosic.esat.kuleuven.be/modinis-idm/twiki/pub/Main/GlossaryDoc/modinis.terminology.paper.v2.01.2005-11-23.pdf, 23 November 2005, accessed 22 December 2005.

[MODINIS IDM 2006] X., 'Modinis IDM Conceptual Framework v.1.1', available at: https://www.cosic.esat.kuleuven.be/modinis-idm/framework/, 18 September 2006, last visited: 18 September 2006.

[PFITZMANN 2006] A. PFITZMANN, M. HANSEN, 'Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology', published online at: http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.27.pdf, 20 February 2006, accessed 20 February 2006.

[RFC3281] X., 'An Internet Attribute Certificate Profile for Authorization', published online at: http://www.ietf.org/rfc/rfc3281.txt, April 2002, accessed 20 May 2006.

[RFC3198] X. 'Terminology for Policy-Based Management', published online at: http://tools.ietf.org/html/rfc3198, November 2001, accessed 20 May 2006.

[ROBBEN 2004a] F. ROBBEN, 'EGovernment: the approach of the Belgian federal administration', published online at: http://www.law.kuleuven.be/icri/frobben/presentations/20041110.ppt, 10 November 2004, accessed 22 October 2005.

[ROBBEN 2004b] F. ROBBEN, 'EGovernment in the Belgian social security sector: a successful combination of back-office integration and an e-portal solution', published online at: http://www.law.kuleuven.be/icri/frobben/presentations/20040513.ppt, 13 May 2004, accessed 25 October 2006.

[ROBBEN 2005a] F. ROBBEN: '1st Modinis Workshop on Identity Management in EGovernment', published online at: http://www.law.kuleuven.ac.be/icri/frobben/presentations/20050504.ppt, 4 May 2005, accessed 30 March 2006.

[ROBBEN 2005b] F. Robben, 'eGovernment in the Social Security Sector', published online at:http://www.law.kuleuven.be/icri/frobben/presentations/20050224.ppt, 24 February 2005, accessed 22 October 2006.

[ROBBEN 2006] F. ROBBEN, 'eGovernment', available at: http://www.law.kuleuven.be/icri/frobben/presentations/20060327b.ppt, 27 March 2006, last visited: 1 April 2006.

[SLONE 2004] S. SLONE, 'Identity Management. A white paper.', published online at: http://www.opengroup.org/onlinepubs/7699959899/toc.pdf, March 2004, accessed 18 August 2006.

[STRICKX 2004] P. STRICKX, J. JOCHMANS, 'Richtlijnen en aanbevelingen voor het gebruik van open standaarden en/of open specificaties bij de federale overheidsbesturen', published online at: http://www.ksz.fgov.be/documentation/nl/documentation/Pers/OpenstandaardenNL_FEDICT.pdf, 6 October 2004, accessed 10 April 2006.

[WESTIN 1970] A.F. WESTIN, 'Privacy and Freedom', New York (USA), Atheneum, 1970, 487 p.

7 Anonymization of official statistical data

7.1 Introduction

Public institutions gather and process data for a wide variety of purposes. In many cases, it is useful for them to process personal data, i.e., data connected to identifiable individuals. However, large-scale processing of personal data allows data-bases to be combined through shared identifiers. This would significantly enhance the identity knowledge capacity of governments; particularly given the sensitive nature of medical data, serious privacy risks are associated with this type of data collection and processing.

Since for certain types of purposes, such as statistical analysis of official data in the health sector, it is not always necessary to process personal data, anonymization of data is an important tool to mitigate the privacy risks of large-scale data processing. In this chapter, we describe how anonymization can be implemented in a particular public sector without threatening the purposes of the data processing. For this, we have chosen two case studies in the health sector. These are relevant for the purpose of this deliverable, since they show that identification data need not necessarily be processed even when exploiting new opportunities of electronic data processing, in this case to generate statistical data throughout patients' lifetimes.

The health sector is a good example to illustrate problems arising from the possession and processing of data. Typically, medical and health data contain identifying data, as well as sensitive data associated with these, namely data about one's health situation, treatments, etc. These data are considered privacy-sensitive for almost everybody, and any misuse may immediately induce privacy concerns. Misuse may provoke serious disadvantages for people, as databases in the medical sector may contain very sensitive information about one's life, one's preferences, one's problems, etc.

Data contained in such a database can be used for multiple, questionable, purposes. For instance:

- An insurance company could use health data to measure the risks a person has to suffer from a certain disease and consequently ban him from any contract.
- An employer may be interested in knowing if a female employee wants to conceive children (easy to find: is she buying contraceptives or not?). In most European countries, female employees are protected when expecting a child, but not when trying to conceive.
- A banker may not grant a loan to a client who has a high risk of contracting a cancer.

Governments try to tackle these problems and to focus on the privacy concerns connected to the processing and possession of health data. Legislation specifies clear procedures to be followed in order to guarantee one's privacy up to a certain point. Additionally, procedures and protocols can help to guarantee anonymity to some extent in this context.

We explain in this chapter that anonymization techniques can, from a technical perspective, help to guarantee non-identifiability in patients' records filed in central medical databases. We show this using two case studies from Switzerland, a federal agency collecting health data, and an international data collection for orthopedic evaluative research.

7.2 Anonymity and anonymization

Considering anonymity, definitions try to focus on a precise idea. Consider for example a definition of anonymous:

"1 : not named or identified

2 : of unknown authorship or origin

3 : lacking individuality, distinction, or recognizability".[93]

Being anonymous according to this definition means that the subject is not identified; it cannot be distinguished from other subjects within a specified group of subjects. Often this reference group is only implicitly specified. The term anonymity is then mainly used to address the state of being anonymous.[94]

The definition from [ISO 15408-2] goes in the same direction: Anonymity "ensures that a user may use a resource or service without disclosing the user's identity. The requirements for Anonymity provide protection of the user identity. Anonymity is not intended to protect the subject identity."

In this chapter, we will focus mainly on anonymization, a technical term which addresses the problem of how one can attain the state of anonymity, while starting from a non-anonymous situation. Typically this problem arises when a large quantity of personalized data is available which needs to be used for statistical purposes, yet actual law does not allow for this data to be given to third parties in order to guarantee privacy protection of individuals or groups.

Typical examples of this are to be found in the large field of profiling and data mining [Hildebrandt & Gutwirth 2007]. Legal regulations in the EU and Switzerland are very strict, and if profiling and data mining are to be done successfully, then legally compliant, anonymization techniques are very helpful.

Many challenges are to be met during the process of anonymization. Some of them are technical in nature like, for example, the question of which algorithms, which protocols are to be used. Many challenges go beyond the technical level, for example the question of whether or not there must be a possibility to reverse the anonymization procedure, and, if so, to what extent and at what "cost"; i.e., only the people allowed to do so should be able to actually do it.

In the following, we will focus on the procedures for anonymization but not on technical details on the algorithmic level of cryptographic procedures and protocols in order to keep it simple and readable for people without cryptological training.

7.3 Statistical information in the health sector

The health sector is a very interesting field for collecting and processing information. There is an enormous amount of information generated by patients treated in hospitals, which bill the respective costs to central insurance agencies, etc. Clearly, most of these data is very sensitive, as personal information is almost always integrated therein. Typically, we have data

[93] From http://www.m-w.com/dictionary/anonymous (25.09.2006). See also Pfitzmann, A., Hansen, M., Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology, v0.28, May 29, 2006, p.6, http://dud.inf.tu-dresden.de/Anon_Terminology.shtml.

[94] Note that we will not enter the discussion of whether, where, when, and why anonymity is advisable. Consider for example [Ström 2005] for a discussion of the problematic profiling issues we're focusing on.

tuples that indicate which person (identified by some number, name, etc.) has had which treatment by which specialized person when and where, and whether the treatment was successful, etc.

These data are useful for the parties generating them, but also for other parties, for example for health insurers. They typically need substantial amounts of data for different purposes; in particular, identifying data are essential to charge the correct party.

Another typical application focuses on statistical evaluation of patient data for general purposes. In this context, the identifying data are typically no longer needed. The concept of data mining is usually the central focus, used for example to provide new information computed from the data, or to make forecasts for planning purposes based on both current and historical data.

Yet often in this very context of data mining and statistics, there is a need to be able to "follow" the same person through different treatments in time and/or space. This is, for example, necessary to gain insight into the success rate of long-term treatments or to compute statistics about the reasons for changing one's personal medical doctor. One wants to find a possible correlation between certain treatments and either complications or health recovery. Clearly, the simplest implementation of this "following" requirement (linkability) is to keep the identifying parts of the data in the collection.

At first glance, providing anonymity as well as linkability could seem to be a contradiction. Pseudonyms make it possible to meet both objectives. In this context, we consider pseudonyms calculated by means of cryptographic protocols, as we will show below.

7.4 Ingredients for anonymization techniques

Different techniques are available to anonymize data. The most effective one is to destroy parts of the data, namely the identifying parts. However, for the applications sketched above, this is not possible, since some linkability has to remain.

The main ingredients for anonymization of data [BFS 1997, Jaquet-Chiffelle & Jeanneret 2001] come from the field of cryptography and we sketch them here without going into implementation or parameterization details which can be found for example in [Bauer 2000, Stallings 1998]. The core idea is to create a portable calculated pseudonym that protects the true identity of the patient efficiently.

- **One-way hash-functions**

  A hash-function takes as input any bit-string (or character string) and produces a bit-string of predefined length. The goal is to produce a "fingerprint", or hash-code, of the input that has characteristics analogous to the human fingerprint, i.e., two persons (usually) have different fingerprints.

  Hash-functions must ensure uniqueness, i.e., given two different hash-codes, there must be two different input strings which have generated them. On the other hand, a hash-function does not guarantee that two different strings do have different hash-codes.[95] Another requirement is that they are computationally easy (i.e., fast to compute). A typical

[95] However, it should be improbable that two different strings have the same hash-code; the actual applications imply the degree of how improbable this is.

application in computer science is the so-called hash-tables where the hash-codes determine the slot the data is put into.

Additionally, here the one-way characteristics requires that there is no (practically feasible) way to create an input which generates a *given* hash-code. A cryptographic hash-function has also the specificity that it is infeasible (in a reasonable amount of time) to create a collision: two elements having the same hash-code. Hence, they are cryptographically meaningful.

Typical example: SHA-1, SHA-2, MD5

- **Symmetric cryptography**

  Also known as secret key cryptography. Very fast algorithms exist for encryption and decryption using symmetric cryptography, which means that the same key is used for both encryption and decryption. The sender and the receiver must know the very same key; they share a common secret. The generation of such a key is computationally trivial given a good source of random numbers, but the “transport” of the key is a major issue.

  Typical example: AES (Advanced Encryptions Standard), IDEA, Triple-DES

- **Public-key cryptography**

  Also known as asymmetric cryptography. In contrast to symmetric cryptography, public-key algorithms have different keys, one for encryption, and another one for decryption. This allows one to distribute freely his public-key so that anyone is able to send him an encrypted message. The receiver keeps secret his private-key that allows him to decrypt the encrypted message. Public-key cryptography makes it possible to create digital signatures. Disadvantages are the computationally expensive key-generation and – in comparison to symmetric encryption – the much slower algorithms for encryption and decryption.

  Typical example: RSA.

These ingredients are used for the protocol described in the next section.

7.5 Case study: a federal agency collecting health data

In this section, we present the current situation in Switzerland, where a central federal agency, namely the Swiss Federal Statistical Office, collects health data, especially from hospitals, from all over Switzerland in order to compute statistical outcomes of different sorts. This means that many data-generating partners must send their sensitive data to the central agency in a secure and reliable way.

All hospitals in Switzerland (more than 400) are connected to this system. Some of them send information directly to the Federal Office while in some cantons (Swiss Federal States) the data are gathered in Cantonal Offices that check the data and forward them to the Federal Office.

The concepts presented [BFS 1997, Jaquet-Chiffelle & Jeanneret 2001] have been used successfully for several years now to protect the medical data during the transmission phase as well as in the databases of the central agency. Note that in Switzerland, data protection is strictly regulated and hence in the context of identifying data, its transmission as well as its storage is not possible without protection measures. In the following, we present the case of Switzerland in a summarized and slightly simplified way.

Consider the situation where a hospital wants to transmit its data records to the Federal Statistical Office. The data consists of tuples, and each such tuple contains on one hand parts which are identifying (like name, first name, date of birth, etc) and on the other hand treatment data which are usually[96] not identifying.

Step 1: Unified representation of identifying data

The data which is identifying is processed in order to remove "noise" introduced by wrong spelling. This is a frequent reason for duplicated entries in databases. One source is just spelling errors, but another typical example which can induce such cases are names which can be written in different ways. Consider for example the German Umlauts: often, the name "Müller" is written as "Mueller".

There are frequently used algorithms to reduce this source of ambiguity from the data: compression algorithms like Soundex[97] or other phonetic algorithms are used for this purpose. They reduce spelling ambiguities by compressing the names to some sort of "spoken normal form". Typically, in such an algorithm, double letters are reduced, silent letters eliminated, etc.

The identifying data is then grouped according to an algorithm to some normal form like for example John Doe, born 1.1.2000, male, is represented by

01012000MDOEJOHN

assuming that the algorithm did not change the name and first name. The Soundex algorithm replaces names of arbitrary length by codes of fixed length.

Step 2: Generating linking code

The linking code is generated from the unified representation using a one-way hash-function (e.g., SHA-2). The result is a bit-string of a predefined length (usually determined by the algorithm, for example 160-bits). This linking code is then used as an identifier for the patient, in other words a digital "fingerprint". Note that anyone in possession of the personal data of some person can create this linking code; there is no secret in the generation of the linking code itself. Hence this linking code can clearly not be used as is for data processing.

On the other hand, given a linking code, one cannot compute the original data anymore because of the one-way property of the underlying hash-function. The only possibility is to make a so-called dictionary attack, which clearly is still possible: find the personal information of every citizen in Switzerland, compute for everyone the linking code and compare. Or just generate all possible combinations of last names, first names, date of birth and gender.

The goal of this procedure is to create a new pseudonymous ID for each individual. Clearly, due to the nature of the hash-function, there is no guaranteed one-to-one correspondence between individuals and linking codes as, with a very small probability (see above), different

[96] Note that treatment data can also be identifying when other data are available. Consider for example the situation where you know that your neighbor has had some very special treatment done last Friday. Using this information, you can make a link from your neighbor to treatment data if the treatment itself is very rare. Techniques to circumvent this are possible, for example by only probabilistically linking identifying data and treatment data, hence by integrating some "noise" in the links.

[97] See http://en.wikipedia.org/wiki/Soundex for a description of this algorithm.

strings are mapped to the same linking code by a hash-function. Yet as this probability is very small[98], the influence on any resulting well-founded statistical outcome is not significant.

Step 3: Encrypted linking code for transmission

For the transmission of the data from the hospital to the Swiss Federal Statistical Office, the linking codes are encrypted. In principle it is sufficient to encrypt only the linking code itself and not the medical data itself as an attacker can link the medical data only to the encrypted linking code, and hence not to an individual. Yet the medical data itself should be protected during transmission as well, in order not to allow attackers to use this information as identifying data in special cases (see footnote 3).

For the encryption, a hybrid encryption approach is used, i.e., a symmetric algorithm is used for encrypting the data (as a large quantity of data is transmitted) using a so-called secret session-key which is typically generated for each session (validity restricted to an hour or to a specific day) individually. After finishing the session, the key is destroyed and another one generated. Clearly, this very session key must also be transferred to the Federal Statistical Office (cf. next step).

Step 4: Encrypted secret key for transmission

The secret session-keys are encrypted with public-key cryptography using the known public-key of the Swiss Federal Statistical Office. Remember that public-key algorithms allow anyone knowing the public-key to encrypt the message, but only the party possessing the private-key to decrypt the message. This transmission is only necessary when a new session key (see step 3) has been created, i.e., at the beginning of a new session. The public-key of the Swiss Federal Statistical Office must be available to all parties through a channel that guarantees the authenticity of the key.

Step 5: Generating uniform linking code

At the Swiss Federal Statistical Office, first the secret session-key is decrypted (using the private-key) and the recovered session-key then used to decrypt the linking codes. The linking codes themselves are then re-encrypted using a symmetric algorithm in order to break the link to the original linking codes and to prevent dictionary attacks. The result of the encryption is called a uniform linking code. The secret-key used for this encryption should be shared between different entities using standard secret sharing algorithms. The original linking codes are destroyed.

The process described using these five steps has several properties:

- The process is simple and needs only standard cryptographic procedures
- The linking codes are generated using a standard and simple procedure which is computationally feasible at a hospital.
- For the same patient at different hospitals, the same linking code is generated. Hence the Federal Statistical Office will be able to "follow" the patient.
- During the transmission phase, linking codes are encrypted, hence an attacker would have to break a standard symmetric encryption algorithm to access the linking codes.

[98] The probability of having at least one collision with 7,000,000 people (Swiss population) is about 0.00000133 for a 64-bits fingerprint and about 0.000000000000000000000000000000000017 for a 160-bits fingerprint!

- The uniform linking codes are anonymous because they cannot be connected to the linking codes anymore. There is no dictionary attack possible. Even when knowing the personal data of the whole population, one cannot link a uniform linking code to a person, as the uniform linking code itself can only be generated using the secret key.
- The security at the Federal Statistical Office depends on the protection of the Secret Key.

Note that uniform linking codes are in general destroyed after 10 years [BFS 1997]: re-encryption takes place according to a predefined scheme in order to change the Secret Key used to generate the uniform linking codes.

While the presented procedures are implemented in practice, an extended framework was presented [Baumann et al. 2005] introducing a recovery authority which can be separated from the central data collector. All tasks regarding generation of universal linking codes (on line) as a well as possible recoveries (only off line) are done by this new party, allowing the central data collector to focus only on the collection of data. This additional party induces more complex protocols, hence the simplicity is partially lost while additional properties are gained. Their theoretical approach, while interesting, is not always adapted to the practical constraints inherent in the wide implementation of such a system, in particular in Switzerland.

In the light of the central question of this deliverable, namely “Is the identity knowledge in the government growing through the development of e-government” (cf. Chapter 1), the present approach is relevant since it tries to minimise the possibility even for the government to expand its knowledge about the citizen using non-identifying pseudonyms, and assuming protection of the secret key at the federal office.

7.6 Case study: international data collection for orthopedic evaluative research

A second example of an anonymous e-health database is the project Memdoc[99]. This project has been initiated by the Institute for Evaluative Research of the University of Bern[100] and is presented in [Roeder et al. 2004]. The goal of this project is to build very large registries of orthopedic cases. The pharmaceutical industry has well-known protocols for testing new medicines: first using animals, then healthy volunteers and finally sick volunteers. There exists no such possibility in orthopedic surgery. Once a hip has been implanted in a patient, we need to wait and see how long it resists inside a patient. The goal of a registry is to follow on a large scale (if possible for all the patients) a type of implants (for instance hip) in a country (or Europe-wide). The statistical results are used to measure the efficiency of a particular implant type or technique, such as to compare physicians or clinics regarding a benchmark (in order to reduce costs or increase efficiency).

The main difference between this case and the previous one is its international aspect. This has two consequences: Since nothing can be assumed regarding the IT infrastructure of the hospitals, data is transfered using a web interface. Moreover, since some European countries restrict the use and export of personal data, we can not have a central database containing all personal and medical data. We have a two-layer architecture, combining a central database and decentralized modules. There is only one central database for all the medical data, in order to have large datasets for statistics, but such a database has to be anonymous. The

[99] See the website http://www.memdoc.org.

[100] Institut für Evaluative Forschung der Universität Bern, http://www.iefo.unibe.ch.

project uses decentralized web servers called “modules” to anonymize the data. The central server collects all medical data, while the different modules are used to store personal data.

When a physician takes part in a registry, he has to submit his cases into the database. Each time he wants to report something, he connects to the module corresponding to his country using a web browser. He fills out a form containing the personal data of the patient, for instance names, date of birth, address or social security number. This information is stored on the module. Then the module creates on the server a new patient linked to the personal information using a random number. The physician fills out forms reporting medical information that is stored on the central server.

This architecture allows a good separation between medical data that can be used for statistical purposes and personal data that is only used to search among one's own patients, and is stored on the module. A patient belongs to one single clinic, this means that no physician outside this clinic can access any data of this person. If someone moves (or simply changes clinics), then he will have two records on the module corresponding to two different users on the server.

The goal of the application is to follow orthopedic implants for a long time. We can not let users be lost when simply changing clinics. We need a solution to track such “moving patients”.

- The first easy solution is to allow any physician having access to the module to see any patient. This solution can not be implemented for evident privacy reasons. Anybody could have access to the data of anyone else!
- The second possibility is to check at the creation of each new patient if a patient having a similar identity already exists in the module; then we would reuse the same patient. This solution is not convenient, because any physician could access all the files stored on the server by forging new empty patients to read data already entered by colleagues.
- Finally, the chosen solution relies on a one-way hash function (e.g., SHA-2) named h(x). For each module, we define a set of identifying attributes. It can be first and last names plus date of birth, it can also be first and maiden names plus date of birth or it could be the social security number of the patient; let us denote this information ID. Moreover, each module generates a unique key denoted by K. Then each patient receives a code generated by the module that corresponds to the result of the function: h(ID+K). This code is transferred to the central server and added to the record of the patient. This number is the same for this patient even if he changes clinics, etc. It is used for the statisticians on the central server, but can not be used by a physician to access the files of a colleague.

Such an architecture allows one to store personal data locally and to centralize all the medical data. A patient can be followed all his life. This will help orthopedic research without harming privacy of the patients. This means also that in the light of the central question of this deliverable, namely “Is the identity knowledge in the government growing through the development of e-government” (cf section 1), this approach guarantees especially that no identifying information is transferred between countries. The data collected in the central database does not help the government to get more information about an individual.

7.7 Conclusion

The medical sector requires more and more data for statistical purposes. The privacy requirements have to be fully integrated in any infrastructure dealing with e-health. But loyalty cards, credit cards or cellular phone operators collect large amounts of information about customers too. Such information can be used for accessing health-related information. For instance, Mrs. X has bought a pregnancy test; Mr. Z was in the left wing of the hospital at 10:00 am and had a scan two hours later, hence he probably has cancer. Such indirect health data have a similar level of confidentiality and should be treated as such by their owners. This risk can be excluded by systematically anonymizing data.

The approach described in this chapter shows how cryptographic methods and non-identifying pseudonyms can, in special situations, be used to guarantee to some extent that identifying information cannot be used for purposes other than originally intended, while still allowing for statistical information to be drawn from the whole set of data. This is a relevant finding for the purpose of this deliverable, since it shows that identification data need not necessarily be processed even when exploiting new opportunities of electronic data processing.

Governments can exploit the potential of ICT to improve service delivery for citizens in eGovernment, but it is not always necessary for governments to process and store identifying data. For certain purposes, the identity knowledge of governments need not necessarily be enlarged, which is particularly relevant if sensitive data – such as medical data – are concerned. The technical approach sketched in this chapter could therefore be used in and adapted to new fields of application in public service delivery, as a privacy-enhancing technology.

7.8 Bibliography

[Bauer 2000] Friedrich L. Bauer, *Decrypted Secrets*, Springer, 2000.

[Baumann et al. 2005] Rainer Baumann, Marcel Rieser, Reto Strobl, FrancoisWeissbaum, AndreaWeisskopf, *Privacy Preserving Data Collection*, Swiss Federal Section For Cryptography, 2005.

[BFS 1997] *Der Datenschutz in der Medizinischen Statistik / La protection des données dans la statistique médicale*, BFS Bundesamt für Statistik / Office fédéral de la statistique (CH), 1997.

[Hildebrandt & Gutwirth 2007] Mireille Hildebrandt and Serge Gutwirth (eds), *Identity in a Networked World*, Springer, to be published 2007.

[ISO 15408-2] ISO 15408-2. Text for ISO/IEC 1th WD 15408-2, *Information technology – Security techniques – Evaluation criteria for IT security – Part 2: Security functional requirements.* 2003 (http://www.commoncriteria.de).

[Jaquet-Chiffelle 2003] David-Olivier Jaquet-Chiffelle (ed.). *Impact of new technologies on privacy and data protection*. TILT 15, special issue, Berne University of Applied Sciences, School of Engineering and Information Technology, 2003 (download from http://www.vip.ch).

[Jaquet-Chiffelle & Jeanneret 2001] David-Olivier Jaquet-Chiffelle and Jean-Paul Jeanneret. *How to protect the rights of patients to medical secrecy in official statistics.* Information Security Bulletin, The International Journal for IT Security Professionals, vol. 6, nb. 8, pp. 41-43, 2001.

[Roeder et al. 2004] Christoph Roeder, Amer EL-Kerdi, S. Eggli, and Max Aebi. *A centralized total joint replacement registry using web-based technologies.* J Bone Joint Surg Am, vol.86-A, nb. 9, pp. 2077-2080,2004.

[Stallings 1998] William Stallings, *Cryptography and Network Security: Principles and Practice,* Prentice Hall, 1998.

[Ström 2005] Pär Ström. *Die Überwachungsmafia, Das gute Geschäft mit unseren Daten.* Carl Hanser Verlag, 2005.

Part III. Conclusion

8 Conclusion

8.1 Context, research question, and approach

Government needs information to govern. Particularly in direct relations with citizens, they need information from citizens in order to do their job, for example, decide upon a request, grant a permit, tell them to remove their illegally built shed, allow them to pass customs, etcetera. In many of these situations, governments want to know exactly which citizen they are dealing with, and hence, they require identification.

This is in certain cases, perhaps even fairly often, not strictly necessary, but understandable. As is illustrated by Jeremy Bentham's theories of social control and identification, identification of citizens is a mode of control that may benefit society at large. But for the obstacle of 'the state of public opinion nowadays', Bentham would suggest every citizen to have his name tattooed on his wrist, as an extreme measure for crime prevention and control.

Whereas Bentham's discourse largely focuses on criminal law and law and order, and hence the citizen is largely viewed in the context of crime and control, government today increasingly is also viewing citizens as 'customers' in the context of service provision. Public organisations are developing personalised ways of interacting with citizens, for example, for efficiency and convenience reasons. This interaction is also a driver of demands for identification and personal information. Technology is a prime facilitator of this process, as it allows speedy and massive personalised data exchange. The transition from a paper-based and brick-and-mortar economy to the information society goes hand in hand with increasing possibilities, in an ever wider variety and depth, of government-citizen interaction.

All this leads to the assumption that citizens are nowadays more and more identified by the government through electronic provision of public services. It is this assumption that is being tested in this deliverable:

> *Does the citizen become more known by the government when digital identification technologies are applied in the process of public service provision? In other words, is the identity knowledge of the government growing through the development of e-government?*

'Identity knowledge' is the collection of descriptive information that is connectable to an individual; it is a bipolar continuum with identifiability at one end and anonymity at the other. In our view, it is more interesting to consider identifiability-anonymity as a continuum than as an either/or issue. The five concentric circles of identity information that Marx has distinguished,[101] of core, unique, sensitive, private, and individual information, can be used to assess the degree to which someone's identity is known. The identity knowledge capacity of an organisation (i.e., the amount of data, the centralisation of those data, the speed of information flows, and the number of points of contact between organisation and subject)[102] is another concept that can be used when answering our question whether e-government leads to increasing identification of citizens.

[101] Marx, G.T. (2005). Varieties of Personal Information as Influences on Attitudes towards Surveillance. In K. D. Haggerty & R. V. Ericson (eds.), *The New Politics of Surveillance and Visibility.* Toronto: University of Toronto Press.

[102] Rule, J.B. (1973). *Private Lives and Public Surveillance.* New York: Schocken Books.

In this report, we have tried to give a first, tentative answer to this question by analysing various case studies. The case studies have been selected to illustrate the broad range of e-government and public services in Europe: different sectors (e.g., environmental planning, border control, health care), from small-scale (tree-felling permit) to large-scale (identity management in general), and across a range of technological mechanisms (password protection, biometrics, identity cards, cryptography), in different countries (Netherlands, Germany, Belgium, and Switzerland). Some cases are more descriptive of current, concrete processes (such as the Dutch cases), while others (such as the Belgian case) focus more on an analysis of more abstract problems in public service provision and how technology could be used to overcome these.

The scope of this study did not allow for a systematic, comparative approach; rather, we have used a heuristic combination of cases. This does not allow for definitive answers, but as mentioned, the objective of this report is to give a tentative first answer, which subsequent research can refine and adapt.

8.2 A minor increase in identity knowledge

A brief roll-call of the case studies provides the following conclusion:

- In the Netherlands, new technologies change the identification process, but this does not necessarily lead to more identity knowledge. One case (tree-felling permits) showed no shift on the anonymity-identifiability continuum, while the other case (Privium border control) shows that individuals are more identified than they are in traditional border passage. The identity knowledge capacity of the royal military police increases, albeit in a relatively minor way.
- In Germany, if a service requires an advanced electronic signature, the certificate shows more data than are often necessary (citizen registration number, date of birth). Pseudonymous certificates are allowed by the law, but half-heartedly, and particularly in governmental procedures, explicitly outlawed for use by citizens. As a result, the identity knowledge of the government grows somewhat through the transition from paper-based to electronic signatures.
- In Belgium, an identity management infrastructure for e-government is being created, with much attention to interoperability and efficient information management. The federal government has decided to use one common global identifier to identify all citizens across several contexts (the National Registry number). Although use of this number is regulated by law and subject to prior authorisation by a committee of the Privacy Commission, it is clear that the government has valued government-centred interoperability and efficiency over user-centred, anonymous or pseudonymous communication with citizens. This creates significant opportunities for increasing the knowledge capacity of the government in future.
- In Switzerland, the case study of medical statistical data showed that technology can facilitate non-identifiability in patients' records in central medical databases, while still allowing the same patient to be followed through different treatments in time or space. Linkability can be effected without identifiability, through cryptography-generated pseudonyms. Some caution is warranted, as the current practice of anonymisation may be extended with a recovery authority who has the power to effect identification if necessary.

Now, what do the case studies suggest in relation to our research question: does the citizen become more known by the government when digital identification technologies are applied in the process of public service provision? This can be answered in the affirmative, if only moderately so. The case studies by and large point in the same direction: the knowledge capacity of governments grows in the transition from paper-based to electronic communications, although currently only to a minor extent.[103] The identity knowledge increases with some more personal data of citizens, such as birth date or a photograph. There may be some cause for concern in this, since the overcollection as well as the move towards sharing of data is contrary to the data-limitation principle of data-protection law, but we should not exaggerate the threat to privacy that this relatively small increase in knowledge capacity poses. At the same time, we should welcome more – and more systematic – research in this area, in order to see whether our tentative conclusion can be affirmed if a wider range of e-government services across Europe are studied.

8.3 Identification infrastructures and the need for further research

The minor increase in identity knowledge that can be observed in several cases is only part of the answer, however. What also emerges from the case studies, is that e-government is very much in development,[104] and that identification infrastructures are slowly being built. Many of these centre on identifying numbers, and it seems significant that at least two of the countries studied here – the Netherlands and Belgium – have opted for single global identifying numbers, rather than sector-specific numbers.[105] The same tendency of citizen-identification desire is voiced in the German federal government's statement that pseudonyms may be used by the public administration but not by the citizen, in order to avoid "abuse of governmental services".[106]

These identification infrastructures with single identifiers facilitate the merging of data bases, data mining, and profiling, to an extent previously unknown.[107] Currently, these kinds of data processing are not happening on a wide scale in e-government, but often, technological possibilities gradually but surely tend to be exploited, and there is a substantial risk that privacy is gradually being eroded as a result.[108] The tendency of governments to call citizens 'customers' and to stress personalisation, and the German prohibition of pseudonymous communication in official government procedures, are signs of the times to come. Perhaps these signs will be contradicted by future developments, but this is at least another important field to research in the coming years.

What should be part of this future research, is to study technological and organisational means to counterbalance the increased identity knowledge of governments. Legislation, including data-protection legislation, will not be a primary tool to keep the knowledge capacity of

[103] The Swiss case is not a counter-example to this conclusion, since the collection of medical data has been enabled by ICT in the first place, and hence, anonymisation of this process does not lead to less identity knowledge than the central agency formerly had.

[104] Cf., Prins, J.E.J. (ed.) (2006), *Designing E-Government*, 2nd Edition, Kluwer Law International 2006.

[105] An overview of ID number policies in various European countries can be found in a forthcoming FIDIS deliverable, D13.3, which will be published on http://www.fidis.net towards the end of 2007.

[106] See *supra*, section 5.4.

[107] For an analysis of the risks of single identifiers, see FIDIS deliverable D13.3, mentioned *supra*, note 105.

[108] Koops, Bert-Jaap & Ronald Leenes (2005), '"Code" and the Slow Erosion of Privacy', *Michigan Telecommunications & Technology Law Review 12* (1), p. 115-188, http://www.mttlr.org/voltwelve/koops&leenes.pdf.

governments in check, since practice shows that the e-government developments in this context are backed up by legislation. A more useful direction to look for checks and balances is technology itself.

Several chapters in this report have stressed the potential of privacy-by-design, by fostering the development of (if necessary reversible) anonymisation techniques, of credentials rather than identifiers, and of smart pseudonym systems, for example. The vision outlined in chapter 6 of privacy-friendly identity management in e-government, using the potential of technology not only to increase knowledge capacity of governments but also to enhance privacy of citizens, is worth elaborating.[109]

8.4 Bibliography

Koops, Bert-Jaap & Ronald Leenes (2005), '"Code" and the Slow Erosion of Privacy', *Michigan Telecommunications & Technology Law Review 12* (1), p. 115-188, http://www.mttlr.org/voltwelve/koops&leenes.pdf.

Marx, G.T. (2005). Varieties of Personal Information as Influences on Attitudes towards Surveillance. In K. D. Haggerty & R. V. Ericson (eds.), *The New Politics of Surveillance and Visibility.* Toronto: University of Toronto Press.

Prins, J.E.J. (ed.) (2006), *Designing E-Government*, 2nd Edition, Kluwer Law International 2006.

Rule, J.B. (1973). *Private Lives and Public Surveillance.* New York: Schocken Books.

[109] Which will be done in FIDIS deliverables D16.1 and further, to be published on http://www.fidis.net towards in 2008.