National Identification Systems

by

Thiên-Lôc Nguyên

Ingénieur de l'Ecole Polytechnique, 2000

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2003

Author .....
Department of Electrical Engineering and Computer Science
March 24, 2003

Certified by .....
Ronald L. Rivest
Viterbi Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

National Identification Systems

by

Thiên-Lôc Nguyên

Submitted to the Department of Electrical Engineering and Computer Science
on March 24, 2003, in partial fulfillment of the
requirements for the degree of
Master of Science in Computer Science and Engineering

Abstract

In this thesis, we investigate the intricate question of technology-enabled national identification systems, from a scientific perspective. We first present a framework to formalize the notion of identification systems in general, before delving into the policy issues and technological challenges specific to national identification systems. With this thesis, we wish to present to non-specialists the different possibilities, often too little known to people outside the research community, enabled by the current knowledge in science and technology (especially cryptology and security), and to lay the foundations for future scientific research on the subject. We hope that this presentation contributes to widen the general public's view on the matter of national identification systems, in particular with respect to the possibility of novel architectures far from those of existing systems.

This thesis assumes no prior technical knowledge in the field of cryptology and is intended to be accessible to any person of college-level education with an interest in security and privacy issues.

Thesis Supervisor: Ronald L. Rivest
Title: Viterbi Professor of Electrical Engineering and Computer Science

Acknowledgments

First and foremost, I would like to express my extreme gratitude to my advisor, Ronald Rivest, for suggesting me this most interesting research topic, as well as for all his assistance and support in producing this work. I also wish to thank Anna Lysyanskaya for her valuable comments and insight about her work on pseudonym systems and privacy-enhancing cryptology in general.

I am very grateful as well to Marilyn Pierce, who makes a great job of taking care of all the graduate students in EECS. Also, this acknowledgements section wouldn't be complete without thanking Be Blackburn, our devoted administrative assistant for the theory group, who is always there to cater to her beloved theory students.

As work has only been part of my experience at MIT, and in Boston in general, I would like to thank all my friends here – old and new – whose friendship and support contributed to make my stay here even more enjoyable. In particular, I am especially thankful to Joan Decker and my roommates Richard Poutrel and Florent Ségonne for their invariable loyalty and moral support throughout my stay, and for making me feel really at home here.

Last but not least, I would like to thank my parents for ... well ... just about everything.

Contents

Chapter 1

Introduction

Following the realization of the threats relating to international terrorism, there has been a revival of interest in national identification systems. Many countries are considering replacing their existing national ID cards with new – technology-enabled – ones, supposedly more secure. Others, that do not currently have a national identification system in use, are debating whether they should adopt such a system at all.

Nowadays, technology has become more and more affordable and pervasive. Technology-based security is present in our daily life through credit card payments, magnetic badges to enter company buildings, or the ever-increasing use of online transactions (e-commerce purchases, online auctions, online banking, etc). As more and more companies are relying on technology to protect their critical physical and information assets, it is natural to wonder to what extent a possible national identification system might benefit from this tried and tested knowledge and experience.

As a matter of public interest, a national identification system ought to be the subject of public consultation and debate. The purposes and needs for such a system should be analyzed carefully. In particular, we do not restrict our analysis to card-based systems using national ID cards. Our study will also address the legitimate concerns of many civil liberties organizations, regarding privacy especially. Yet, throughout this thesis, special consideration will be given to issues more particularly relevant to the United States.

The role of the scientist in this affair is not to individually design the best system – technically speaking – but rather, to inform the public about current state-of-the-art technology (what is presently possible, what is still being researched, and what is most

likely impractical or definitely impossible), assess the possible needs and constraints of such a system, and then eventually propose different alternatives that would be debated publicly before possible adoption. To that end, a good system needs to be flexible and customizable; particularly, it should be possible to set certain tradeoffs independently of the system design to meet the requirements and needs of different countries, or to smoothly upgrade the system to include new features or fulfill new needs without disrupting the normal functioning of the system.

In chapter 2 we briefly survey existing literature on national identification systems, and present our approach to the problem. We analyze the possible purposes of identification systems in general in chapter 3, and introduce a framework for analyzing the problem from a technical perspective. We also introduce some new terminology to reflect what we think are the essential concepts to think about. In chapter 4 we focus on the specifics of national identification systems, and examine the main issues to be tackled and the major challenges to overcome. Chapter 5 surveys many existing applications related to our problem and shows how some successfully achieve many of the desired goals of national identification systems, and to what extent others fall short of their objectives on certain aspects. In chapter 6 we present current state-of-the-art knowledge in technology – and more specifically cryptology – and try to give the reader more insight into the different current technical possibilities and how we could apply them to the specific problems of national identification systems. We conclude in chapter 7 by explaining to what extent a national identification scheme may address the major concerns and reservations expressed by the public opinion, in particular regarding privacy.

1.1 Disclaimer

Scope and goal of this thesis

We wish to emphasize here that our goal in this thesis is not to give a comprehensive study of national identification systems. Rather, we are introducing a framework to think about these systems in a more general way than what may have been done so far. In particular, we aim at changing the way people traditionally envision these systems, by presenting illustrative examples of different original scenarios or designs that are little considered. Nonetheless, we tried in this thesis to cover the essential issues, regarding both

policy and technology, and to present the relevant technology, especially the recent advances in privacy-enhancing cryptology.

Personal opinion

This thesis is the outcome of a very fruitful work with my advisor, Prof. Ronald L. Rivest. Yet, although he has fully reviewed the material presented here, he may disagree on some aspects of it. Therefore, this thesis represents my own vision on national identification systems (and in some parts, my personal opinion), and the reader should not consider it to reflect Prof. Rivest's own opinion. In addition, my personal opinions are certainly influenced by my background as a Frenchman, and thus may differ, for example, from the typical American point of view.

Chapter 2

Background

National identification systems – most specifically national ID card systems – have been much studied recently and the current attention on the topic has stimulated further debate and analysis, with much of the concern being on whether they can help fight terrorism and enhance national security. Also, many people are especially worried about the possible loss of privacy a national identification system could bring.

2.1 Related work

Numerous articles from newspapers or magazines have expressed personal opinions and feelings about national ID cards. For instance, the Electronic Privacy Information Center (EPIC) lists a number of relevant news articles on the subject since September 2001 [82]. It is interesting to notice that most authors tend to oppose the very idea of a national identification system, based on sensible arguments, mainly about privacy issues.

Among the few proponents of national ID card systems are people who may have interest in their deployment, such as Larry Ellison, CEO of Oracle, who offered to donate the whole infrastructure to the US government for free (but the maintenance and upgrades will not be free) [65, 124].

A good many public or non-profit organizations have been reflecting on the subject:

- The Electronic Privacy Information Center (EPIC) is a "public interest research center in Washington, D.C. It was established in 1994 to focus public attention on emerging civil liberties issues and to protect privacy, the First Amendment, and constitutional

values"[1]. EPIC maintains a webpage [82] referencing an extensive list of news articles, web resources and reports on national IDs and related issues. Also, EPIC produced a report [85] in February 2002 assessing the current project of the American Association of Motor Vehicle Administrators (AAMVA) to integrate the different state (resp. province) driver's licence systems into a national identification system in the United States (resp. Canada). This project and report will be examined in section 5.2.

- "Privacy International (PI) is a human rights group formed in 1990 as a watchdog on surveillance by governments and corporations"[2]. PI maintains its own webpage [108] on the subject of national ID cards, containing also original work from the organization. In particular, its ID Card FAQ (Frequently Asked Questions) section [110] (August 1996) offers an excellent introduction to the topic, addressing concisely most of the main issues raised in that matter. More recently, on July 13, 2002, in response to UK's Home Office's consultation paper on an "entitlement card" on July 3, 2002, Privacy International created a new section on its website [109] devoted entirely to this affair.

- The CATO Institute "is a non-profit public policy research foundation"[3]. It analyzed the matter of a national ID system as a solution to illegal immigration [136] in September 1995.

- Computer Professionals for Social Responsibility (CPSR) "is a public-interest alliance of computer scientists and others concerned about the impact of computer technology on society"[4]. It issued a FAQ section [91] in September 2001 investigating the use of National Identification Schemes (NIDS) for fighting against terrorism.

- The American Civil Liberties Union (ACLU) is a non-profit organization with nearly 300,000 members and supporters, which "mission is to fight civil liberties violations wherever and whenever they occur"[5]. It explained in 1996 concisely why it strongly opposes National Identification Cards [155].

- More recently, the Computer Science and Telecommunications Board from the US

[1]Description taken from EPIC's website: http://www.epic.org/epic/about.html
[2]Description taken from PI's website: http://www.privacyinternational.org/
[3]Description taken from CATO's website: http://www.cato.org/about/about.html
[4]Description taken from CPSR's website: http://www.cpsr.org/cpsr/about-cpsr.html
[5]Description taken from ACLU's website: http://www.aclu.org/about/aboutmain.cfm

National Research Council has issued a report on nationwide identity systems [113] in April 2002. This excellent study investigates the essential policy questions that need to be addressed, along with some of the technological challenges that need to be overcome. This report has significantly fuelled our reflection on national identification systems, especially as regards policy issues.

National governments in many countries have also initiated various investigations about the desirability (and sometimes established project committees to start working on actual proposals) of a national identification system. Some of them already deployed such a system in response to the tragic events of September 11, 2001.

- The Privacy Commissioner of Canada issued a report [96] to the Canadian Parliament in January 2003 on Canada's *Privacy Act* and *Personal Information Protection and Electronic Documents Act.* In particular, he strongly opposes recent initiatives mounted by the Government of Canada in its fight against terrorism, which represent severe threats to the people's privacy rights. In his analysis, he pays special attention to the issue of national ID cards.

- The adoption by Japan in August 2002 of an ID system with an ID number for each citizen was highly controversial [64, 111]. In particular, the data privacy and protection legislation that was to accompany the deployment of the system has not yet been enacted.

- UK's Home Office issued a proposal in July 2002 for an entitlement card, and expected public comments by January 31, 2003. Although UK's government is publicly claiming that there is majority support for ID cards, the proposal has faced strong opposition. We refer the reader to the already mentioned Privacy International's webpage on the subject [109] for a comprehensive treatment of this affair.

Finally, we wish to mention here Simson Garfinkel's book *Database Nation* [95], which provides an excellent treatment of the privacy implications of ongoing technological developments in our modern society, and especially those related to computer databases. This book is a landmark in the campaign for public awareness of the privacy risks brought by technology.

2.2 Our approach

We contribute to the already flourishing discussion on the subject by providing a rational analysis of national identification systems, including both policy and technical issues, using a novel approach. CSTB's report [113] investigated the essential policy questions that need to be addressed, along with some of the technological challenges that need to be overcome. However, unlike CSTB's policy-oriented report which focuses mainly on raising the issues, we adopt a more scientific approach[6]:

- We provide a framework to analyze identification systems in general from a technical perspective, and introduce appropriate terminology as needed. We hope this framework can contribute to lay out the foundations for further scientific research on the subject.

- We try to emphasize the different non-traditional forms a national identification scheme could take (cf section 1.1): a national identification system without actual national ID card, a group of independent systems working together as a national identification scheme but fulfilling separate purposes, an "anonymous" identification process revealing only selected information at the discretion of the individual, etc.

- We analyze the possible purposes of a national identification scheme, as well as the probable motivations, needs, interests and concerns of the different parties concerned. In particular, we break up the discussion into the two main aspects of the problem: the actual identification of the people, and the applications that will use (and possibly misuse) this identification feature.

- We strongly differentiate between the issues that can be effectively addressed by technology (and give elements of solutions or research directions for these questions) and those that need to be settled by policy.

With this thesis, we aim at bridging the gap between the policy community and the scientific research community. Many policy experts are unaware of the currently available

[6]Also, unlike CSTB's report, which investigates *nationwide* identity systems, we focus on *national* identification systems: although we do not consider only a government-run system, we are interested in a system overseen by the government – and possibly designed using government guidelines – that would have its sovereign endorsement. We will explain more precisely our position and reasons in section 4.1.

technology (and may therefore have a biased assessment of the problem), while many scientific experts may be too little aware of how their remarkable results may be used for the design of national identification systems.

We hope that this thesis will foster the discussion on national identification systems and contribute to focus the debate on the key issues, while providing the general public with a better insight into today's technology landscape. We also wish to convince the reader that a well-designed national identification scheme, while not solving all the problems, could be an improvement over the present situation. We believe however that haste should be avoided in trying to expedite the deployment of such a scheme to face the new security risks related to international terrorism. On the contrary, we deem worthy to further study the social implications of the adoption of a national identification scheme, in order to find a solution (or an alternative) to the problem before a poor de facto such scheme is used – and misused – widely.

It is outside our province to take side in favor or against the actual adoption and deployment of a national identification scheme. Our goal is to help the policymakers and stakeholders make an informed decision on the subject by offering them a better picture of the problem – especially the new possibilities enabled by recent research advances, too often overlooked – and a better assessment of the critical tradeoffs they have to consider when deciding whether to adopt a national identification scheme for their country.

Chapter 3

Purposes and structure of identification systems

In this chapter, we explore the possible purposes of an identification system, and present the framework in which we will later analyze the different issues and challenges to be addressed by a national identification scheme. We also introduce some terminology we will use throughout the thesis. Finally, we explain how the use of cryptology can substantially change some aspects of identification systems, and shed new light on the way we traditionally look at such systems. We will focus on the general aspects of an identification system in this chapter, and will address the specifics of national identification systems only in subsequent chapters.

3.1 The notion of identity

Prior to discussing the different issues involved in an identification system, let us first address the very notion of identity itself. For reference, the Oxford English Dictionary[1] gives the following definition of identity:

> *" The sameness of a person or thing at all times or in all circumstances; the condition or fact that a person or thing is itself and not something else; individuality, personality. "*[2]

[1] We used here the online version of the Oxford English Dictionary available to the MIT community at http://libproxy.mit.edu:8181/entrance.dtl.

[2] Definition taken from the Oxford English Dictionary: identity, 2.a.

Starting from this common notion of identity, we will clarify in this section what we mean by identity in our discussion of identification systems, as well as define clearly how we interpret other related notions.

3.1.1 Natural persons and artificial persons

In many countries, the laws make a fundamental distinction between a natural person and an artificial person. Following is the definition[3] given on the Natural-Person non-profit website [22]:

> " Here are the exact definitions from Barron's Canadian Law Dictionary, third edition:
>
> - natural person. A natural person is a human being that has the capacity for rights and duties.
> - artificial person. A legal entity, not a human being, recognized as a person in law to whom certain legal rights and duties may attached – e.g. a body corporate. "

In many applications, one may want to "identify" an artificial person: when signing contracts on behalf of an organization, cashing checks for a company, etc. The identification of artificial persons is definitely an interesting problem, with many useful applications. However, we will focus in this thesis on the identification of natural persons, i.e. human beings. Also, although the identification of computer machines or processes may seem out of the scope of a discussion on identification of human beings, we will explain in section 3.1.6 why this matter is of interest to us.

3.1.2 Persons, bodies and identities

When one wants to describe a (natural) person, one can do so by giving characteristics of the following nature:

- physical. These characteristics pertain to the person's physical body: gender, height, eye color, etc.

[3] Although this definition is only fully applicable to the laws of Canada, the same notions are present in the laws of the United States, and many other countries.

- character. These characteristics pertain to the nature of the person: honesty, integrity, loyalty, courage, etc.

- social/legal. These characteristics pertain to the person as an active member of human society: name, citizenship, residence, etc.

One person, one body

As the assessment of somebody's character is highly subjective, we will focus on the objective characteristics of the two other categories. We thus define the following notions:

Definition 3.1

- The *person* is defined by the Oxford English Dictionary as "an individual human being: a man, woman or child"[4]. Without delving further into any philosophical discussion, we define the person to be characterized by his/her mind: he/she is the one who thinks and decides upon his/her actions.

- The *body* is "the physical or material frame or structure of man [...]: the whole material organism viewed as an organic entity"[5].

Until fundamental discovery about the human being or revolutionary technological invention that would enable people to have possibly many bodies[6], we will assume for the time being (and our discussion here) that each person has one unique body, and that conversely each body is inhabited by a unique person.

One person, but possibly many identities

Definition 3.2

- A *person*'s *identity* is a set of characteristics – physical and social/legal – that fully describe and characterize that person as an active member of human society, and differentiate him/her from the rest of the population.

[4]Definition taken from the Oxford English Dictionary: person, *n.*, II.2.a.

[5]Definition taken from the Oxford English Dictionary: body, *n.*, I.1.a.

[6]According to Kurzweil [114], we may be able in the soon future to capture a comprehensive map of the human mind within a computer, and then possibly "download" our mind into our favorite computer. This might be in the realm of the possible, yet remains presently highly speculative.

Most persons live their lives happily and peacefully by relating to others with the same identity for their entire existence.

However, there is a legitimate need for enabling persons to have different identities. In some circumstances, a person may need to start a new life afresh by becoming a totally new individual. In the case of key witnesses to critical trials[7] for instance, one may actually want to "bury" the old identity, and the person will be known through the new identity for ever since the change. In other cases, as for undercover agents, one may want a person to be able to take on both identities, according to his/her choice, while having these identities unlinkable.

Depending on the identification system considered, one may allow persons to have multiple identities in the system, or actually prevent it. The ability to possess multiple identities could be regulated, and also limited to exceptional cases.

Conversely however, by definition, only one unique person can assume a given identity.

3.1.3 The need for identity and identification

Where does the need for identity come from? Why would a person want to take on a given identity? In many situations, one can conduct transactions with other people without being identified: when buying a newspaper, dining at a restaurant, etc. In fact, many "one-time" transactions can be performed anonymously. However, the need for identity (and identification) arises as soon as a person has a recurring interaction with another party: along with his/her identity, a person carries a history.

In many situations, assuming a given identity entitles a person to benefit from privileges granted to that identity (cashing a paycheck, entering an office building, boarding a plane, etc). Yet, in most cases, a full identification is not needed: to benefit from a due privilege or right, a person often only needs to prove his/her membership of a group or that he/she satisfies some requirements: to buy tobacco or alcohol for instance, all he/she needs to prove is that he/she meets the legal age requirements.

In our fast-paced modern society, we have to deal more and more often with people we haven't met, and sometimes will never meet in person. Being able to identify the other party with strong confidence has nowadays become a precondition to many interactions in

[7]Information on the United States' federal Witness Security Program can be found at http://www.usdoj.gov/marshals/witsec.html

our daily life.

3.1.4 Identities, profiles and names

To reflect the different forms of identification mentioned in the previous section, we introduce the following notions:

Definition 3.3

- A *person*'s *identity* is a set of characteristics – physical and social/legal – that fully describe and characterize that person as an active member of human society, and differentiate him/her from the rest of the population. This is a reminder of the definition introduced earlier.
- A *profile* of a *person's identity* is a set of properties about the person's identity. We also define the notion of profile of a person, when the person's identity is clear from the context.

How does a profile relate to the identity?

As we have seen in the previous section (3.1.3), identification is often used to prove membership of a group or fulfillment of certain requirements. For that purpose, a person need not reveal his/her full identity. Rather he/she only needs to reveal part of it, which constitutes a profile. We show in figure 3-1 an example of identity and profiles.

We wish to emphasize here the fact that, by definition, the profile is not just a subset of the characteristics composing the person's identity, but rather a set of properties about these characteristics. Let us consider for instance the following example: the purchase of alcohol in the United States. All the store clerk needs to know in order to sell you a bottle of wine is that you are over 21, but he/she does not need to know your exact age, and even less your date of birth. Therefore the profile you need to disclose need not contain any age or birth related characteristics of your identity, but only the property that you are indeed over 21.

When taking on a given identity, a person can yet have multiple profiles corresponding to that identity (the citizen who votes, the employee who works at MyCompany, the person who owns a bank account at MyBank, etc); the same person/identity is actually known

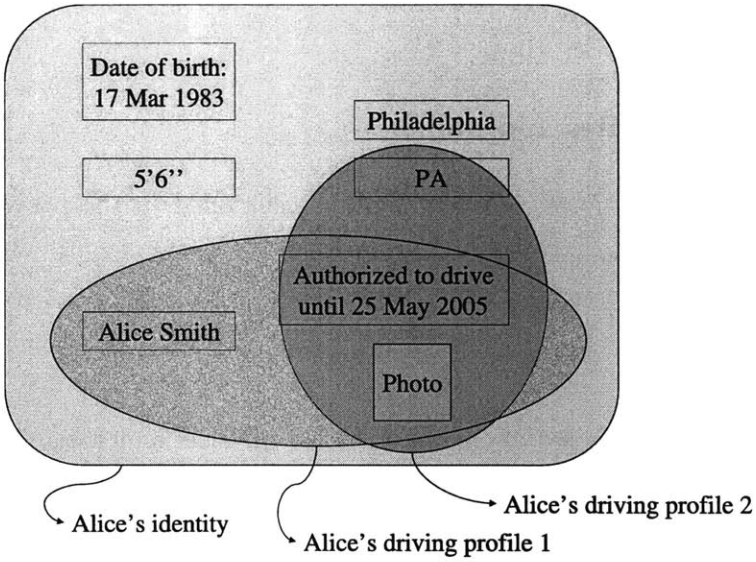


Figure 3-1: Identity and profiles

to different organizations through different profiles. These organizations need not actually know the person's full identity, but only certain pieces of information about him/her. In fact, most organizations will never know the person's full identity, but only a profile about him/her, containing – in principle – only the information that is relevant to their businesses.

What is a profile, really?

The profile actually encompasses many different notions:

- A profile can describe the role or status of a person: it can "say" that a given person is the current President of the United States, or recipient of a past Nobel Prize, etc.
- A profile can state the membership of a person in a group: it can "say" that the person is a male, is a member of the government, or an adult age-wise (or in other words is over 18 of age), etc.
- A profile can also describe the fulfillment of a certain number of criteria: it can "say" that the person has a valid driver's license and has not committed any major traffic violation so far.
- Finally, a profile can represent the personal dossier an organization has about an individual, containing personal information about him/her (which usually only captures part of his/her full identity).

Names and identifiers

Historically the need for identification has been mainly social. Men and women have been using names to identity their fellows. In the early ages, men only needed identification within their limited communities: the use of one-word names was then largely sufficient. With the growth of human communities, the development of travel and trade between different populations, the need for more complex names appeared. Nowadays, people in Western countries usually have a first name, one or many middle names, and a last name. Many people even have nicknames, which they may prefer to their actual names. The means of identification have been refined over the centuries to now combine a variety of characteristics: name, photo, date of birth, social security number, etc.

We wish to emphasize here the following point to prevent any possible misconception. Although identifying a person may have historically translated to determining his/her name, this is not true any more in our modern society. Nowadays, a person's name need not uniquely determine his/her identity: many people today have names that are identical or homonymous. Determining a person's name is actually neither required nor sufficient for identifying that person: while a person's name won't always determine his/her identity, the latter may be determined without knowledge of his/her name (in biometrics identification for instance).

A person's name is by no means an identifier leading to his/her identity.

3.1.5 The cryptographic notion of identity

Since the invention of public-key cryptography, the notion of identity in cryptography has traditionally been equivalent to the possession of some cryptographic key. In a public-key infrastructure, each party – human person, organization, computer machine, etc – possesses a pair of associated keys, one public and one secret: the public key is known to all parties while the secret key remains the private property of the key holder.

The possession of the secret key enables the key holder to perform various cryptographic actions nobody else can do without the knowledge of the key. For instance, using her secret key, Alice can digitally sign a document so that anybody can verify the authenticity of this signature using Alice's public key, while nobody can forge Alice's signature without knowing her secret key. Alice's key pair therefore identifies Alice. More precisely, in one

traditional approach, Alice's public key represents her "digital identity" (it is actually a unique identifier leading to her digital identity) while her secret key enables her to prove it.

Public-key cryptography will be presented in more detail in section 6.2.1, while the various variants of identity used in the cryptography field will be exposed in section 6.3.2.

3.1.6 The identity of machines

While the cryptographic notion of identity presented in the previous section (3.1.5) does not capture entirely the notion of human identity, it is well-suited for computer machines or agents. In fact, various schemes aimed at identifying computer systems are currently deployed (cf section 5.7).

This identification of computer systems could actually be leveraged in the near future for an identification system of human beings. One can indeed imagine each person having a computer agent that he/she trusts perform his/her identification: provided that the access to the computer agent is securely restricted to its owner, a successful identification (and authentication) of the agent implies the identification (and authentication) of its owner, who initiated the identification process. This computer agent could be running on a personal handheld device such as a cell phone or Personal Digital Assistant (PDA) for instance.

3.2 The notion of identification system

We describe in this section the purpose of an identification system: an identification system enables the storage and update of personal information for future revelation, through the use of digital identities. An identification system uses one or many databases to store the information, and can be run as a stand-alone system or together with one or more identification systems to form a more complex identification scheme. We also categorize the information contained in the system according to its function, and introduce the essential notion of credential. Finally, we distinguish between voluntary and involuntary identification systems, the former being our focus in this thesis.

3.2.1 The purpose of an identification system

In the various scenarios presented in the previous section, the need for identification comes down to the following: learning more information about a targeted person. This

information learning process actually breaks down into two steps: information about the person is remembered for future use, and information about the person remembered from the past is revealed. For our personal use in social interactions, these procedures are performed subconsciously by our brain and memory. When building an identification system however, the personal information needs to be stored physically so that it can be retrieved later.

An identification system should therefore implement the following functionalities:

- Registration: for each targeted person, one or more dossiers are created in the system to record personal information.
- Information storage/update: each time information about a person needs to be remembered for future use, it is stored in the system in one of his/her dossiers, and possibly updates previously stored information.
- Information revelation: each time information about a person is sought, it is retrieved from one of his/her dossiers.

Yet, what differentiates an identification system from a regular information system storing personal information is its ability to *identify* the targeted person while achieving the aforementioned functionalities. Rather than just containing descriptive personal information, a person's dossier will contain information enabling the identification (and authentication) of the person by the system: the dossier is really a *digital identity* of the person. We wish to note here that the identification-authentication performed by the identification system could be really strict (using state-of-the-art biometrics for instance), or else rather loose (such as the US Social Security Number system for instance), yet is a fundamental component of identification systems.

3.2.2 Digital identity

Each person who participates in the identification system will have his/her personal information stored somehow somewhere in the system. All the data relative to a given person in the system – data that may or may not be accurate – form a digital entity that is intended to truthfully represent the person's actual identity. However, there may be some differences between the actual information about the person and the recorded information stored in the system. This motivates the introduction of the following notions:

Definition 3.4

- The *digital identity* of a *person* is the part of that person's identity that it is recorded in the identification system. Note that the digital identity need not capture the person's "full" identity.

- A *digital profile* of a *person* is a profile of that person, as it is recorded in the identification system. We wish to insist here on the fact that a digital profile need not be a static set of data present in the identification system: it can be created dynamically when actually needed, based on the personal information contained in the digital identity.

When the person is registered for the identification system, a digital identity is created in the system to capture (part of) one of his/her real identities. While the digital identity need not be a digital representation of the "full" real identity (in other words, it can capture only certain aspects of this real identity), it ought to contain accurate information.

We will discuss the problems specific to the inaccuracies in the personal information stored in the identification system in section 3.6.1.

3.2.3 Identification scheme, identification system and databases

For numerous reasons, one may want the personal information relative to a targeted person to be distributed throughout the system. This may come from the desire to maintain multiple (unlinkable) identities (cf section 3.1.2), from the desire to categorize the information by nature (medical, financial, etc) or for security reasons. This separation of the information could be achieved through the use of different physical supports, or even different physical systems.

For instance, the personal records of a driver's license identification system run by a state Department of Motor Vehicles (DMV) in the United States are likely to be stored in several different databases. Though running independently of each other, these state identification systems, when taken together, form a "national" driver's license identification scheme.

In summary, an identification scheme provides for the aforementioned identification functionalities through the use of one or many independent physical identification systems,

which in turn could use one or more databases. We define these notions more precisely as follows:

Definition 3.5

- An *identification scheme* is a formal system providing for the identification functionalities: registration, information storage/update and information revelation. An identification scheme can be implemented as a single identification system or a group of identification systems.

- An *identification system* is a physical (paper, microfilm, computer, etc) system which can be run as a stand-alone system and implements the identification functionalities: registration, information storage/update and information revelation. An identification system can be used together with one or many other identification systems to meet the needs of an identification scheme.

- A *database* is a physical support for the storage of the personal information contained in an identification system. A database could be a library of paper files, a computer database, or a (smart) card carried by the person. A database does not provide any identification functionality, but only the physical means for information storage.

3.2.4 The nature of the information contained in the digital identity

The information contained in a person's digital identity can be broken down into the following categories, according to the function it serves:

- An authentication template. This template contains the information necessary to authenticate the person to the system, and determine his/her identity.

- Pointers to other information systems. These pointers enable the retrieval of personal information relative to the person's digital identity recorded in other information systems (which can be identification systems as well).

- Personally identifiable information. This is the additional personal information contained in the digital identity.

3.2.5 Credentials

Depending on which database the personal information is stored, one may decide whether he/she trusts the information or not. For instance, a piece of information stored in a government database is likely to be more trustworthy than one contained in a personal database. There is therefore a need in the identification system to have some personal information certified by some appropriate authority. David Chaum [52] introduced the notion of credentials in 1985 to illustrate this:

Definition 3.6

- David Chaum defines *credentials* as "statements based on an individual's relationship with organizations that are, in general, provided to other organizations." [52]. For our discussion here, a *credential* is more generally a document certified by some credential authority.

A very common example of credential in an identification system is the ID: whether it is a passport, a driver's license or some other identification card, an ID contains usually the essential personal information about the person, and is certified by the ID issuer.

We will see later in section 3.9.1 the various purposes a credential may serve.

3.2.6 Voluntary identification system vs. involuntary identification system

An important question to address when studying identification systems is whether the functionalities (registration, information storage/update, information revelation) are performed with the voluntary participation of the person:

Definition 3.7

- A *voluntary identification system* is an identification system involving the voluntary participation of the person in the information revelation functionality.
- An *involuntary identification system* is an identification system that does not involve the voluntary participation of the person in the information revelation functionality, or in other words, does not require his/her consent.

Most identification systems are voluntary identification systems. To travel abroad, a person needs to show his/her passport to the immigration and customs services of the foreign country. To purchase alcohol or tobacco, a person needs to provide a proof that he/she meets the legal age requirements.

Some identification systems however run without the consent of the person. In fact, they may even exist without his/her consent, approval or even knowledge. The quintessence of such identification systems is the "surveillance system". Every casino for instance runs an extensive (and expensive) surveillance system to maintain its security and prevent any fraud; for that latter purpose, it tries to identity known cheaters among its clientele. Also, there has been some controversy about the security measures used at Super Bowl XXXV in 2002 [123]: the authorities set up a face-recognition system to "identify" people susceptible of threatening the security of the event.

While involuntary identification systems pose some interesting challenges, especially determining a person's identity without his/her active participation, we will focus in this thesis on voluntary identification systems. While not addressing the issues specific to an involuntary identification system, most of our definitions and discussions are still applicable to such a system.

3.3 The notion of identification

What is identification exactly? The Oxford English Dictionary gives the following definition of identification:

> *" The determination of identity; the action or process of determining what a thing is; the recognition of a thing as being what it is. "*[8]

While identifying a person is essentially differentiating him/her from his/her fellows, the actual notion of identification is somehow ambiguous and has been used to designate many different (though related) notions. We will clarify here the sense we give to the term "identification" and to other related notions.

[8]Definition taken from the Oxford English Dictionary: identification, 2.

3.3.1 Identification vs. authentication

The term "identification" is equivocal: it sometimes refers to the whole identification process, revealing some personal information about a targeted person – at the end of which a certain right or privilege may be granted – while at other times it really stands for (only) the determination of the identity of the targeted person. When talking about identification, one needs to make the following distinction:

Definition 3.8

- *Identification* consists in determining the identity of a person from within a given population of possible matches.
- *Authentication* consists in verifying an identity either determined by identification, or claimed by a person.

Henceforward, we will use the terms "identification" in the aforementioned sense of identity determination, and "identification process" to designate the whole process.

3.3.2 Partial identification vs. total identification

As seen above in section 3.1.3, we may distinguish two main forms of identification processes in an identification system, according to the amount of personal information revealed in the process:

Definition 3.9

- A *partial identification process* only reveals part of a person's (digital) identity, or in other words, only reveals a (digital) profile of the person.
- A *total identification process* reveals the whole (digital) identity of the person.

As a person's digital profile could be his/her full digital identity, a total identification process is no more than a special case of partial identification process, where the digital profile revealed is the full digital identity. We will therefore focus hereafter on the harder problem of partial identification process.

3.4 The different parties involved

Before reviewing in more detail the different functionalities of the identification system (registration, information storage/update, information revelation), we define in this section the parties involved in these functionalities. We present here the roles and purposes of these (generic) parties in an identification system in general, delaying to section 4.2 the analysis of the interests of the real-world parties and stakeholders involved in a national identification scheme.

3.4.1 Overview of the parties by functionality

The registration involves the following two parties:

- The *person* who has a digital identity created in the identification system.
- The *identity authority* who carries out the actual creation of the *person*'s digital identity in the identification system.

For the information storage/update, the parties involved are:

- The *registered person* who has some personal information stored/updated, and attached to one of his/her digital identities.
- The *identity authority* who carries out the actual information storage/update of the *registered person*'s digital identity in the identification system.
- One or many *information authorities* who certify the authenticity of the personal information to be stored/updated.

The information revelation however involves four different parties:

- The *registered person* who has a digital profile – corresponding to one of his/her digital identities – revealed to an *examiner*.
- The *examiner* who obtains a digital profile of a *registered person*.
- The *identity authority* who provides for the technical means of this digital profile revelation.
- One or many *information authorities* who certify the authenticity of the information contained in the digital profile of the *registered person* that is revealed to the *examiner*.

3.4.2 The person/registered person

The person is at the heart of the identification system: the very purpose of an identification system is to provide reliable personal information about the person.

Yet, the person need not be a passive participant in the identification system: he/she could indeed be play a dynamic role in all the identification functionalities (registration, information storage/update, information revelation), and control the conditions and specifics of these processes. He/she could for instance decide when and how to update his/her personal information, what information he/she discloses at the information revelation stage, etc. We will explain in chapter 6 how technology enables for very flexible information storage/update and information revelation functionalities, preserving many an aspect of his/her privacy.

3.4.3 The examiner

An identification system contains personal information, which is inherently sensitive. Being able to obtain some personal information about other persons obviously needs to be regulated carefully. Since the very purpose of the identification process is for an examiner to learn some personal information about the registered person, the system needs to enforce an access control policy. In fact, it may be desirable to define different categories of examiners, which would have different privileges as regards the access to personal information: for instance, police and law enforcement officers would be able to request driving profiles, while liquor store clerks could only request age profiles.

When deciding on the levels of authorization to give to different categories of examiners, one needs to keep in mind that the information obtained by the examiner through the identification system could be combined with information from other systems, especially those privately run by the examiner. Therefore, not only the scope of the information obtainable by an examiner needs to be considered, but also the possible use of this information needs to be regulated carefully.

3.4.4 The identity authority

The identity authority is responsible for supervising the establishment and functioning of the identification system. As such, it needs to be trusted by all other parties for its ability

to maintain the security of the system, as well as its integrity in performing its function of administering the person's digital identities.

In many common applications, the identity authority is also the main (or only) examiner (for private identification systems), or is a legal authority (such as the government for some national identification schemes) having some sovereign legitimacy.

This trust in the identity authority for the proper administration of the system and the digital identities need not extend to other functions however: in particular, the registered person may distrust the identity authority as regards the use of his/her sensitive private information, or its willingness to track the registered persons as they use the identification system.

Finally, the identity authority also often serves as information authority for part of the personal information contained in the system.

3.4.5 The information authorities

An identity authority is responsible for certifying some part of the personal information contained in an identification system. As such, it needs to be trusted by both the examiner and the registered person for its authority to certify personal information.

While the identity authority is usually also an information authority, other parties could assume this function as well: in the case of the (state) driver's licence in the United States for instance, while the Department of Motor Vehicles (DMV) would act as identity authority, and as information authority for the driving credentials, the civil information could be certified by some other information authority (for instance, the date of birth could be certified by a vital records agency, which issues birth certificates).

As for the examiner, one may want to create categories of information authorities, depending on the type of information each is allowed to certify. While the vital records agency would be a natural information authority for dates of birth, it would not have any legitimacy in certifying address-related information for instance.

Also, an information authority may or may not be authoritative of the personal information it certifies. While the management (storage, update, etc) of physical attributes or biometrics data may befall the identity authority itself, virtually anybody can act as an information authority to certify this type of personal information. It will be up to the examiner to decide whether it trusts the information authority as regards the certification.

Finally, we note here that we will often use the term "credential authority" instead of "information authority" in the future, when the certification of the personal information takes the form of a credential.

3.4.6 Multiple roles and authentication of the parties

When considering the different parties involved in an identification system, one needs to keep in mind that a single person can have multiple roles: any examiner is likely to be also a registered person, the identity authority often also acts as an information authority, etc.

Also, how does an examiner, identity authority or other information authority authenticate to the identification system to get the authorization to perform his/her function? This in turn is an identification application in itself, which would most likely need an authentication method as least as secure as the one used in the identification system. Should the identification system itself be used, or another system?

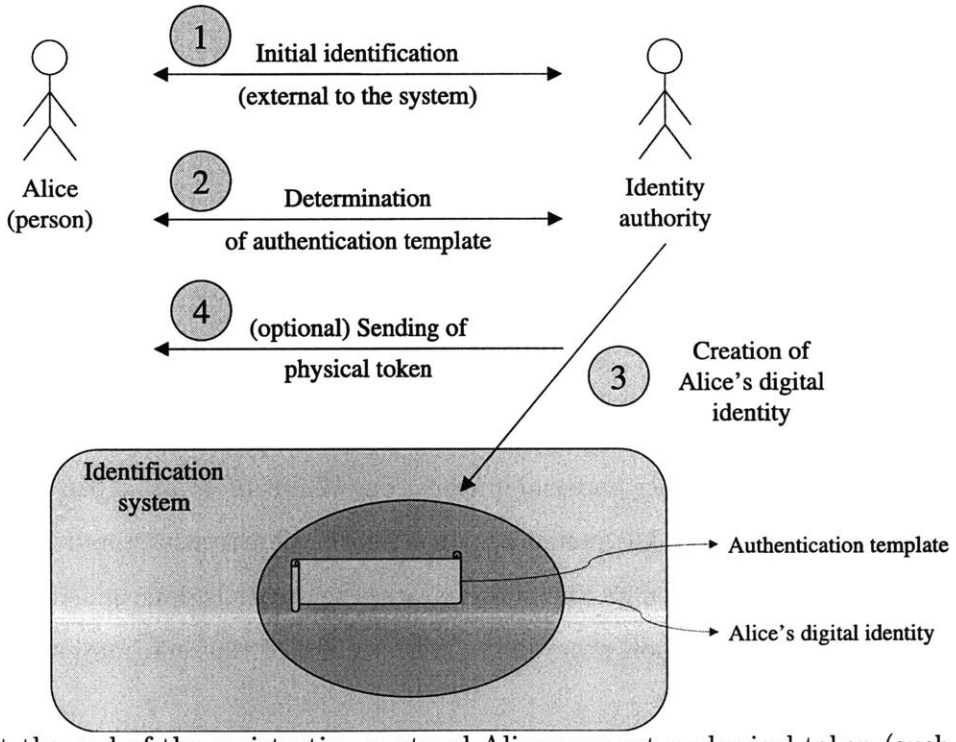
- Should the identification itself be used for authentication of the persons with examiner, identity authority or information authority privileges, this would require them to be registered persons themselves. Then, whether the examiner/identity authority/information authority privileges should be stored in a specific profile of the registered person having such privileges, or new identities be created for this purpose deserves much consideration. While the former would reduce the need for multiple identities (or actually be motivated by the desire to prevent multiple identities), the latter preserves the privacy of examiners/identity authorities/information authorities as private individuals.

- Should a specific "authentication" system be used for the authentication of examiners, identity authorities and information authorities, one should consider the following: since examiner privileges may be granted to a large population, and do not endanger so much the security of the identification system, this specific "authentication" system could be limited to persons needing identity authority or information authority privileges. Obvious deployment considerations – including cost – come with this alternative. Besides, this would introduce the same authentication/authorization problems for that specific "authentication" system, which is an identification system by itself.

Finally, examiners, identity and information authorities may actually be machines: one can envision the use of a machine at the entrance of bars and nightclubs to perform the age verification, the certification of biometric data could be automated, etc.

3.5 Registration

The very first step in the establishment of an identification system is the registration process, where the digital identities are created. This registration process actually breaks down into two separate stages: the initial identification of the person to be registered, and the actual creation of a new digital identity. An important matter to take in consideration in this process is how to handle multiple identities. Figure 3-2 shows a possible scenario for the registration process.



Note: At the end of the registration protocol Alice may get a physical token (such as an identity card) that may contain her authentication template and/or additional personally identifiable information.

Figure 3-2: Registration

3.5.1 The initial identification

In order to properly register a person in the identification system, one needs to "identify" him/her properly (in this case, differentiate him/her from other persons) to make sure the digital identity created corresponds to the person to be registered. For that purpose, and also to enable later identifications within the identification system, one needs to define an authentication template, containing the information necessary to authenticate the person to the system, and determine his/her (digital) identity. This authentication template could be different for different identification systems, according to the system goals. The actual determination of what constitutes an apt authentication template will be addressed in section 4.5.1.

As we will explain later in section 3.6.3, the identification system may have to rely on some identification system (itself or another) for the authenticity of the personal information to be included in the system. In particular, the identity authority may decide to rely on some other information authority (outside the information system) for the authenticity of part the information constituting the authentication template.

This initial identification is of utmost importance for the security of the identification system since all further identifications in the system will rely on the authenticity of the information recorded at registration.

3.5.2 The creation of a new digital identity

The very purpose of the registration process is to create a new digital identity for the person, in order to store his/her personal information. While information can still be added or updated later (cf section 3.6), some personal information is also stored in the digital identity at creation. In particular, the authentication template is determined, verified and recorded during the registration procedure. The specifics of this information storage will be addressed in section 3.6.

Along with the creation of a new digital identity in the identification system, there may be the creation of a physical token containing part of the information contained in the digital identity. For instance, when a person's first registers for a passport, he/she is issued a passport at the same time a new digital identity is created in the passport system.

3.5.3 Multiple identities

When considering the registration procedure, one needs to define how to handle the problem of multiple identities. The central question to be addressed in this matter is: when a person registers for a new digital identity, does the system check whether he/she has already a digital identity in the system?

Identification systems with multiple identities

If the identification system is to allow for multiple identities for anybody, whether the system should check for existing digital identities at registration depends on whether one wants the old and new digital identities to be linkable: indeed, the determination of an existing digital identity (if there exists one) in the system would enable its linkage to the digital identity newly created.

In the case of an identification system for adopted children for instance, one may want both the old and new identities to be linkable to allow for future opening of an adopted child's record. In the case for an undercover secret agent on the contrary, one may want his/her cover identity to be unlinkable to his/her original identity to protect his/her personal life.

Identification systems with single identity

A design goal of many identification systems is to prevent the possibility for a person to have multiple identities. In fact, many national identification systems were established in part to achieve this goal: the impossibility of having multiple identities is an effective way to prevent fraud in many applications. An identification system for Social Welfare entitlement for instance may want to prevent multiple identities to prevent fraud: indeed, a person registering with multiple identities could benefit from Social Welfare the corresponding number of times.

Hybrid identification systems

Some identification systems are somehow hybrid as regards multiple identities: they may want to prevent multiple identities for the majority of the persons, while allowing this possibility for a selected few. A national identification scheme for instance could prevent

ordinary persons from having multiple identities while allowing for intelligence agencies to issue a secondary identity to some of its field personnel.

3.6 Information storage/update

While much personal information can be stored at the registration of a new digital identity, we will explain here why an identification system still needs to allow for an update of the personal information. This information storage/update process actually breaks down into two separate stages: the identification (and authentication) of the person whose information needs to be stored/updated, and the actual information storage/update. The identification-authentication stage will be addressed later in section 3.8, and we will focus here on the issues relating to the actual information storage/update: the databases, and the authentication of the information to be recorded. Figure 3-3 shows a possible scenario for the information storage/update process.

3.6.1 The need for information update

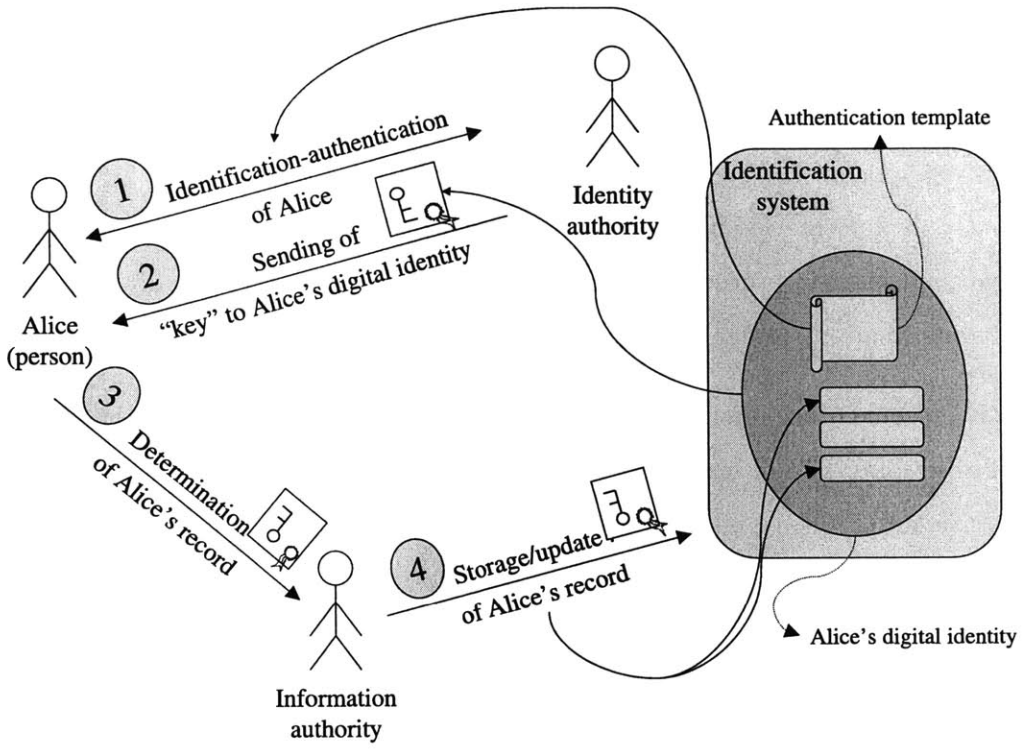
Whatever the purposes and goals of an identification system, there is an inherent need for the possibility of updating a registered person's personal information. Indeed, some pieces of information about the person are bound to change in the course of his/her lifetime: address, last name (for a woman), height and other biometric information, etc.

Also, one also needs to cope with the inevitable errors in recording the information. As Garfinkel points out in his book *Database Nation* [95], errors in credit records are legion in the United States: Associated Credit Bureaus conceded that "errors critical to the decision of offering credit turn up in fewer than 1% of all consumer files". Still, 1% represents more than two million American people, and does not take into account minor inexactitudes, which could make the figure raise to as high as 50% according to privacy activists.

Finally, there may be a legal requirement for the registered person to have access to his/her personal record and update any error. We will come back to this point in section 4.6.

3.6.2 Databases

Although many identification systems involve the use of an identification card, the main component of the system is the set of databases containing the digital identities and record-



Notes:

- During an information storage/update process, the identity authority may also be an information authority.
- There may be many information authorities involved in the same information storage/update process.
- The key to Alice's digital identity need not be the same as that to the corresponding record in which the information authority stores/update her personal information. The latter could be derived from the former by a procedure known only by Alice.

Figure 3-3: Information storage/update

ing the personal information. Even in an identification system with a physical identification token (passport or driver's license for instance), although the token may contain all the personal information needed for identification, it is very likely that the identity authority still records all that information in a separate database, if only to keep a log of all transactions for possible future investigations in the case of fraud, etc.

These databases represent the core of the physical information system. Their security and access control are essential to preserve the authenticity and confidentiality of the personal information of the identification system. The actual problems posed by these databases for a national identification scheme will be later analyzed in section 4.5.4.

3.6.3 Authenticating the information recorded

To ascertain the authenticity of the personal information to be recorded in the system, the identity authority can:

- verify the information by himself/herself. For instance, the determination of height or eye color is a live procedure in the case where the person is physically present during the information storage/update protocol.
- rely on some other information authority. For instance, to determine the authenticity of a person's date of birth, most national identification systems usually rely on birth certificates.
- implicitly trust the person, for personal information for which the person has no interest in falsifying (for instance a shipping information) or that is of little importance in the identification application (such as the person's name for a library card, used mainly for convenience).

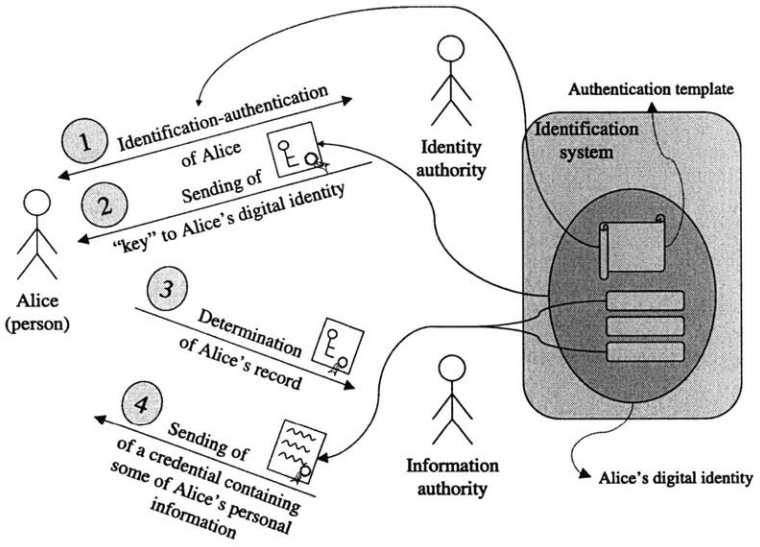
It is important to note here that most identification systems rely in some way or another on some other identification system to establish the authenticity of the information that will be recorded in the system. Even birth certificates, which are issued at a person's birth, rely on some other identification system, to determine the identities of the person's parents. We wish to note here that the "other" identification system can actually be the same system: to determine the person's parents' names for instance, it could rely on the information it has previously authenticated in the parents' digital identities.

3.7 Information revelation

Information revelation is the essence and very purpose of the identification system. Most identification systems function according to the identification-authentication-authorization[9] model. Since many so-called identification applications do not need an actual identification (identity determination) or do not lead to an authorization, we present here an alternate model, which is a slight modification to the traditional model to take into account the aforementioned scenarios. This new approach of revealing personal information without

[9]We use here the terminology commonly used in the literature.

fully identify oneself was first introduced by David Chaum [52]. His model will be detailed in section 6.3.3. Figures 3-4 and 3-5 shows a possible scenario for the information revelation process.



Notes:

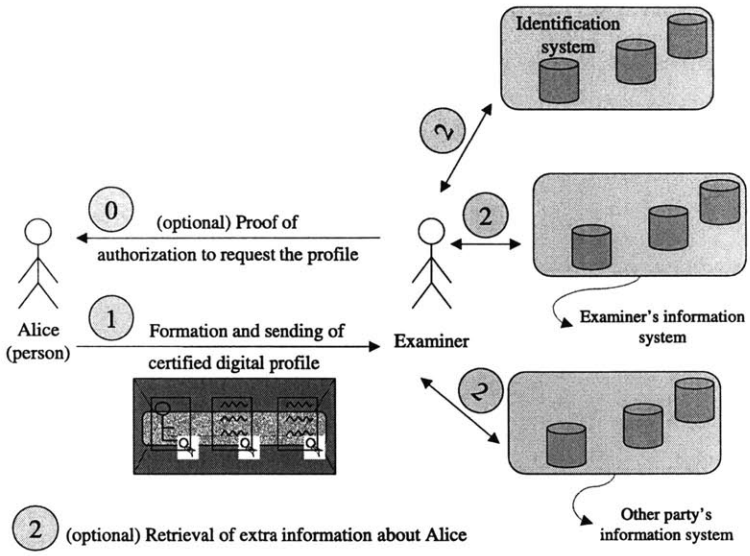
- During this first part of the information revelation process, Alice collects the credential(s) necessary for the formation of the certified digital profile she wishes to show.
- The credential(s) need not contain personal information uniquely identifying Alice.
- Part of this first phase may occur at some indefinite earlier time than that of the actual information revelation process.

Figure 3-4: Information revelation, part 1: credential collection

3.7.1 The traditional model for identification

For most identification systems nowadays, the identification process is usually decomposed into the three steps: identification, authentication, authorization. Identification consists in determining the identity of the person, authentication consists in proving that identity to the system, and authorization consists in getting the authorization for a certain right or privilege based on the authenticated identity.

As we have seen in section 3.1.3, the real purpose of the identification process is actually to reveal (and prove) some personal information (in the form of a profile) about a person registered in the system, without necessarily identifying him/her completely. For that



Notes:

- During this second part of the information revelation process, Alice combines the credentials she obtained in the first part to form the certified digital profile she wishes to show to the examiner. Using David Chaum's analogy with windowed envelopes) [52], Alice puts her credentials in a window envelope, which she passes on to that the examiner, so that he/she can only see the partial information (and the seals) visible through the window.
- Upon reception of the profile, the examiner may get some extra information about Alice by correlating her digital profile with some other information in the national identification system, in his/her own information system, or in another party's information system.
- Upon completion of the profile revelation (and of the extra information retrieval), Alice may be granted some right or privilege by the examiner.

Figure 3-5: Information revelation, part 2: profile creation and revelation

purpose, one needs not go through the three steps aforementioned. We therefore propose to adopt a different framework, which we describe now.

3.7.2 Identification-authentication and profile revelation

The goal of the information revelation functionality can be described as follows: first, the person has to be authenticated to the identification system as a valid, registered person – without necessarily revealing his/her whole (digital) identity – then one of his/her (digital) profiles is revealed through the system – without necessarily having the system (or any other party) getting more information than that contained in the digital profile. We thus

define the following:

Definition 3.10

- An *identification-authentication* is the process by which a person registered in the identification system is proven to possess a valid digital identity in the system. This identification-authentication can be initiated by the person himself/herself, or by some other party, and does not need to reveal his/her whole digital identity.
- A *profile revelation* is the process by which a digital profile of a person is revealed through the identification system, after he/she has been authenticated with a given identity. This profile revelation can be initiated by the person himself/herself, or by some other party, and does not need to reveal his/her whole digital identity. We remind the reader here that the digital profile need not be present in the system prior to this process: it can be computed during this process based on personal information contained in the person's digital identity (and possibly discarded right after revelation).

3.7.3 The nature of identification: our model

In our framework, the identification is composed of the two following steps: identification-authentication and profile revelation. We explain here how this decomposition encompasses all the "identification" scenarios presented above.

- The identification-authentication step usually consists of an identification, followed by an authentication. Yet it does not need to reveal or use a digital identity, or any other personal information: it could just consist in the possession of a token. For example, the possession of the key to the office of a company's CEO could authenticate the key holder as one of the persons authorized to have the key (who could be for instance either the CEO or his/her secretary). More generally speaking, an identification-authentication may consist in only proving that a person's digital identity belongs to a certain group: this group could be defined by its authorization to certain rights or privileges (having the CEO's key in the example above) or else could simply be the entire population of the identification system. *The identification-authentication step does not require full identification.*

- The profile revelation step essentially comes down to disclosing some personal information about a person's digital identity. In a "surveillance" identification system, no right or privilege is granted to the person (in fact, he/she could even be unaware of the very existence of the system), but the party administering the system gains some information about the person, which can take the form of a negative statement (the person is not present in the "suspects" database). In other identification systems, this profile revelation step could prove the membership of the person in a group or the fulfillment of certain requirements, leading to an authorization to perform some action. *The profile revelation step consists in revealing some personal information about a person, without necessarily granting him/her an authorization to perform some action(s).*

3.8 Identification-Authentication

In the security field, one traditionally distinguishes the following sources of information upon which (identification-)authentication can be performed:

- who the person is.
- what the person has.
- what the person knows.
- where the person is.

Strictly speaking, identification is about who the user is. We will see however in this section why one may want to use the other sources of information for an effective identification-authentication. After exploring further the relationship between identification and authentication, we review traditional authentication methods (used by human beings or machines), then explain how the development of biometrics has significantly changed the problem of authentication, and finally present the key points to consider when deciding on the right authentication method.

3.8.1 Identification-authentication without identification?

As mentioned in section 3.7.3, the identification-authentication need not include a full identification, i.e. the revelation of the person's (full) digital identity. In fact, the identifica-

tion part has often been used in identification systems only in order to achieve identification-authentication. However, *identification-authentication does not need prior identification.*

The actual ability to achieve identification-authentication without identification could be based on the paradoxical notion of zero-knowledge proof, first introduced in 1985 by Shafi Goldwasser, Silvio Micali and Charles Rackoff [103]. With a zero-knowledge proof, a prover can prove convincingly to a verifier a property or result without passing to the verifier any more knowledge than the property/result itself. For instance, a zero-knowledge proof of "I am over 18" would not reveal the prover's date of birth or age, but just this very fact.

More generally speaking, a zero-knowledge proof enables the proof of a property about a digital identity without revealing any information identifying that digital identity, and thus provides the technical means for identification-authentication without identification.

The notion of zero-knowledge will be presented in more detail in section 6.2.6.

3.8.2 Traditional authentication methods

We will review here some commonly used authentication methods, using the following sources of information: who the person is, what the person has, what the person knows, where the person is.

Passwords and secrets

In many common applications, the authentication is performed through the use of a password or secret.

MIT for instance has a computer network called Athena with currently 377 workstations[10] located in Athena clusters throughout the campus, accessible at virtually any time to any MIT student, faculty member, or on-campus staff member. If you are a registered user, all you need to know to log in to any of these workstations is your username and password.

For authentication over the phone (to identify yourself to the customer representative of an organization for instance), widely used "secrets" include the individual's date

[10]Information taken from http://web.mit.edu/olh/Clusters/.

of birth, social security number and mother's maiden name. These "secrets" may be specific to the organization, such as the individual's account number for instance.

The main advantage of this type of authentication is its voluntary aspect: you need to provide what you know in order to authenticate yourself. Consequently, you cannot be authenticated against your will.

Besides, as long as you keep your password/secret confidential, nobody can pose as you but by guessing that password/secret. While nothing can prevent an adversary from being lucky at this task or trying all possibilities, one can make the likelihood of the former as low as desired and the cost of the latter as high as desired.

Cards, badges and other physical tokens

Another traditional authentication method is the use of physical tokens such as cards or badges.

MIT issues an MIT card for all members of its community. After initial registration for the card, the cardholder can access all the premises on campus (certain buildings such as the athletic facilities require the possession of the card for entrance), pay for most of his/her expenses on campus and benefit from all other advantages granted to the MIT community[11].

Police officers carry a specific police badge. Although a police officer usually also carries an ID containing more information (name, photo, etc), in many situations, the mere possession of the badge authenticates the badge holder as a police officer.

Unlike the use of passwords and secrets, there is theoretically absolutely no chance at bypassing the authentication without a valid token. However, this introduces the need to remedy the loss or theft of valid tokens, as well as that of using adequate – possibly very strong – anti-counterfeiting measures.

[11] Although the MIT card is strictly speaking an ID, for most of the uses, the mere possession of the card entitles the cardholder to all the benefits: indeed, the card readers only read the data contained on the card magnetic stripe – which does not contain the photo – and does not verify whether the person presenting the card is the valid cardholder. For more information on the MIT card, see http://web.mit.edu/mitcard/

Personal information

In many common situations, sometimes even without our noticing, there is a tacit authentication based on personal – especially physical – information.

In a residence with a concierge, while he/she will always question any visitor or stranger for the reason of his/her presence on the premises, he/she will never stop any resident: he/she simply recognizes him.

Looking for fingerprints at a crime scene has become common procedure for police detectives. Indeed, the presence of fingerprints on the scene "authenticates" (in theory) the owner as being present on the premises at some point in the past, an evidence which could lead to the solution of the case.

The ways in which we – human beings – identify people are varied: from purely physical traits through dressing to behavior, we pick up hints (sometimes subconsciously) that lead to an identification(-authentication). Until recently however, identification-authentication to computer systems relied little on personal information: the technology for enabling computers to effectively get information about who you are has only become mature lately. We will cover this in more detail in the following section about biometrics (3.8.3).

This form of identification-authentication is by definition bound to the person as a physical human being. This may seem at first all we need for an identification system. We will see however in section 3.8.5 the benefits of using the other sources of information.

Location-based authentication

Although not so common, an authentication based on where the person is can prove to be useful.

Since the access to military restricted areas are strictly enforced, the mere presence on such premises authenticates (in theory) the person as an authorized person.

Likewise, the presence of a person in a bar in the United States authenticate him/her (in theory) as a person over 21 years of age.

The recent development of technologies such as the Global Positioning System (GPS) has enabled a totally new range of authentication applications [77].

Two-factor authentication

To combine the benefits of the previous types of identification, many existing applications demand a two-factor authentication:

- what you know and what you have. To withdraw cash at an ATM (Automated Teller Machine), you need to provide your debit/credit card along with the corresponding PIN (Personal Identification Number).

- what you have and who you are. Most current identification schemes rely on the possession of an ID along with personal information on it to match it to the ID bearer.

- what you know and where you are. To set up a meeting with a person you have never met before, a common solution is to agree on a location (and time) for the meeting as well as a code. The presence at this location (and time) and the knowledge of the code then authenticates the other party as the person you were to meet.

3.8.3 Biometrics: the future of identification-authentication?

The traditional identification-authentication techniques presented above have proven to be effective and suitable for their respective domains of application.

Information about people's physical characteristics is currently widely used on all sorts of IDs to enable the identification(-authentication) of an individual with adequate accuracy, or in other words, to use our terminology, the matching of the person's physical body to the individual's digital identity. These characteristics used to be mainly verbal description of physical attributes such as height, eye color, etc.

But the combination of science and technology have led to the increasing use of biometrics for identification/authentication purposes. As defined in a brief study on biometrics by the US Air Force Material Command [66]:

> "Biometric identification is a broad category of technologies which provide precise confirmation of an individual's identity through the use of that individual's own physiological or behavioral characteristics. A physiological characteristic is a relatively stable physical characteristic such as a fingerprint, retinal scan, hand

geometry, or facial features. Behavioral characteristics are influenced by the individual's personality. These include voice print, signature, and keystroke."

This study also includes a brief overview of different biometrics technologies. A good description of the more popular ones can also be found in chapter 3 of Simson Garfinkel's book *Database Nation* [95], along with a good discussion of their use for identification.

Biometrics performance

Many studies have been conducted to assess the performance of actual biometric technologies: for the European Commission in 1997 [135], for the Federal Highway Administration by the U.S. National Biometric Research Center in 1997 [159], for the International Civil Aviation Organization in 1999 [105], and more recently by Deutsche Bank Research in 2002 [138]. In particular, the U.S. National Biometric Test Center [28] has done extensive work on biometrics [158, 49].

To give the reader an idea of the current performance of biometrics, we reproduce here in figure 3-6 parts of the results of the recent study conducted by Deutsche Bank Research in May 2002 [138]:

Biometrics and identification

The use of biometrics identification has been studied by many. Some [31] study more specifically the issues regarding face recognition, while others [3] address various aspects of the use of biometrics in general, and identification in particular. Roger Clarke, an Australian privacy expert, published many papers regarding the issues related to the use of biometrics for human identification [59, 60, 62], as well as to datasurveillance and information privacy [61].

As shown in figure 3-6, even the most efficient biometrics techniques are not perfect and their use for a large population like that of the United States will inevitably lead to false authentications (0.0001% of 250 million people still represent 250 persons). Besides, the system would actually need to achieve an "intelligent match": the biometric representation has to be matched to the corresponding physical characteristic of the individual, that may and *will* change over time. For instance, in the case of face recognition, there can be a tremendous difference between the way a person looks before and after he/she radically

Biometric System	False Acceptance Rate	False Rejection Rate	Failure to Enroll Rate
Iris scan	0.0001%	0.25%	0.5%
Fingerprint (2)	0.008%	2.5%	1%
Voice recognition	0.03%	2%	0%
Fingerprint (1)	0.08%	6%	1%
Hand geometry	0.70%	0.5%	0%
Fingerprint (optical)	0.45%	11%	2%
Face recognition	0.45%	17%	0%

Following are the definitions of the performance indicators used, as defined in the study:

- Not everybody can necessarily be enrolled in a given biometric system. The Failure to Enrol Rate (FER) measures the percentage of users that can not use a specific biometric system. For example, manual laborers sometimes have abraded fingerprints that can not be detected by a sensor.
- Not every legitimate user is necessarily recognized by a biometric system. The False Rejection Rate (FRR) (see chart) measures the percentage of valid users that are incorrectly rejected. For example, a gardener might have varying cracks in the skin of his fingers that are mistaken as minutiae.
- Not every illegitimate user is necessarily barred by a biometric system. The False Acceptance Rate (FAR) measures the percentage of impostors that are wrongly accepted. For example, a face recognition scheme might not be able to discern identical twins.
- Both the false rejection rate and the false acceptance rate that are finally realized when the system is in operation depend on where the decision threshold is set (i.e. on what the system considers as similar enough). To remove this dependency, one often uses the Equal Error Rate (EER), where the system is as likely to reject a legitimate user as it is to accept an illegitimate user. That is, the decision threshold is chosen in such a way that FRR=FAR.

Figure 3-6: False Acceptance Rates (FAR), False Rejection Rates (FRR) and Failure to Enrol Rates of typical biometric systems

changes his/her haircut.

We will not delve in this thesis into the technical details of current biometrics techniques. Rather, we refer the reader to the abundant resources available, that we present briefly in the next subsection. However, we will address the implications of its use for a national identification system in subsequent chapters. Although very appealing because of the physical binding it offers (which may prevent numerous types of fraud), biometrics – if not used properly – may seriously threaten the privacy of the persons registered in the identification system. One needs to keep in mind however that, as the only means for a machine to get some information about a person's physical characteristics, the use or non-use of biometrics is a matter that must be addressed by any proposal for a national identification scheme.

Biometrics resources

For more information about biometrics, the reader can consult EPIC's Biometrics Identifiers section [81], which lists news items and resources on the subject. Also, the Avanti Biometric Reference [3] presents a good overview of biometrics (*The Biometric White Paper* [36]) and interesting studies of biometric-related issues, such as *The Distinction Between Authentication and Identification* [37]. If you are looking for FAQ (Frequently Asked Questions) sections, we recommend those by the International Biometric Industry Association (IBIA) [104] and Dr. Manfred Bromba [43].

Other organizations maintaining websites with resources on biometrics include: Biomet - the Biometric Center [4], the Biometrics Institute [7], the Biometric Digest [6], the Biometric Consortium [5] and the International Biometric Industry Association (IBIA) [14].

Finally, both the Biometric Consortium and MSU (Michigan State University) maintain good lists of publications: the former [68] is focusing more on policy issues and industry practices, while the latter [35] is focusing mainly on research publications.

3.8.4 Remote authentication and human eligibility

An interesting problem is the ability for a person to remotely authenticate himself/herself. A remote authentication is an authentication where the person and the examiner are not physically present in the same place. For example, all authentications over the internet or over the phone are remote authentications. Since the examiner cannot physically verify that a person does not pass the authentication on behalf on another person, the problems of identity sharing, lending or other renting need to be addressed by suitable regulations and policies. Strong deterrents to prevent this fraud include the provision of heavy sanctions (legal, penal, financial, etc) in case of fraud detection, or as we will see in section 6.3.4, the possibility for the beneficiary of the fraudulent authentication to impersonate the accomplice (be authenticated as him/her) whenever he/she wishes to do so in the future.

Also, one needs to address the problem of who owns/controls the hardware used at the remote location (if some computer system is used remotely). In particular, the examiner may or may not trust this hardware.

Another problem arising when addressing remote authentication is the problem of whether the party who authenticates himself/herself is actually human. Indeed, with the

advance of technology, a computer machine or agent could be programmed to pose as a human being as regards the authentication process. For instance, since an authentication over the phone essentially consists in answering a series of questions, a computer agent, equipped with voice-recognition and voice-synthesis software and a natural-language processing tool to process the questions and select the appropriate answers, could pass the authentication provided that it knows the information constituting the answers to the questions. An interesting research project in this domain of checking the human nature of the authenticating party is the CAPTCHA project [8]. "A CAPTCHA is a program that can generate and grade tests that:

- Most humans can pass.
- Current computer programs can't pass."

3.8.5 What is the right authentication?

To decide on what authentication method (or combination of methods) meets the best the needs of an identification system, one needs to take into account the following:

- The accuracy of the authentication is likely to be an important measure of the effectiveness of the system: for most identification systems, the better the accuracy the better the system.
- In a voluntary identification system, the registered person may want to be able to control the amount of personal information revealed in the authentication step. Therefore, using what he/she knows or has could prove to be critical to the application concerned.
- While the use of biometrics may lead to a very accurate and physically binding identification-authentication, it remains uncertain whether one can reconcile the use of such highly identifying authentication techniques with the cryptographic privacy techniques that would enable a rather anonymous identification-authentication.
- Whatever the authentication method, there is most likely going to be false positives (accepting invalid authentications) and false negatives (rejecting valid authentications). Since a lower rate of false positives would lead to a higher rate of false negatives (and vice versa), care must be taken in determining the adequate tradeoff. Also,

the security implications of a false positive need to be addressed, while there needs to be provisions in the case of false negatives.

3.9 Profile revelation

As mentioned earlier in section 3.1.3, a major purpose of the identification process is to obtain a (digital) profile of a registered person. We will explain in this section how to use credentials to authenticate the information contained in the (digital) profile, and present some features one may want to have regarding the profile revelation process: a selective disclosure of personal information, a non-discarding of negative personal information, and the proof of negative statements. Also, we address the matter of whether the procedure should be online or offline.

3.9.1 The use of credentials

Whether the digital profile involved in the profile revelation step is disclosed by the registered person himself/herself or by some other party (such as a database system), the examiner needs to get some assurance about the authenticity (and accuracy) of the information contained in the digital profile. The certification of this information may be achieved through the use of credentials, a notion introduced by David Chaum [52]in 1985:

Definition 3.11

- David Chaum defines *credentials* as "statements based on an individual's relationship with organizations that are, in general, provided to other organizations." [52]. For our discussion here, a *credential* is more generally a document certified by some credential authority. This is a reminder of the definition introduced earlier.

A certified profile will then consist of one or many credentials. These could be issued at the moment of the profile revelation itself or at some indefinite earlier time.

For a credential issued some indefinite time before the actual profile revelation, an important question to address is its validity period. In many situations, a credential needs to be revoked before the end of the validity period (for example, a "driver's license" credential needs to be revoked in the case of suspension of revocation of the license). Credential

revocation however is a complex problem that does not have a good and effective solution in all contexts. This matter of credential revocation will be further examined in section 3.10.7.

The duration of the validity period of a credential is therefore likely to be an important design decision, involving the following tradeoff: short-lived credentials decrease the risk of mistakenly accepting revoked credentials, while creating some inconvenience on the registered person who would then need to renew his/her credentials frequently.

3.9.2 The selective disclosure of personal information

In the case of a voluntary identification of the registered person, we have seen in section 3.1.3 that his/her ability to carry out a partial identification by revealing only a digital profile and not his/her whole digital identity was essential to preserve some level of privacy.

The actual choice of what information to disclose in the profile revelation process can take place at different moments:

- He/she can make this choice at the issuance of the credential(s), which can be some indefinite amount of time before the actual profile revelation. For instance, the registered person can decide to partition his/her digital identity as follows: when requesting credentials to the credential authority, he/she could obtain credentials containing only a specific category of characteristics (for example age or address information). To reveal a profile containing characteristics of many categories, he/she would use a combination of the corresponding credentials.

- He/she could also decide at the time of credential disclosure to hide part of the personal information contained in the credential, or more generally only reveal a set of properties about this information. We will see in section 6.3.3 that current technology actually enables this surprising possibility.

3.9.3 The non-discarding of negative personal information

The ability for the registered person to prove his/her membership of a group or the his/her fulfillment of certain criteria without necessarily revealing his/her whole digital identity is a feature that would be desirable in many identification systems. On the other hand, the impossibility for him/her to hide some "negative" personal information in certain situations may be desirable as well.

Suppose for instance that the identification system carries out the function of providing driving authorization. One may then want to record any major traffic violation or offense in the registered person's record. While this may be kept undisclosed for "ordinary" identifications, this could be part and parcel of a "complete" driving profile, which the person would be required to disclose in exceptional circumstances (if he/she is caught for a traffic offense for instance). We will see that the technology presented in section 6.3.3 enables also this useful feature.

Nevertheless, this feature, if adopted, needs to be implemented with care to avoid abuse (since it diminishes the level of privacy attained by selective disclosure), and also to prevent unfair discrimination.

3.9.4 The proof of negative statements

Sometimes, an identification process is mainly intended to determine a negative statement: for instance, most security checks aim at determining that you are *not* dangerous. We will see that the technology presented in section 6.3.3 provides for the technical means to achieve this unexpected possibility.

3.9.5 Online or offline?

Although the profile revelation theoretically involves the identity authority and some information authority(ies) in addition to the registered person and examiner, the aforementioned authorities need not be present at the actual time the process takes place:

- An online process is an interaction involving a third party, depository of some sensitive information. A typical example is when you pay with your debit card: you swipe the card, enter the PIN number, then the terminal asks a central server whether you have enough money in your checking account and, should this happen, authorizes the payment.

- An offline process on the contrary is an interaction without any other party, but the registered person and the examiner. A typical example is when a law enforcement officer asks you for your driver's license. He/she determines the authenticity of the license and information it contains, based on the license alone.

3.10 The different protocols

We briefly describe here the different protocols involved in an identification system. A protocol is an interaction between two or more parties during which some function is performed.

3.10.1 Digital identity creation

The purpose of the registration functionality is to create a new digital identity. During the digital identity creation protocol, a person is created a new digital identity, which he/she will use consequently in any interaction with the identification system. Depending on whether the system allows for multiple (digital) identities, the protocol may include the determination of whether the person has already some existing (digital) identity(ies) in the system. An authentication template is created along with the digital identity, and stored in the system. Some other personal information (information pointers, descriptive personally identifiable information) is usually added to the digital identity at its creation.

3.10.2 Digital identity update: personal information update

The digital identity update protocol is the core protocol of the information storage/update functionality. In this protocol, the registered person's digital identity is updated with some new information. Proper identification-authentication needs to be performed prior to the protocol to ensure the updated information actually corresponds to the targeted digital identity. The specifics of the update protocol depend on the category of the information to update.

Authentication template

The possibility of updating an authentication template cannot be avoided: for instance, the height (which is likely to be included in an authentication template) of a growing child or teenager is bound to increase over time.

However, insofar as a change in the authentication template may in theory change the binding between the person and his/her digital identity, this very possibility needs to be strictly regulated. Besides, while in some cases (growing height for instance) an update of the authentication template seems natural, in others, the revocation of the current digital

identity and the creation of a new one may be a better solution (change of gender for instance).

Pointers to other information systems

While a change in the pointers to other information systems does not change the way a person authenticates to the identification systems nor his/her digital identity, it still needs to be handled with care, since this may mean a change in the binding between the person and part of his/her digital identity. Indeed, this process could lead to the digital identity containing a pointer to a part of someone else's digital identity.

Personally identifiable information

Personally identifiable information is the type of information the most prone to a change, while being the least sensitive. No special attention needs to be paid to this case, but to make sure to properly authenticate the targeted person, and the information to store/update before the actual update of the information.

3.10.3 Digital identity revocation

The matter of digital identity revocation is common in identification systems allowing for multiple identities. But it also naturally arises for single-identity or hybrid identification systems, if only in the case of natural death.

One then needs to decide on whether to keep the revoked digital identities in the system. While in the case of a decease, there could be an apparent reason to simply delete the digital identity to avoid identity theft, there are many reasons why one may want to just deactivate a revoked identity in the general case: one may want to keep a history of all transactions in the system for audit or archive purposes, or may want to allow for the possibility of a reactivation of a previously revoked identity (for instance in the case of an adopted child's original identity, or the correction of a mistaken death notification).

3.10.4 Credential issuance

In this protocol, the credential authority issues one or more credentials containing personal information contained in the registered person's digital identity. Upon issuance, one can wonder where this credential should be stored. Should it be on a token you keep with

you (typically a national ID card) or remotely in some other database, in which case you would just hold the pointer indicating the storage location in the database? We will come back to this issue in section 4.5.

3.10.5 Credential disclosure

While the examiner needs to be sure of the authenticity of the information contained in the credential(s) disclosed by the registered person, the latter also needs to get evidence of the authorization of the former to request these credentials. Indeed, you don't want for instance to disclose your driver's and vehicle information to an insurance company marketeer who pretends to be a law enforcement agent for instance.

3.10.6 Credential renewal

For numerous reasons, it may be desirable to have a limited period of validity for the credential. Then upon expiration, you will need to renew it. An interesting question is whether this renewal is a prolongation of validity of the present credential, or a reissuing of a new credential containing the same information.

For instance, when staying abroad with a visa, in some cases, you can get an extension of your current visa, while in others, upon expiration of your visa, you are reissued a new one, containing exactly the same information, but for some "identifying" information (such as the visa number) and of course the expiration date.

3.10.7 Credential revocation

Even with credentials with limited validity, it may be needed to revoke someone's credential(s). As mentioned in section 3.9.1, a "driver's license" credential needs to be revoked in the case of suspension of revocation of the license.

Unfortunately, there is no perfect answer for the issue of credential revocation. A good solution would aim at minimizing the risk of accepting revoked credentials while avoiding too much inconvenience.

Chapter 4

Policy issues and technology challenges in a national identification scheme

We discuss in this chapter the main policy issues and technology challenges facing to the establishment of a technology-enabled *national* identification scheme. We present here the pros and cons of such a scheme as well as the essential tradeoffs to consider in the deployment of a national identification scheme, should such a scheme be adopted. Also, we present many design possibilities and scenarios; these examples are not to be considered as design suggestions (and thus will not be analyzed thoroughly), but rather their very purpose is to illustrate the vast range in the various forms a national identification scheme could take. Although we try to cover the essential policy issues to be considered, the purpose of this section is not to provide a comprehensive account of these, but rather to focus on how the recent advances in technology shed new light on some traditional views of national identification systems, and to what extent this impacts the corresponding policy issues. We refer to CSTB's excellent report [113] for a more complete exposition of the important policy issues and technological challenges involved in a national identification scheme.

4.1 The notion of national identification scheme

Prior to exploring the issues and challenges relative to a national identification scheme, let us first define in this section how we interpret the very notion of national identification scheme. Loosely speaking, a national identification scheme is a system (or group of systems) aiming at providing reliable identification for a nation's people.

Insofar as a proof of identity (or lack thereof) may lead to some legal consequences, the natural entity for supervising a national identification scheme (and act as an identity authority) would be the government or some government agency. Although industry initiatives could be conceivable for some applications, we focus here mainly on schemes that would have the sovereign endorsement of the government. The government may then have some limited legal accountability for the effectiveness and security of the scheme. Such a scheme would be tailored to meet the needs of a given country (and thus be quite different for various countries), and could be established either for the sole use of the government and government agencies, or as a basis for other identification systems in use throughout the country.

4.2 The possible interests of the stakeholders

To get a good idea of the problems and challenges any national identification scheme would have to solve, we ought to consider the divergent interests of the different parties who would be involved.

4.2.1 Private individuals

As we have seen in section 3.1.3, the private individual may want to identify himself/herself to benefit from a due privilege or right. Having an identification system record the personal information he/she wants to disclose at some later time however may represent a real threat to his/her privacy, insofar as he/she may have little or no control over who has access to his/her personal information.

With the advent of technology, the use (and misuse) of this personal information has become easier. For instance, by just scanning the magnetic stripe of a driver's license, you can instantly read *all* the digital information contained in it: the New York Times recounted in March 2002 the story [153] of a bar owner in Boston, who could use the data collected

when checking for IDs at the door of his bar to derive demographics conclusions about his clientele and possibly use it for marketing purposes. What most people don't know however is that there has been active research in privacy-enhancing technology. While for current paper IDs, it is still possible for people who check your ID to learn your address even in a check of a couple of seconds, it is actually possible to prevent the examiner from even seeing it with an electronic ID. This feature, along with several others, will be detailed in chapter 6.

Even though technology might enable the person's control over what information he/she decides to disclose, policy safeguards need to be set up to enforce the privacy of private individuals. The use of a national identification scheme needs to comply with the fundamental right to personal privacy and anonymity in a democracy. Also, one needs to make sure the national identification scheme is not abused for discriminatory purposes[1].

4.2.2 The government

Who is the government? In this thesis, by government, we mean the executive authorities at the country/federal level as well as at the state or local level. In the following, when we want to mention either of those specifically, we will explicitly say "federal government" (or country government), "state government" or "local government".

What could be the interest of the government in deploying a national identification scheme? As we will see in section 4.9.5, the cost of such an undertaking could be enormous, and a substantial part of it will be most likely to be borne by the government itself. Will it be worth it?

As described in Privacy International's FAQ on identity cards [110], the governments' motivations for setting up a national identification scheme are varied: they go from proving citizenship in France to establishing Social Welfare entitlement in New Zealand to helping to improve government administrative efficiency altogether in the Netherlands.

We can broadly divide these goals into the following categories:

- Increasing the government's control over its people. From enhancing national security to controlling illegal immigration, a national identification scheme is maybe one of the

[1] A very tragic example of such abuse is the use by the Nazis in World War II of the population data systems in Germany, Poland, France, the Netherlands, and Norway as an operational tool of genocide during the Holocaust [70].

most effective tools a government has in its possession to monitor its people.

- Improving government administrative efficiency. With a national identification scheme in use throughout the country, government agencies – such as those handling social security, taxes, immigration or customs – could improve their productivity.

- Providing a valuable service to society. From building highways to providing medical care to maintaining justice, the role of the government is to help improve the quality of the life of its people. Facilitating the life of the people as regards their daily interactions with various organizations is a benefit worthy of consideration for the government.

We present here a possible allocation of the prerogatives of the different legal authorities (federal, state, local) in a federal country like the United States, and address the issues specific to police and law enforcement in section 4.2.3, and to government agencies in section 4.2.4.

A possible allocation of prerogatives for a federal country

Although the allocation of the prerogatives of the federal, state and local governments comes within the competence of lawmakers, and highly depends on the Constitution and other legal usages of the country adopting the national identification scheme, we describe here a possible scenario that might be applicable to a federal country like the United States.

In a federal country, the role of the federal government may be limited to a supervisory role. In the United States for instance, apart from the passport which is intended for international use, internal IDs (driver's licenses, liquor IDs, etc) are issued at the state level. Each state has its own system of IDs that may be incompatible with another state's system. A reasonable model for the United States could be letting the federal government define the guidelines and directives for the system, as well as the standards and main specifications to ensure the compatibility of the different state subsystems, while letting the different states run their own subsystems.

State and local authorities on the other hand would also play an important role in a national identification scheme. Even in the extreme case of a completely centralized system where the federal government would control every single portion of the system, the registration process and issuance of identity would most likely fall into the hands of state

or local authorities, if only for practical reasons such as efficiency or cost. A more realistic scenario in a federal country would be, as we mentioned above, a system overseen by the government with federal guidelines, but actually maintained and managed by the states. An actual solution we might be heading towards in the United States is the standardization of state-issued driver's licences and other state-issued IDs with a linkage of the underlying databases (cf section 5.2). In any case, any proposal for a national identification scheme needs to address the delicate issue of the future of current state-issued IDs. This question will be further discussed along with other deployment issues in section 4.9.

4.2.3 Police and law enforcement

The police and law enforcement officers seem to be the people who would most benefit from a national identification scheme. Wouldn't it be a dream come true for a police officer to be able to have the complete dossier of a person – including his/her police record – on his/her screen by just scanning that person's ID? It is no surprise that mandatory IDs have been a major tool for maintaining home security during wartime or for totalitarian regimes.

Towards a Big Brother society?

In an identification system, through the registration process, the identity authority can set up a gigantic database of all the registered persons. Then by frequent identifications – be it by explicit identity checks, or indirectly by requiring proof of identity for common transactions – the identity authority could possibly track any registered person. Were this identification system a "national identification system" with the government acting as identity authority, it could be the most sophisticated surveillance system ever invented.

Can the private individual then do anything to prevent a pervasive surveillance of his/her every move, and the nightmare of having some Big Brother constantly looking over his/her shoulders? The answer is: Yes. As mentioned in section 3.9, it is technically possible in a voluntary identification system for an individual to disclose only selected information about his/her digital identity, at his/her discretion. We will present some of the techniques allowing this feature in section 6.3. Nonetheless, the individual's actual ability to protect his/her privacy does not actually reside in technology only but in suitable policy. Indeed, as noted by Privacy International [110], "even in democratic nations, police retain the right to demand ID on pain of detention". What is the point of being able not to disclose one's

name if the consequence is being arrested? Technology can only prevent the examiner from seeing personal information the individual does not want to disclose (which is already a plus), but not the (possibly dire) consequences of not disclosing some personal information.

Towards a totally secure society?

The other common question that comes in mind when talking about law enforcement is: can a national identification scheme prevent terrorism? As it is impossible to prevent anybody from suddenly becoming mad, it is likewise impossible to prevent a person who has been until now an upstanding citizen to turn into the most evil terrorist overnight. The national identification scheme is not the miracle cure to all the security problems.

If nothing can effectively eradicate all threats of terrorism, the next question would be: can a national identification scheme dramatically decrease the risk of terrorism? The actual question many opponents of national IDs want answered is the following: could the tragic events of September 11, 2001 have been prevented if there were a national identification scheme in place? Probably not, but it could have made it harder. However, no matter how high the security standards are, a wealthy and determined terrorist organization could always find a way to bypass even the highest security procedures. "Zero risk" is unfortunately an unreachable goal. Security has always been a matter of compromise: the bottom line is to decide on the appropriate tradeoffs between high security, reasonable cost, ease of use, convenience, privacy and other factors.

We will further discuss the related issues of fraud, and identity theft in section 4.8.

What is then the benefit for police and law enforcement?

If a national identification scheme could actually enhance the individual's privacy against police surveillance, and may not help law enforcement bodies to prevent terrorism, could it then be of any help for police and law enforcement personnel?

First, a technology-enabled national identification scheme should dramatically decrease the risk of identity forgery. For instance, with 50 states in the United States, it is really difficult for a police officer to detect fake out-of-state IDs. Not only is it hard for him/her to tell whether an ID issued by another state is actually a real one, it is almost as hard for him/her to tell whether it actually *looks like* a real one.

Although deliberate tracking of the every moves of a private individual should be prevented, the individual still ought to reveal a rather complete profile in the case of sensitive situations such as the purchase of weapons, explosives or chemical products, or boarding a plane. Although such situations do not require complete identification per se at the moment, they would most surely demand the possibility for complete identification at a later time in exceptional cases such as criminal investigations. To prevent the abuse of this feature though, one may require the authorization of multiple parties to enable it (see section 6.2.4 for the technical details).

Also, even if showing a driving profile may hide your actual identity, one may require it to show a driving "score". Different levels could correspond to the gravity of your driving record: whether you have committed no traffic offense or violation, only minor ones, or repetitive major ones for instance. There could also be more specialized categories: parking violations, speeding, drunk driving, etc. Likewise, one could envision a criminal "score". Although such scores could help improve law enforcement while providing some privacy to upstanding citizens, the pros and cons of such a feature need to be weighed carefully: that all persons having committed some major offense be branded as dangerous criminals could severely hinder any attempt at the social reintegration of past criminals for instance, or even provoke some form of discrimination. A compromise could be the discarding of detrimental information after a certain period of time: in the United States for example, a personal bankruptcy disappears from one's credit history after ten years. The potential risks of the adoption of such a "score" are more further examined when we present the case of the credit reporting industry in the US in section 5.8.1.

Even though these issues come within the competence of policy and law makers, they illustrate the extent to which a national identification scheme could be used to improve national security while maintaining some level of privacy.

4.2.4 Government agencies

As seen in section 4.2.2, the motivation behind the adoption of a national identification scheme in many countries has been the improvement of government administrative efficiency. Many of the uses of the system for tax, social welfare or immigration and customs purposes need to be considered when discussing the adoption of a national identification scheme. Since the operations and budget of public organizations are directly or indirectly controlled

by the government, an improved government administration could prove to be a good payoff for the colossal investments incurred by the deployment of a national identification scheme.

Although some government agencies (such as the FBI or the CIA in the United States) may want to maintain their own identification system for private use, most of them could greatly benefit from leveraging a national identification scheme set up by the government. Such a use should be carefully regulated however: the history of the Social Security Number (cf section 5.4) and its omnipresent use and misuse by public organizations – federal, state and local government agencies – (and by private organizations as well) until the adoption of the *Privacy Act* of 1974 (and now still in many cases) indicates that feature creep is a major risk of a system with goals of a rather "universal" nature. A system designed and run by the government or some government agency is indeed likely to be used (sometimes misused) by many private organizations.

4.2.5 Private organizations and corporations

One of the main players in the discussion over a national identification scheme may actually be private organizations and corporations. Indeed, in countries like the United States, they are the parties with the largest budget. They are maintaining today immense databases with personal information, and perhaps make those more efficient with a national identification scheme.

Personal data and marketing. Nowadays, information has become a valuable commodity, and the demand for more and more detailed personal information has fuelled many a business. Marketing for instance is by itself a multi-trillion dollar business in the United states alone. What does the marketing industry do? It essentially collects personal information about anybody they can get a hold of, in order to determine (among other things) people's spending habits and try to infer the likelihood of the response of a given individual to a personalized commercial solicitation. Garfinkel describes in his book *Database Nation* [95] how this practice has drifted and poses today a serious threat to the private individual's privacy.

Customer Relationship Management. Each company selling to the public maintains some form of database of its customers. Customer retention and loyalty has become a

major concern for companies today as the cost of acquiring new customers may prove to be substantial. By collecting more personal information about their customers, companies expect to improve customer service and hope that a more personalized service (including targeted discounts) would keep its customers away from its competitors. Frequent flyers programs for instance represent a good example of this practice, which furthermore presents a voluntary enrollment.

Rewards programs and other sweepstakes. Many private organizations actually exchange the disclosure of personal information against some concrete benefits. This is one of the purposes of credit cards and frequent flyers programs, and the main goal of many rewards programs and sweepstakes offers. By enrolling in a rewards program of a supermarket for instance, you benefit from year-round discounts in exchange for providing your identity when checking out. This information is usually used to derive some spending patterns and the impact of pricing on the supermarket's products (in which case the actual identity of the member does not really matter, just his/her spending history). Many a sweepstake would ask for some personal information about the participant in exchange for a chance at winning a grand prize. This information could be use later for marketing or other targeted commercial offers.

Industry initiatives Personal information is so valuable that there have been industry initiatives to standardize the use of personal dossiers. We will analyze the industry's two main projects to date (Microsoft .NET Passport and the Liberty Alliance Project) in section 5.5.

4.3 Who should be the different parties?

4.3.1 The person/registered person

When considering the place of the registered person in the system, important policy issues arise: Who should be registered in the national identification scheme? Should the national identification scheme system be compulsory or voluntary?

Who should be registered in the national identification scheme? At first, one may think of *any* person related to the country. This definition is unfortunately very vague.

Should it be all citizens (as in France where the national ID card attests of the person's French citizenship), including those living abroad, or all legal residents (as in Honk-Kong), or maybe both? Should we consider only permanent residents or also temporary ones? In the latter case, where do we set the threshold? For instance, many students come to study in the United States with a student or exchange visa which can last for any duration from a few weeks to many years. What about visitors?

Should the national identification scheme be compulsory or voluntary? Should all the people concerned have to register in the scheme, or should they be allowed to choose whether or not to enroll?

- If the system is compulsory, how do you enforce this obligation? What would be the legal sanctions for not being registered?
- If the system is voluntary, what happens if you are not registered? Are you considered as suspect? And if you are registered, are you relieved from all suspicion? Would the cost (financial, inconvenience, etc) of not being registered be so high that the system would be virtually compulsory in practice?

These questions can only be answered according to the declared goals of the national identification scheme. The determination of these goals is not the object of the present thesis, but the matter of public consultation and debate. It is important however to keep in mind that the answers to these questions will greatly influence the design of a possible scheme, as well as the technical choices.

4.3.2 The examiner

Many people or organizations may be willing (or required) in the course of their businesses, to check a person's identity, or other information about the person. This need for identification may come from a desire to improve their relationships with their partners or a legal requirement: for instance, any business (restaurant, supermarket, liquor store, etc) serving alcohol in the United States is required by law to check if the customer meets the legal age requirements.

Natural candidates to the role of examiners would be police and law enforcement personnel, as well as the staff of government agencies. Private organizations and corporations

may also express interest in participating in the national identification scheme as examiners, or else be legally required (an age verification for instance could be made to be legally valid only if performed through the national identification scheme).

Also, as mentioned in section 3.4.3, should the examiners have different "examination privileges", policy and law makers need to define them appropriately. Furthermore, the use of information obtained or derived from the national identification scheme needs to be regulated to avoid abuses or any other feature creep.

Finally, an analysis of the alternatives to the national identification scheme can prove to be useful. Potential examiners may prefer for a variety of reasons some other system – existing or in project/development, in replacement or complement of the national identification scheme – that may cater to their specific needs. For example, the Social Security Administration in the United States could use a national identification scheme when it has to actually identify physically a person while keeping using its own identification system based on the Social Security Number for all of its other operations. The national identification scheme could on the other hand be the primary identification scheme used by the police, or else possibly subsume the system of state driver's licenses and IDs currently in force in the United States. We will more thoroughly discuss other deployment issues in section 4.9.

4.3.3 The identity authority

As mentioned in section 4.1, we focus our attention here on identification systems with the endorsement of the government. Therefore, a natural candidate for the role of identity authority would the government itself or some government agency, or else an organization expressly authorized by the government for that matter. Insofar as the legal liability of an identity authority not controlled by the government may seriously hinder the legitimacy of the government's endorsement of the identification system, it seems undesirable that the identity authority be some organization not under the supervision of the government.

4.3.4 The information authorities

An information authority is accountable for the personal information it certifies. Some types of information may have a natural candidate for information authority: for instance, a vital records agency would likely to be authoritative on dates of birth. Others may be

certifiable by a large range of information authorities: photos for instance could be certified by any party.

While appropriate policy decisions need to be made regarding who should be authoritative of what data regarding its management (recording, storage, update, etc), one need not regulate the ability of a given organization to act as an "certifier" information authority. Indeed, as mentioned in section 3.4.5, the examiner can decide on his/her own whether or not he/she trusts a given information authority.

4.4 A single system or a group of systems?

Before delving further into more specific considerations, one may wonder whether there should be a single national identification system or a collection of co-existing national identification systems.

As we have seen in section 4.3.1, the scope of the population concerned by a national identification scheme may vary according to its goal. It could be conceivable for instance to have different systems for citizens and foreigners. While both would probably contain basic demographic information, each could contain more specific information: the "citizen" system could contain for instance the date and reason for obtaining the citizenship (birth, naturalization, etc) while the "foreigner" system would certainly contain the country of origin of the individual. Also, one could want to have different authentication procedures for both systems – with for instance the "foreigner" system having a stronger procedure. Last but not least, these systems could be run by different organizations: the "citizen" system could be under the supervision of the ministry of the Interior, while the "foreigner" system could be the responsibility of the ministry of the Foreign Affairs (or the Immigration Services).

Different identification systems could also coexist for different purposes. For instance, different systems could be used to hold, one the full legal names, another the dates of birth, yet another the addresses, etc. In this scenario, the access to each system could be through pseudonyms, which only the registered person knows. Furthermore, the group of systems could be designed so that the different pseudonyms of the registered person are never used at the same time, and thus unlinkable provided that the registered person follows adequate security policies. This idea of using pseudonyms for this purpose was first introduced by

David Chaum [52], and a technical treatment of how this could be achieved is presented in section 6.3.3.

This last scenario could fit into the following possible framework. A "core" identification system would carry out only the functionality of binding the registered persons to their digital identities. Nonetheless, this core system would not contain any actual personal information about these digital identities, but rather some "master key" for each registered person. This master key would serve as a proof of valid authentication, as well as the basis for access to other "informational" identification systems containing the actual personal information. For instance, to retrieve the date of birth of a registered person, his/her master key could be used to derive a "date-of-birth pseudonym" following some verifiable procedure, to unlock the targeted data. For the address, the procedure would be different and yield a different "address pseudonym" that would be unlinkable to either the master key or the "date-of-birth pseudonym".

This idea of a "core" identification system that would provide for just the binding between physical persons and digital identities (with possibly the enforcement of a single digital identity per person as another goal) has already been suggested by David Chaum [52]: Chaum's "is-a-person" organization's purpose is to ensure that each person has at most one digital identity. Also, Anna Lysyanskaya, Ronald Rivest Amit Sahai and Stefan Wolf [117] already introduced the idea of a master key, which possession enables the effective impersonation of the key owner. We will examine these schemes further in section 6.3.

4.5 What information?

We focus in this section on the nature of the information that would be recorded in the system, on where and how it would be stored, and finally on the information that would likely be stored alongside the system by various organizations.

4.5.1 Nature of the information

As we have seen in section 3.2.4, the information contained in an identification system can be of the following categories: identifying information constituting the authentication template, pointers to other information systems, and other personally identifiable information.

While the determination of the data to form the authentication template is a policy question involving significantly the technological state and cost of actual authentication techniques, the definition of the appropriate information pointers and personal identification information that an identification system should include definitely comes within the jurisdiction of lawmakers.

In the probable case of a group of co-existing identification systems, lawmakers need to decide not only what information each system should handle, but also the terms and conditions in how and to what extent personal information in one system is accessible to another system. Indeed, a total linkage of all the digital identities of a person in all the identification systems may defeat the very purpose of distributing his/her personal information across independent systems. To that end, the information pointers present in an identification system may be designed not to directly enable the access to some personal information in another system: rather it could be a key that the person would use to unlock this link (at his/her discretion) to retrieve personal information contained in the other system, yet without enabling a linkable of his/her digital identities in the two systems.

4.5.2 A card-based system or a card-less system?

Contrary to popular belief, a national ID card system is not the only form a national identification scheme could take. Even though most national identification schemes so far include an ID card or some other sort of token, this does not exclude the adoption of card-less systems in the near future. Indeed, even in a card-based system, the card is only the visible part of the iceberg: it would be naive not to take as granted that all the personal information contained in any issued ID would also be recorded in a gigantic database at the time of ID issuance. Indisputably, the main component of a national identification scheme is the underlying set of databases of personal information. We will discuss the issues relative to these databases in more detail in the following section (4.5.4). Let us examine here the pros and cons of a card-based system versus a card-less system.

The benefits of a card-based system

- First and foremost, a card (or any other token) can contain lots of data for a very reasonable cost. More precisely, it can contain confidential data that are hard for a human to remember, such as cryptographic keys. A card-based system could im-

plement the following feature: assuming that biometrics data alone do not suffice to reveal an individual's digital profile, the card could contain the key(s) necessary to unlock this profile revelation.

- A card-based system would allow for off-line identification protocols.
- The card could be also of psychological importance. It increases the confidence of the person in the voluntary aspect of the identification. In particular, it could contain data not present anywhere else in the system, such as the person's secrets and cryptographic (secret) keys.

The benefits of a card-less system

- If there is no card, there is no card to lose. With a card-less system, there is no need to make provisions for lost or stolen cards.
- If there is no card, there is no card to produce. Besides the actual costs of the card, there are also savings for the infrastructure needed to produce the cards.
- In a card-less system, the identification would rely mainly on who the individual is, and what he/she knows. Providing that the individual carefully keeps confidential the secrets necessary to his/her identification, this is likely to reduce the risk of identity theft, and would most surely eliminate that risk subsequent to a loss or theft of the card.

4.5.3 Machine-readable vs. human-readable information

While in an electronic national identification scheme, the concern is much centered around what information is digitally stored in the system in a machine-readable format, one may still wonder about human-readable information present in the system, and especially that printed on the ID card (if the system were to include such a card). At the very least, there should be some information allowing a person to distinguish between two such cards, or identify the cardholder in case of loss. To that effect, most cards today carry the name of their legitimate holder. This is however not a necessity. For instance, Citibank used to issue payment cards without the cardholder's name but only his/her account number and photo.

Also, the nature of the human-readable information is related to the intended use of the new electronic national identification scheme, should it be adopted. In particular, one may want the same kind of information printed on the new ID cards as that printed on the existing IDs, thus enabling the new ID cards to be used in the same way as the existing IDs. On the contrary, one may consider this to be too much information in a future wired world where computer systems would be pervasive, thus enabling the easy access to machine-readable data. This issue, along with others regarding the relationships and possible interactions between the existing system(s) and the new system will be further addressed in section 4.9.2.

4.5.4 Databases

Whether the system will include a card or not, its main component is the underlying set of databases storing all the personal information. A database can take several forms, notably depending on who controls it.

Identity database. The identity database is the main database in an identification system. It is maintained by the identity authority to administer the digital identities registered in the system. This identity database may also contain personal information, insofar as an identity authority also usually acts as information authority.

Information database. Information authorities may maintain their own databases for the personal information for which they are authoritative. These databases, along with the identity database, form the bulk of the infrastructure for the recording and storage of personal information.

Personal database. Contrary to popular belief, the person may maintain his/her own database of personal information. This may take the form of a personal smart-card, a personal computer database accessible remotely, or a part of a bigger database maintained by an organization that the person trusts. Unlike the two former kinds of database, the information contained in here will likely need to certified by some identity/information authority: rather than containing raw information, a personal database will most likely contain credentials.

Examiner's database. Although not part of the identification system per se, databases maintained by the examiner can be used in the information revelation procedure. We will examine in the next section (4.5.5) the implications of such external databases of personal information, which can be tied to the identification system (for instance through the use of the identification system's Unique Identifier if there is one).

Other databases. Other databases, which are not part of the identification system, may be used by the examiner as well in the information revelation procedure: these could be private databases run by some partner, or a publicly available database (such as Google for instance).

4.5.5 What information is stored alongside the system?

Should a national identification system be adopted, it will become inevitably the cornerstone of countless other systems. Almost any company nowadays maintains some form of database containing personal information, if only for its employees. But more in the spotlight are customer databases. Any company from which you purchase some merchandise may potentially have your profile in its customer database. This profile could be anything from mostly anonymous to highly personalized: at one end of the spectrum lies your supermarket which may collect anonymous data about your grocery shopping habits; at the other end, your favorite airline company has very personal information including your name, residence, membership number, favorite seating, etc, which it uses for your convenience to expedite the purchase of plane tickets.

When assessing the privacy of personal information in a national identification scheme, one needs to keep in mind that the most serious threats come from the linkage of this information with related personal information present in other systems. As we have seen in section 3.7, the profile revelation procedure may involve databases of personal information external to the identification system – be they maintained by the examiners themselves or some third party (such as a credit bureau for instance). To avoid some severe consequences of feature creep (cf section 4.8.6), the designers of the system must take into account such "external" factors.

4.6 Who has access to what information?

An interesting question is whether the registered person has access to his/her own information stored in a national identification system. In France for instance, organizations maintaining databases containing personal information are required by law [76] since 1978 to inform the person of the very existence of the database, as well as to give him/her the right to access his/her personal information and correct any inaccuracies.

A key question here is the availability of the databases: Who should have access to these databases? For what type of queries? etc. Being able to query a database to retrieve the full digital identity of an individual who just showed a profile would indeed compromise his/her will to only disclose partial information about his/her digital identity.

Privacy laws throughout the world regulate the use of personal information by the organizations maintaining databases, as well as define the individuals' rights to that information.

Privacy Law

Privacy concerns are of foremost importance in the European Union. France for example passed the *Law No. 78-17 relative to information technology, files and liberties* [76] as early as 1978. Also, the Council of Europe concluded in 1981 the *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data* [125]: "The Convention [...] was the first legally binding international instrument in the data protection field. Under this Convention, the parties are required to take the necessary steps in their domestic legislation to apply the principles it lays down in order to ensure respect in their territory for the fundamental human rights of all individuals with regard to processing of personal data." Later, the European Parliament issued in 1995 *Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data* [133]. These concern *any* database containing personal information, be it operated by a government agency or a private organization.

On an international level, the Organization for Economic Cooperation and Development (OECD) issued in 1980 *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* [89] and the United Nations (UN) adopted in 1990 *General Assembly Resolution 45/95: Guidelines for the Regulation of Computerized Personnel Data Files* [154].

In the United States, the government regulated the use of personal information by its

federal agencies: the *Privacy Act* of 1974 [127, 128] is a "code of fair information practices that attempts to regulate the collection, maintenance, use, and dissemination of personal information by federal executive branch agencies"; the *Freedom of Information Act* of 1966 (amended in 2002) [126, 129] "generally provides that any person has a right, enforceable in court, to obtain access to federal agency records, except to the extent that such records (or portions of them) are protected from public disclosure by one of nine exemptions or by one of three special law enforcement record exclusions".

For more information about privacy law, we recommend EPIC's *Privacy and Human Rights* [48] and *Privacy Law Sourcebook 2001* [142].

4.7 A Unique Identifier?

A most controversial question about national identification systems is whether they should include a *visible* Unique Identifier. Do we want to have a serial number that would enable any person with the right to access the appropriate database(s) to know everything about us? Do we want a computer to be able to instantly display on a screen all our history after scanning a bar code, like at a supermarket cashier? Obviously not. So why even bother discussing the issue? The privacy implications of a Unique Identifier go much beyond these ethical considerations: the unfortunate example of the various misuses and abuses of the Social Security Number in the United States crystallize the depth of the problem, and is the quintessence of feature creep.

Without speculating further about the possible nightmares a society where any human being would be just another bar code or serial number, let us examine more sensibly what could be the benefits and drawbacks of a Unique Identifier, and the essential privacy implications.

4.7.1 A Unique Identifier is underlying in any database

From a technical point of view, the benefits of the existence of a Unique Identifier are obvious: this would enable a one-step access to any given record. After all, doesn't your bank representative access your bank account instantly if you bring with you your bank card or a check that contains your bank account number?

But this not the only efficient way to access a record from a database. Many databases

do not actually have a visible Unique Identifier used to access records. When you want to rent a movie at your local video store, when the clerk wants to check in the store database if your favorite movie is in-store, all he/she needs is the title of the movie (and the year, if there are several movies with the same name). Supposing there are no two movies with the exact same title released the same year, the pair (title, year) is going to uniquely determine the movie. This will form the *primary key* of the table in the database holding all the records for the movies: the knowledge of the value of this primary key alone will enable the user to access the unique movie in the database with that title and year, if there is one. This primary key is nothing else but some form of Unique Identifier: it enables the unique identification of a given record.

Is there any other way to know if your favorite movie is available for rental? If you have a look at the Internet Movie Database's website[2], you can see a plethora of search options. You can retrieve for example all the movies that feature Francis Ford Coppola as director and Al Pacino as actor[3]. Did we need any Unique Identifier or other primary key? No. But the reality is that the server searched through all records for suitable matches, and internally still uses a primary key to identify records.

Although the search capabilities of a database may be various (and may hide the very existence of an underlying Unique Identifier), the truth remains that any (computer) database uses primary keys to manage the access to its records. Therefore, the real question is not about whether a national identification system should include a Unique Identifier (since it most surely does, in some way or another, if only for technical purposes), but rather whether this Unique Identifier should be made *visible* to the users of the system.

4.7.2 Do we need a Unique Identifier to identify people?

We have seen that the presence of a Unique Identifier (or alike) is inherent to any database. But do we need it? Aren't machines (and more generally speaking, technology) supposed to serve the needs of humans, and not the contrary?

Let's stand back and think about how we identify people in our daily life. As we described in section 3.1.4, people usually use names to identify on another. If this is not enough, one may provide some other personal information in order to achieve the identifica-

[2]The Internet Movie Database's website can be found at http://www.imdb.com/.

[3]This search yields *The Godfather Trilogy: 1901-1980* (1992), *The Godfather* (1972), *The Godfather: Part II* (1974), and *The Godfather: Part III* (1990).

tion. Rather than using some form of Unique Identifier, the human process of identification usually uses gradually more information until the identification is complete. For instance, if you want to refer to your friend John Smith, its name alone may be insufficient for identification, since your interlocutor may know many persons named John Smith. So the identification process could unfold as follows[4]:

— [...] My friend John yesterday ...
— John who?
— John Smith!
— Which one?
— Oh, you know many John Smiths? ... OK, I am talking about the blond guy working at *La Taverne*.
— Never been there.
— Well, you know, Anne's brother.
— Oh, I see [...]

We can see that we (humans) do not actually need a Unique Identifier to identify people: we have lived millenniums without resorting to it. Why would we use it today? The main reason might be convenience. But is the convenience worth the risks of using a Unique Identifier?

4.7.3 The privacy implications of a visible Unique Identifier

With the cost of electronic storage plummeting nowadays, we need to take as granted that any information stored digitally today may remain available forever. Have you ever posted a silly message on a bulletin board on the internet during your high-school years? Well, too bad for you: it will probably still be around decades later. Have you ever tried to search for your main email address on the internet? The results may be quite amazing. And this is only the visible part of the iceberg: the Internet only contains a rather small part of the digital information created by man. There are many more databases to which you don't have access that store much more personal information about you, not to mention the databases you are not even aware of.

[4]Actually, this process of identification by humans is more generally applicable to other contexts: for instance, when searching for some information on the Internet with a search engine such as Google, a human first tries general queries, which he refines progressively to get gradually more specific answers.

One common belief is that by clustering the information you provide to different organizations (i.e. partitioning it into several independent pieces to be handled by different organizations), you are protecting your privacy. Yet, the reality is totally different. Giving out your Social Security Number for opening an account with any business you deal with has become common practice in the United States. Once two organizations have your Social Security Number, it is easy for them to link their information about you or merge their corresponding records. The main obstacle to linking records of different databases today is the technical difficulty of matching two records corresponding to the same person without a common Unique Identifier. So wouldn't a Unique Identifier for the people, endorsed by the government itself (or some affiliated agency), be the sought-after Unique Identifier that would enable the merging of all databases containing personal information into a gigantic information-rich database?

The Unique Identifier itself is not so much a threat to personal privacy, but it is its misuse that leads to privacy invasions. As long as a Unique Identifier for a system is not recorded in another system, the danger of linkage of personal information records across systems remains limited. Therefore, should a visible Unique Identifier be included in a national identification system, appropriate policy needs to be determined to regulate its use. Yet, insofar as such policy may be hard to enforce, one needs to weigh carefully the few benefits of a visible Unique Identifier against the dire implications as regards a sure loss of privacy.

4.8 Security of the system

We define in this section the notion of security for a national identification system. Prior to assessing the actual security of a system, one needs to define the security objectives to be attained. We will therefore present in the first part of this section a set of security objectives that are likely to be part of the security policy of a national identification scheme.

We then examine the main security vulnerabilities. As everyone knows, the security of a system is as strong as the "weakest link". In particular, we do not limit our analysis here to the evaluation of the technical security of the system, but also consider the following factors: the level of trust in other parties, the reliance on external systems, the human factor, and feature creep.

4.8.1 Security policy

The definition of the security policy of a national identification system (the set of security objectives to be pursued) is a critical part of the design of such a system. Since these security objectives directly depend on the actual goals of the system, it is up to the appropriate policy and law makers to decide upon a suitable security policy. Therefore, we do not aim here at providing an example security policy, but rather suggest some possible security objectives that we think may be desirable. In addition, some of the security objectives listed below may not be compatible; in particular, those regarding single and multiple digital identities are exclusive of one another.

Security Objective: It should not be possible for a person to "identify" himself as another person:

- It should not be possible for a person to register a digital identity corresponding to another person, real or fictitious.
- It should not be possible for a person to achieve the storage/update in his/her digital identity of personal information corresponding to another person, real or fictitious.
- It should not be possible for a person to achieve the storage/update of his/her personal information in a digital identity corresponding to another person, real or fictitious, and more generally, to gain any other access to a digital identity corresponding to another person, real or fictitious, than those intended and provided for by the system.
- It should not be possible for a person to reveal personal information corresponding to another person, real or fictitious.

Security Objective: It should not be possible for anyone to forge any component of the system or falsify any information present in the system.

Security Objective: It should not be possible for anyone to alter any information in the system without proper authorization.

Security Objective: It should not be possible for anyone to gain, by working alone or collaborating with other parties, any privilege or rights other than those intended and provided for by the system.

Security Objective: It should not be possible for anyone to use the system for other purposes than those intended and provided for by the system.

Security Objective: Any party should not have to put more trust, when performing a given function, in another party than the minimum necessary to carry out the function.

Security Objective: The reliance on factors external to the system in the realization of any function within the system should be minimized.

Security Objective: An individual should have total control over which elements of his/her personal information are revealed in an information revelation procedure.

Security Objective: The disclosure of authenticated personal information should not need the revelation of the person's full digital identity.

Security Objective: The different disclosures of personal information performed by the same person should not be linkable.

Security Objective: There should be mechanisms to reveal the full identity of a person who is trying to misuse the system.

Security Objective: No person can register more than one digital identity in the system, except in extraordinary cases, provided for by the system.

Security Objective: The multiple identities of a person should not be linkable, except in extraordinary cases, provided for by the system.

4.8.2 Who do you trust?

Before assessing the security of a system, one needs to define the underlying trust model. Put in simple words, this comes down, for any party, to the following: Who in the system do you trust?

For instance, in the model presented in chapter 3, there is an implicit trust in an information authority with its legitimacy to certify personal information. This does not mean however that the individual should blindly trust it with respect to other functions, especially as regards surveillance issues: while certifying the credentials used by the individuals to reveal a profile, the information authority may add some extra information enabling the later tracking of the credential.

Also, for a two-party interaction like the profile revelation between a registered person

and an examiner, there is a natural distrust of each other. In the case of a voluntary profile revelation by the registered person, on one hand, the examiner may suspect the registered person of revealing a false profile, while on the other hand, the registered person may question the examiner's authorization to request the profile. For instance, should a national identification system serve for driving authorization, a police officer would be entitled to request a driving profile, while the tax government agency (the IRS in the United States) should not get access to any driving data, but only tax-related information. The management of different rights and privileges of examiners has been discussed already in sections 3.4.3 and 4.3.2, and we will explain in section 6.2.5 how current technology enables the two parties to still perform the aforementioned profile revelation while ensuring that the other party is not cheating.

4.8.3 The reliance on external systems

The vulnerable registration

When you first register in a national identification system, the identity authority issuing your digital identity authenticates the personal information it includes in your digital identity. To ensure the veracity of this information, it mainly relies on the following methods:

- It trusts its own perception and considers as true the characteristics it can directly determine. These characteristics are mainly physical attributes (such as gender, current height, eye color, etc), but may also be legal/adminitrative attributes (such as date of birth if the issuance occurs at the actual birth of the individual). It will thus authenticate these attributes.

- It relies on "documents" issued by a trusted party, and authenticates the characteristics they contain. For instance, it could rely on the parents' digital identities to determine the parents' names.

Consequently, there are two main ways to deceive the identity authority at the registration[5]: you can either deceive its perception or provide false documents.

[5]The question of the malicious behavior of a person acting on behalf of the identity authority during the registration process will be addressed later in section 4.8.5

- From wearing colored contact lenses to using gummy fingers, the methods for cheating biometrics are multiple, and rather affordable. A report by researchers at the Yokohama National University [118] for instance shows how one can produce very cheaply (a little ingenuity and US $10 worth of household supplies) "gummy fingers" that effectively simulate live fingers with respect to fingerprinting techniques.

- The problem of the reliance on the authenticity of other documents is different. While the issuing authority will have no hesitation for trusting documents it has issued itself earlier, relying on documents of uncertain source could be risky. The problem becomes all the more intricate when the document has been issued abroad, or by a foreign authority. The clear definition of what documents the authority should trust is no doubt an important task in the agenda of the law and policy makers.

4.8.4 Technical security

A system is not just the juxtaposition of its individual components: its security is not measured by the security of its weakest component, but the combination of the components can introduce new security risks that each component alone would not raise.

While the security of each individual component will be addressed when it will be analyzed, we investigate here the security vulnerabilities of the system as a whole.

Forgery

One of the most common ways to cheat is to forge a fake identity. As preventing forgery is one of the first and foremost security objective of any identification system, system designers are not likely to forget to assess the strength of their proposals against forgery.

The forgery of stand-alone documents has been a cat-and-mouse game for a long time. The efficiency of anti-counterfeiting techniques for current currency [131, 132] or IDs has been a well-studied area. We will therefore focus here on the aspects specific to an electronic national identification system: is it possible to falsify the digital information contained in the system? Since the digital information revealed to the examiner is contained in credentials, its unforgeability results from the security of the relevant credential protocols.

Identity theft

In our era of digital information, identity theft has become a more and more costly burden to society. What is identity theft? We define identity theft as the ability of a given individual to convince an examiner that he/she is someone he/she isn't. The claimed individual could be real or fictitious. The risk of identity theft is not specific to an identification system, yet the adoption of a national identification scheme would significantly increase the gravity of the consequences of a successful identity theft.

For more information about identity theft, we refer the reader to the webpage maintained by the US federal government [139] providing resources on the subject, and to the detailed explanation of the problem of identity theft by the US Federal Trade Commission [93].

Multiple identities

Related to the identity theft problem, the question of multiple identities is more problematic. Indeed, as mentioned in section 3.1.2, the ability for a person to have multiple identities may actually *not* be a feature one wants to prevent. While in the general case, one might want most people to have only one digital identity, but as mentioned in section 3.5.3, there might be exceptional cases where some persons may be allowed to have multiple identities.

The ability to create multiple identities therefore has to be a regulated privilege. The very existence of this needed feature introduces a major risk as regards the security of the entire system. How can one prevent the misuse of such a powerful right?

For a cryptology specialist, the answer is quite a reflex: if you want to divide too powerful a right, you use threshold cryptography. The founding principle of threshold cryptography is the following: instead of giving one party full latitude to perform a critical action, you divide that privilege among several people, a given number of which is necessary to carry out the above-mentioned action. Threshold cryptography will be explained in more detail in section 6.2.4.

This practice is actually not so uncommon: as popularized by the movie *Crimson Tide*, in order to launch a nuclear missile from a nuclear submarine, one needs the approval of both the captain and the second-in-command of the battleship.

4.8.5 The human factor

Although perfect technical security is an utopia that will never become true, the security of computer systems nowadays is less at risk because of technical reasons than because of human factors. Indeed, most of the successful attacks today are based on the exploit of these human factors. We will briefly analyze here the main forms of "human threat": collusion, insider attack, and social engineering.

Collusion

Collusion consists in the active collaboration of multiple parties in sharing their knowledge and privileges for purposes of deceit or fraud.

In a national identification system for instance, should the examiner pass on the information he/she got from a person's digital profile to the information authority that certified it, they could possibly get more information about the person than that intended, and maybe even identify him/her – provided that the information authority puts some tracking information in the digital profile. Another common example would be the case where the information authority would certify false information to a person for future revelation to an examiner to deceive him/her and make him/her grant undue privileges.

In fact, collusion is a major threat to computer systems and is actually addressed in most designs. Indeed, unlike for an insider attack or social engineering, current technology provides many means to limit – if not prevent – collusion. For instance, threshold cryptography (cf section 6.2.4) provides for the sharing of a privilege among many persons, thus preventing the collusion of a limited number of people.

Insider attack

Insider knowledge can prove to be extremely valuable in many situations. In a society where information and knowledge may be more worthy than any material good, man has created laws to regulate the use or sharing of this knowledge for "malicious" purposes: insider trading is for instance a punishable offense, many companies make their employees sign non-disclosure agreement for sensitive information, etc.

The core of the problem is how to prevent legitimate users of the system from improperly using the privileges they have due to their roles. If not monitored, what would stop a person

having some identity authority privileges from issuing to himself/herself profiles containing false information?

There are many known ways to limit the risks of an insider attack. Sharing privileges with parties of divergent or contradictory interests is quite effective, though not totally immune to collusion. Also, keeping an audit trail of all the operations performed in the system remains a sure deterrent to the evil temptations of insiders.

Social engineering

In the young history of the computer age, a good many hackers have tried to break in computer systems and quite a few succeeded. Kevin Mitnick, one of the most famous of them – if not the most famous – describes in his book *The Art of Deception: Controlling the Human Element of Security* [121] one of the main techniques he used, which he refers to as social engineering.

As he defines it, "social engineering uses influence and persuasion to deceive people by convincing them that the social engineer is someone he/she is not, or by manipulation. As a result the social engineer is able to take advantage of people to obtain information with or without the use of technology."

Why try to guess a password or code, when you can just ask for it? Why bother trying to find some technical vulnerability in a complex computer system to break into it, when you can just ask someone to let you in? Mitnick describes in his book in a very lively fashion many approaches and methods to achieve such feats. The most astonishing maybe is that these attacks would work on most sound – yet not computer-savvy – people.

When assessing the security of a computer system, one needs to keep in mind that it is just as good as its weakest link. And many times, this weakest link resides precisely in the very nature of human character.

4.8.6 Feature creep

Last but not least, a major source of security vulnerability in a system is what is referred to as feature creep: while designed to fulfill certain purposes, a system might be used for unprojected other purposes. Since these were not planned by the designers of the system, serious security vulnerabilities may arise from the misuse of the system.

The most preeminent examples of feature creep in the United States are probably the use of state driver's license as a de facto national ID (cf section 5.2), and the use of the Social Security Number as a Unique Identifier in too many databases of personal information (cf section 5.4).

4.9 Adoption and deployment considerations

When designing a practical national identification system to be in wide use, besides the technical robustness of the system, one needs to pay special attention to adoption and deployment issues.

4.9.1 Adoption process

First and foremost, should a government consider to establish a national identification system, the technical study is only a small part of a very long legal and political process. This process would probably involve the participation of technical and policy experts, as well as law makers. Also, it should leave room for an appropriate period of public consultation and debate. The matter of the adoption process will be further discussed in section 7.6.

4.9.2 Setup and integration of existing systems

Should a new technology-based national identification scheme come into effect, one needs to define what would become of the existing identification systems, and how the transition will occur.

The answer to that question depends highly on whether the system is to be compulsory or voluntary. Should it voluntary, provisions need to be taken for the old and new systems to coexist in harmony. Should it be compulsory, one needs to decide whether the new system is to replace the old ones or coexist with them. Even in the case where the new system is going to subsume some existing system(s), there needs to be a smooth transition period where both old and new systems would be in force.

4.9.3 Updates and extensions to the system

A new national identification system, if in use, needs to be upgradable easily and smoothly. If some cryptographic keys or functions need to be changed or some features

are to be added, those changes to the system need to take place without disrupting its normal functioning. In designing policies for updates of the system, one needs to remember that there needs to be backwards compatibility: some examiners and many individuals will possibly update their hardware/software components in the system only every now and then.

Also, as we will see in the next section (4.9.4), one needs to take into account the technological changes that will occur during the lifetime of the system.

4.9.4 Rapid changes in science and technology

Since the process of adoption of a national identification scheme is likely to be a long one, one needs to take into account the rapid changes in science and technology that are likely to occur between the initial design stage and the actual deployment. For instance, the most powerful personal computers now are virtually obsolete after a couple to a few years. Also, some of the techniques in the privacy-enhancing cryptography field presented in chapter 6 are fairly recent (1999 to today).

Although it is hard to predict what new scientific or technological breakthrough might happen in the near future, a large-scale project such as a national identification scheme cannot ignore the future technological landscape. A totally wired world with a pervasive network access, or the development of a very affordable handheld device assuming the functions of telephone, personal digital assistant (with calendar, address book, notebook, etc), are for instance in the realm of the possible, if not likely.

4.9.5 Cost

Recall the four generic parties we described in section 4.3: the individual, the examiner, the identity authority and the information authority. Ideally, each party should only bear the costs related to its role in the system. While the identity authority should assume the fixed costs of setting up and maintaining the system, the marginal costs could fall into the scope of the individual – for his/her own identity and profiles, as well as the physical/electronic support – and the examiner – for the device he/she uses to read the individual's profile.

Although most examiners will be organizations and thus may afford quite expensive equipment, one needs to remember that the actual cost, if too high, may become the first

– and maybe only – barrier to adoption. For instance, does a small liquor store in a small town really need to have an electronic ID reader, when most of its customers are local, and that checking ID is more of a legal constraint (and actually an inconvenience in the course of its normal business) than a way to get more business?

Finally, the cost to the individual needs to be minimal: people are likely to reject a new identification system they might be skeptical about, if it is going to cost them too much money.

4.9.6 Ergonomics and usability

We will see in chapter 6 that current technology enable lots of interesting features to be implemented in a national identification scheme. Nonetheless, the different protocols involved need to remain easy and convenient to use for the individual.

The designers of the system need to keep in mind that most people are not well-versed in technology. For instance, many people don't program their VCRs and avoid electronic devices as much as possible. A reasonable design goal is to make the system not much more difficult to understand than how to use a debit card for instance.

Also, as we have emphasized already, an identification system where the person would be in control of what personal information he/she wishes to disclose is in the domain of the possible. Should such a system be adopted, a good initiative would be to facilitate this disclosure by providing preprogrammed functions for some typical common scenarios, such as showing the minimum driving credentials, proving that you are over 18 (or 21) of age, etc.

4.9.7 Scalability

For a large-scale system such as a national identification system, scalability issues will inevitably arise: will the system support smoothly hundreds of millions of users and possibly tens to hundreds of billions identifications a day? However, the industry has extensive experience of large-scale projects, and what is more problematic here is how the system is going to be distributed to cater the needs of all the users.

The generic identity authority for instance is not a single party, but may actually consist of authorities at the federal, state and local level. Also, states may want to keep control of the issuance and management of the identities in the state (as they currently have over

driver's licenses and state IDs), while the federal government may just want to enforce a standard to diminish fraud for out-of-state identities.

Chapter 5

Analysis of existing applications of interest

In this section, we analyze briefly some existing applications that have aspects similar to the national identification scheme we consider. In particular, for each application of interest, we will focus on its strengths – especially the challenges common to both this application and a national identification scheme, that have been overcome – as well as its shortcomings – aspects that would need to be improved or avoided for a national identification scheme.

5.1 Passport

When reflecting on the design of a national identification system, one can examine the most widely used ID in the world: the passport. Although each country issues the passports for its own citizens, the design of all passports follows an international standard set by the International Civil Aviation Organization (ICAO) [15], a specialized agency under the United Nations. For more information on the exact specifications, the reader can refer to ICAO's Doc 9303 [17]. Also, a brief history of the passport can be found in ICAO's guide to Doc 9303 [12].

It is interesting to note that each country still keeps its own sovereignty by administering and maintaining its own system of passports while, by following an international standard, each country's officials can easily assess the authenticity of a foreign passport (as a physical document) while relying on the issuing country for the authenticity of the information

contained in the passport[1].

This is a remarkable working distributed system, where the different sovereign members (countries) share the responsibility of issuing and verifying the individuals' identities. An apt national identification system could benefit greatly from the careful analysis of the passport system, especially from a legal point of view: who issues the passport? who is ultimately accountable for the authenticity of the passport? on what basis is the passport issued to a requestor? how is an individual authenticated to be who he/she pretends to be in the first place before he/she is issued his/her first passport? etc.

The historical success of the standardization of the passport shows that even countries with very divergent interests[2], different laws and traditions, could agree on a common international standard for identification for travel purposes. A set of basic demographic identifying information was agreed upon, and language barriers were overcome.

The example of the passport proves that the legal structures and policy regulations for a national identification scheme need not overshadow the sovereign rights of the states/provinces in a federal country, especially regarding the administration of the "identities" of their people. For instance, the federal government could decide on the standards to be followed, while each state would be sovereign in issuing and managing its own "identities".

5.2 State driver's license (in the United States)

When asked to show identification, most US residents show their driver's license. In the absence of a national ID card in the United States, although it was at first only designed to authorize people to drive, the American state driver's license has become the most widely ID used in the United States, and thus a de facto general-purpose ID. This is a very obvious example of feature creep: designed to meet the needs of a license to drive (and thus with the corresponding security implications), its use has drifted to include general-purpose identifications. Also, in many states, a driver's license can even be withheld for a wide number of offenses not related to driving at all.

The American state driver's license suffers from the following drawback: since each of

[1]In the case where a country does not totally rely on another for the authenticity of the information contained in the passport, it can require the possession of a visa delivered by its own immigration services. Some further background checks can thus be performed on the person prior to the issuance of the visa.

[2]These countries could even be political enemies: recall that the ICAO was founded in 1946, right after the worst war in human History.

the 50 states is responsible for issuing and managing its own set of driver's licenses (and other state IDs for non-drivers), IDs from different states follow different designs, and it is hard for anyone to detect fake out-of-state IDs. Besides, many people have taken advantage of the system to get a new driver's license with a "clean" record in another state if they have a bad record in some state.

The American Association of Motor Vehicle Administrators (AAMVA), which is "is a voluntary, nonprofit, tax exempt, educational organization, [...] represent[ing] the state and provincial officials in the United States and Canada who administer and enforce motor vehicle laws"[3], has tried to address these different issues by issuing a proposal for a national standard for driver's licenses and identification cards [130].

In my opinion, this proposal suffers from some major flaws:

- The AAMVA tried to address the security issues raised by the tragic events of September 11, 2001 by proposing a hasty fix – mainly driven by operational considerations, such as security, interoperability, efficiency – without considering the consequences of its adoption in the long run, especially regarding privacy issues.

- The purpose of the Departments of the Motor Vehicles (DMVs) – who currently issue driver's licenses – is to administer licenses to drive, not national identification cards. The establishment of a national identification system is well beyond the jurisdiction of the state DMVs, and even more of the AAMVA which is not even a government agency: their contribution to the matter should be limited to an advisory role regarding the more specific driving issues.

- Whether the driver's license should serve as the basis for a national identification system is actually a question that deserves to be addressed on its own.

EPIC gives a more detailed assessment of this proposal [85] to transform the state driver's license into a de facto national ID card, and presents a very compelling case *against* it. We will retain of this report the following credo: "there must be a full assessment of the risks and consequences of a system of national identification in the United States. Appropriate legal and technical safeguards should be established before should a project goes forward." Our thesis aims at contributing to this assessment.

[3]Description taken from http://www.aamva.org/about/.

5.3 International driving permit

The situation about international driving is quite confused. Indeed, each country has its own driving regulations, and requirements for obtaining a driver's license, and these may even change from state to state in a federal country like the United States. Then, if you obtained a driver's license in a country, are you authorized to drive in another? In the United States, the states have their own driving rules, and each state grants driver's licenses to its residents if they pass a series of tests, specific to the state. However, the American state driver's license entitles you to drive anywhere in the United States.

The United Nations clarified the situation at the Convention on Road Traffic in 1949 by creating the International Driving Permit (IDP), often also called an international driver's license[4]. Yet it is *not* a driver's license. It does not have any legitimacy on its own: its only purpose is to provide a translation for your own country's driver's license, and has no value unless accompanied by the corresponding "real" driver's license.

Unfortunately, its ambiguous name has created an undesirable situation: as described by the US Federal Trade Commission [67], many companies have exploited the ambiguity and set up scams to sell fake international driver's licenses: they claim that the international driver's license they sell subsume a national (or state-issued in the case of the United States) driver's license, and charge "from US $65 to US $350" for these fake documents, while the legitimate document can be acquired for a mere US $10. Although a wise person wouldn't buy a legal document from a private company, their claim sounds sensible: if a passport can in many countries be used internally as an ID – thus subsuming one of the main purposes of a national ID, why wouldn't an international driver's license subsume a national one? Besides, many such companies go even further by putting a reference to the United Nations' Convention.

Although the International Driving Permit is a totally legal and valid document, its ambiguous name and the lack of information about it have led to a confusion about international driving. This is an unfortunate example where poor policy introduces the possibility of serious exploits of the system.

[4] A FAQ on the International Driving Permit can be found at http://www.aaa-calif.com/travel/idp.asp.

5.4 Social Security Number in the United States

Designed at first to be an identifier for the use of the US Social Security Administration [26], the US Social Security Number (SSN) has become a de facto identifier. As explained by Garfinkel [95], not long after its creation, the SSN started to be used by many federal and state organizations as a Unique Identifier[5] on for their records to save the costs of designing and maintaining a new one.

Although the *Privacy Act* of 1974 now regulates the use of the SSN by federal, state, or local government agencies[6], the SSN is still present in many a database, and represents the quintessence of feature creep, as well as the origin of many a privacy infringement: the very nature of identifier of the SSN makes it an ideal tool to link databases of personal information.

The controversy about the misuses and abuses of the SSNs stigmatizes the privacy implications of the existence and wide use of a visible Unique Identifier. These matters have been well-documented: CPSR maintains a comprehensive FAQ section on SSNs and privacy [90], while the Privacy Rights Clearinghouse's FAQ [63] focuses on the individual's rights regarding SSNs and Network USA's FAQ [156] is a strong indictment of the use of SSNs as a numbering tool. As for any important policy issue, EPIC maintains a resource page on SSNs and privacy [83]. Also, for technical information, the reader may refer to the Social Security Administration's official website [26], which contains also a basic FAQ section [30].

5.5 Various industry initiatives on the Internet

5.5.1 Microsoft .NET Passport

Brief overview

Microsoft .NET Passport [18] is a suite of Web-based services launched by Microsoft in 1999. It aims at facilitating the user's experience by providing the following services: a single sign-in (SSI) feature, "Passport wallets" and Kids Passports. Although the third

[5] We refer the reader to section 4.7 for a more detailed treatment of Unique Identifiers.

[6] "It [is] illegal for federal, state, and local government agencies to deny any rights, privileges or benefits to individuals who refuse to provide their SSNs unless the disclosure is required by Federal statute, or the disclosure is to an agency for use in a record system which required the SSN before 1975." [90]

service is a very laudable initiative to comply with the *Children's Online Privacy Protection Act* in the United States, we will focus here on the first two, which form the core of the suite.

The SSI feature enables the Web user to authenticate to all Passport-enabled websites using the same Passport. A Passport can contain as little as an email address, which will serve as the user sign-in name, and a password. The user can also include some basic demographics information in his/her profile. When authenticating to a Passport-enabled website, he/she can choose to either share his/her whole profile, his/her first and last names and email address, only his/her email address, or even no information at all (in which case, the website only knows that the user has a Passport[7]).

The Passport "wallet" contains additional information used for online purchases, including name, telephone number, credit card information, and billing and shipping addresses. At his/her choice, the user can decide to share his/her "wallet" information with a Passport-enabled website, thus avoiding retyping it.

For a more complete description of Microsoft .NET Passport, we refer the reader to the official Review Guide [20].

Benefits

The benefits for the user are mainly the ease of use and convenience of dealing with a single authentication procedure. Besides, the Passport is not tied to a single machine, but he/she can use it with any machine connected to the Internet. Most importantly, he/she is in control and decides what information he/she wants to disclose to a website he/she visits. For more information regarding what personal information is disclosed by Microsoft .NET Passport, please consult Microsoft .NET Passport's Privacy Statement [19]

Yet, the main beneficiary of the Passport is the website using the feature. While outsourcing the delicate problem of authentication of the users to a renowned company, they can leverage the Passport database by using the unique Passport ID (PUID). The PUID is a unique identifying number for the Passport system, that is transmitted to any Passport-enabled website the user visits when he/she is signed in with his/her Passport.

[7]We will see shortly to what extent this is true.

Summary

The very existence of the PUID is a major threat to the user's privacy. When using his/her Passport, the user has the impression that he/she can disclose whatever information he/she wishes to the websites he/she visits. However, using his/her PUID, different Passport-enabled websites can share information about him. Although Microsoft forces websites using his/her Passport to agree not "to assign, transfer, share, transmit, or publicly disclose Passport profile information [...] to any third party without the user's consent" [20], it is unclear what sanctions they are facing if they do. Besides, this does not apply to personal information they get by other means.

While being an interesting initiative at facilitating the user's experience and preserving his/her privacy, Microsoft .NET Passport fails at attaining this objective: just by using the unique Passport ID, Passport-enabled websites can efficiently link their databases; and they are outside the legal privacy restrictions imposed by Microsoft when sharing non-Passport information such as purchasing history, musical preferences, etc. Unlike the SSN which use is somehow regulated (cf section 5.4), the PUID can serve as a Unique Identifier for all Passport users who cannot keep it secret.

Also, the initiative suffers from the lack of legal structures: there is little Microsoft can do to force a Passport-enable website to enforce a security policy. As mentioned in page 14 of the official Review Guide [20], there is a possible security breach if the Passport-enable website does not require the user to reenter his password when accessing sensitive information: if the user forgets to sign out of his/her Passport, another person using the same machine later could access this sensitive information on the website while being authenticated as the user.

In summary, while being a success at attaining its technical objectives, Microsoft's .NET Passport fails to achieve its purpose due to some policy and legal shortcomings. This example highlights the need for proper policy and legal structures to accompany the technological achievements.

5.5.2 Liberty Alliance Project

The Liberty Alliance Project [16] is an industry initiative with roughly the same goals as Microsoft's .NET Passport: facilitating the Web user's management of his/her personal

information on the Internet (online identity, digital profiles, personalized online configurations, spending habits and history, shopping preferences, etc). The main features the project wishes to provide are: opt-in account linking across participating websites, simplified sign-on for linked accounts, authentication context (where organizations would communicate the type and level of authentication that should be used when the user logs into different accounts), and global log-out. It released its first version (1.0) of its specifications on July 15, 2002, which was updated after a period of public comment into a new version (1.1) released on January 15, 2003.

The Liberty Alliance Project however differs substantially from Microsoft's .NET Passport:

- Unlike Microsoft whose aim was to ship as soon as possible an operational product targeted mainly at corporations, the Liberty Alliance Project is an open standard, which aims at complying with a wide range of operating systems, programming languages, and network infrastructures. In particular, to prevent haste, there is no tight timeline for the project.

- The Liberty Alliance Project does not rely on a single, central identity authority, which would provide for the sharing of the individual's personal information. It envisions a distributed network of multiple identity authorities maintaining "network identities" of individuals, which can be linked to the accounts maintained by service providers.

- The Liberty Alliance Project considers the privacy of the users to be one of its top priorities, and thus pays appropriate attention to privacy issues in its design.

The Liberty Alliance Project has all the makings for a successful online identification system, but lacks two essential features needed in a national identification system:

- The project only provides the user with the option of linking separate accounts. In particular, the user has no control about what information is exchanged between the corresponding organizations. These organizations will then, forever since the linking, view the individual as being the same whenever he/she is on any of the organizations' websites. In particular, one can assume that the organizations concerned fully merge the personal information they have about the individual.

- The project relies mainly on the respect of private privacy policies. In particular, the individual needs to trust the identity authorities to which he/she delegated the management of his/her network identities. Besides, it is unclear to what extent these identity authorities are legally accountable for the misuse or abuse of this entrustment.

However, this project is a good example for the actual design (and definition of specifications) phase of a project prone to open discussion and/or debate. In particular, the timeline adopted (with the allotment of a period of six months after the release of the preliminary specifications for public comment) is rather appropriate. This public consultation however only involves technical issues, and a suitable time allocation for public debate on a preliminary design proposal for a national identification scheme is likely to be much longer, since it involves also the participation of policy experts and lawmakers.

To get more information on the Liberty Alliance Project, we refer the reader to the Project's website [16], containing a good FAQ section introducing the Project in broad outline, as well as the actual Liberty Alliance specifications.

5.6 Identification of objects: from bar codes to RFID technology

While identification of human beings has been a controversial topic, there is another domain where technology has been used for decades for identification: the identification of commercial goods.

The bar code, invented in the early 1950s, has revolutionized the way many industries are conducting business. The bar code enables the identification of products by machines. A bar code can identify a series of items presenting the same characteristics (such as most industrial goods one can find in a supermarket) or a unique one (such as bar codes present on documents and packets from shipping companies). Nowadays, any industrial good – from food items to books through consumer electronics products – has a bar code identifying its producer, category, and type of product.

Nonetheless, the Automatic Identification and Data Capture (AIDC) industry is now increasingly using a new technology: Radio Frequency IDentification (RFID). Unlike bar codes, RFID tags are wireless systems that allow for non-contact reading and are effective in manufacturing and other hostile environments where bar code labels could not survive.

Common commercial applications using RFID technology include electronic article surveillance, electronic toll collection, and tracking.

Just like the bar code, one of the main benefits of RFID technology is its incredibly low cost. RFID is thus a mature technology offering a very affordable solution for identification in many applications.

RFID technology is however well-suited for applications needing absolute and precise identification, such as tracking. At first, one would not imagine using RFID tags to identify human beings in a national identification system. After all, we don't want to be just another serial number. However, its strengths – the possibility of transmitting some information in a limited range without requiring line of sight, its very reasonable cost[8] and above all the ability to turn the RFID tags on and off – may prove to be useful in some limited situations.

Although RFID technology would not be of interest if used as is in a national identification system, it is definitely a technology that deserves attention and may be used to perform some specific functions within the system.

For more information on current aspects of these technologies, we refer the reader to the Association for Automatic Identification and Data Capture Technologies (AIM), "the worldwide authority on automatic identification, data collection, and networking in a mobile environment"[9]. Its website contains a section on bar code technology [32] as well as on RFID technology [33]. The latter is described as "a link to happenings in the RFID world"[10]. It also contains a short history of RFID technology [24].

Also, the Auto-ID Center [2] "is a unique partnership between more than 87 global companies and three of the world's leading research universities" aiming at "creating the standards and assembling the building blocks needed to create an "Internet of things""[11].

5.7 Authentication of computer machines and agents

In our modern society, computer systems are becoming more and more ubiquitous. Many individuals and organizations rely today on computers to operate their businesses. Yet, with the increased connectivity of computer systems through local networks or the

[8]The Auto-ID Center (http://www.autoidcenter.org/) is currently working on the feasibility of producing RFID tags for 5 cents.

[9]Description taken from http://www.aimglobal.org/aboutaim/.

[10]Description taken from http://www.aimglobal.org/technologies/rfid/.

[11]Description taken from http://www.autoidcenter.org/aboutthecenter.asp.

internet comes an increased risk of obtaining untrusted software or data. The various risks of running malicious software or manipulating suspicious data range from installing and running viruses, worms or other trojan horses, to compromising the sensitive data on the computer or the whole computer itself. The increased need for an authentication of computer machines and agents has led to various industry initiatives.

We will present in this section the SSL/TLS standard for authenticating (and securing) communications over the internet: SSL/TLS enables a reciprocal authentication of the Web server and the Web client in communication. We also examine the TCPA and Palladium projects aiming at providing for the authentication of computer systems and programs.

5.7.1 Secure Sockets Layer/Transport Layer Security (SSL/TLS)

As the internet becomes more and more pervasive, many organizations have tried to leverage the value of this network to build various applications. However, the decentralized nature of the network makes the internet an insecure communication channel. To meet the need for authenticated and secure communications, Netscape designed and developed in 1994 the Secure Sockets Layer (SSL) protocol. SSL is a Web protocol for establishing authenticated and encrypted sessions between Web servers and Web clients. Since its inception, SSL has been widely used over the internet for applications ranging from online banking to e-commerce. Following the large success of SSL, the Internet Engineering Task Force (IETF) worked on standardizing the protocol and published a new proposed standard based on SSL: the Transport Layer Security (TLS) protocol. SSL/TLS is nowadays an internet standard.

In SSL/TLS, the authentication of the two parties (the Web server and the Web client) relies on certificates issued by a trusted certificate authority. Each party can decide whether or not to trust a given certificate authority.

For technical information about SSL/TLS, we refer the reader to Netscape's introduction to SSL [122], to IETF's charter on TLS [107], and to Eric Rescorla's book *SSL and TLS: Designing and Building Secure Systems* [137].

5.7.2 Trusted Computing Platform Alliance (TCPA) and Palladium

The TCPA and Palladium[12] projects, announced in 1999 and 2002 respectively, aim at providing for the authentication of computer systems and programs. The Trusted Computing Platform Alliance (TCPA) is an industry work group focused on enhancing trust and security on computer platforms. "Palladium is the code name for an evolutionary set of features for the Microsoft Windows operating system. When combined with a new breed of hardware and applications, these features will give individuals and groups of users greater data security, personal privacy, and system integrity."[13].

Through the use of a tamper-resistant Trusted Platform Module (TPM), which is uniquely bound to a single platform, TCPA[14] provides for platform authentication and attestation: the platform properties[15] can be attested to any challenging party. Also, TCPA can reliably measure and report on the platform's software state. These two functionalities enable the attestation of a given hardware/software configuration. Finally, TCPA provides the means for protected storage: the TPM can tie sensitive data to a specific computer process (or a group of them) on a given platform.

Palladium[16] itself offers a protected trusted environment, highly resistant to tampering and interference, for running applications. Relying on TCPA or similar hardware, a computer platform running Palladium will be able to authenticate the software it is running – for its own use, such as assigning adequate rights to these processes, or to a remote client. Also, like TCPA, it enables the authentication of a software configuration and a sealed storage, tied to a given computer process on a given platform. Finally, Palladium provides for a secure input and output for the user.

Applications for TCPA / Palladium are numerous: from digital rights management to the protection of confidential data, to the strict association of data or processes to a given configuration, these technologies enable a totally new range of possibilities. In particular, as we mentioned in section 3.1.6, by using computer agents/proxys, this may allow for a novel means for human identification.

[12]Microsoft discontinued the use of the term Palladium for its project: Palladium is now .

[13]Description taken from Microsoft's official Palladium webpage [21].

[14]We base our analysis partly on a presentation of TCPA by Joe Pato, HP Labs, given at MIT on October 17, 2002.

[15]Note that no platform-identifying information (such as a serial number) is disclosed by TCPA: TCPA only provides for the attestation of platform properties, to ensure the platform is in some expected state.

[16]We base our analysis partly on a presentation of Palladium by Brian LaMacchia, Software Architect for Windows Trusted Platform Technologies, given at MIT on October 17, 2002.

Nonetheless, many privacy activists have started to strongly oppose TCPA [1, 23] on privacy grounds. Indeed, the privacy concerns over a national identification system also apply to the identification of machines. And since a computer system may be tied to a person, or group of persons – its user(s), the privacy threats are affecting in turn these people.

For more information about the TCPA and Palladium projects, we refer the reader to the TCPA official website [27], the Palladium official website [21], and Ross Anderson's FAQ on TCPA/Palladium [34].

5.8 Linkage of databases: credit reporting, Total Information Awareness (TIA)

In a society where information has become a very valuable commodity, virtually any organization maintains one or several databases to record the information it possesses. The value of this information to the organization increases when it can be linked to relevant information from other information systems. This represents a major privacy risk of any information system – and of a national identification system in particular. We briefly present in this section the case of the credit reporting industry in the United States, which first raised the privacy issues relating to computerized information and databases and led to a series of privacy policies and laws, to illustrate the need for appropriate policy as regards privacy. We also examine the TIA project initiated by the US Department of Defense, which aims at analyzing aggregated data gathered from multiple sources to derive some useful information (for fighting terrorism). We pay special attention to the Department's concern about privacy and its effort to research and develop technologies for monitoring and controlling the access to the data in the system.

5.8.1 The credit reporting industry in the United States

In the United States, any decision regarding the granting of credit is based on the person's credit history. Credit bureaus maintain a list of the person's credit accounts – credit cards, mortgages, loans, etc – and aggregate credit data kept by the various credit institutions, which the person is related to. These companies then sell credit reports to credit institutions, which use this information to decide whether to accept an application

for credit (and under what terms if so) or to try to attract new potential customers.

This financial information is very sensitive, insofar as credit plays a central role in the United States' modern society. As Garfinkel writes in his book *Database Nation* [95], "in a society where credit is required by all but the richest families to buy a house, to buy or lease a car, or to get an education, denying somebody credit effectively denies that person the privileges of being a member of society."

The computerization of the data in the late 1960s definitely changed the industry: any personal data could now be stored virtually indefinitely and would be very easily accessible later. This introduced many new problems and technology-related policy issues: Who has access to the information? For what purpose(s)? What happens in case of an error in the data? etc. This also raised many privacy concerns, which would be settled by the *Fair Credit Reporting Act* in 1971.

The credit report example remains the quintessence of the use of databases of personal information, and of the possible dire privacy consequences of errors and abuse (in particular, identity theft has plagued the industry for decades).

For a more thorough discussion of the privacy implications of databases, we refer the reader to *Databanks in a Free Society: Computers, Record Keeping and Privacy* [160] by Alan F. Westin and Michael A. Baker and *Database Nation* [95] by Simson Garfinkel.

5.8.2 Total Information Awareness (TIA) System

The Total Information Awareness (TIA) program is a research program initiated by the Defense Advanced Research Projects Agency (DARPA), the central research and development organization for the US Department of Defense (DoD). As described on the FAQ [10] on its official website, the goal of TIA is "to create a prototype network that integrates innovative information technologies for detecting and preempting foreign terrorist activities against Americans". More precisely, the TIA system is "an experimental prototype system that consists of three parts: language translation technologies, data search and pattern recognition technologies, and advanced collaborative and decision support tools." [11].

The TIA program is not an effort to create a gigantic database of personal information to help fight terrorism, but rather is aimed at developing appropriate technology to leverage the data currently available from the many databases maintained legally by US intelligence, counterintelligence, and law enforcement agencies (and possibly from commercial and public

databases as well).

If the preservation of privacy is a major goal of a national identification scheme, one needs to consider what information could be derived from it by the use a system such as TIA.

On the other hand, the program also plans to "research and develop technologies to protect the system from internal abuses and external threats. The goal is to achieve a quantum leap in privacy technology to ensure data is protected and used only for lawful purposes" [11]. This research effort include the evaluation of some of the Information Science and Technology (ISAT) Panel's *Security with Privacy* study's recommendations, such as immutable audit and self-reporting data. While immutable audit enables a unforgeable tracking of all activities regarding the data present in the system, self-reporting data would allow for auditors to know who accessed it. The results of the TIA program in this area could be used for enhancing the privacy of the personal information of a national identification system.

For more information about the TIA program, we refer to reader to its official web-page [74], and to EPIC's section on TIA [84].

Chapter 6

Science and Technology

In this section, we explore briefly the current knowledge in Science and Technology – cryptology especially – that is relevant to national identification systems. We start by providing a brief history of cryptology, especially with respect to how it became a fundamental Science in the security domain. We then proceed to present succinctly various areas of cryptology and how they contribute to achieving some of the goals of a national identification system. In particular, we insist on a recent research area of particular relevance: privacy-enhancing cryptology. We conclude by providing general resources for the reader more interested in the scientific or technical details of cryptology.

6.1 The role of cryptology in the security domain

First and foremost, the security of a national identification system lies in the confidentiality of the information it contains. The encryption and decryption of information has long been the main object of cryptology. Today however, cryptology has evolved into a more comprehensive science enabling a vast range of applications, ranging from digital signatures to credentials through secure multi-party computation and other zero-knowledge proofs.

6.1.1 A little bit of History

Codemaking and codebreaking: the dawn of cryptology

Since antiquity, the ability to keep secrets has been essential to mankind. Quite early, man has realized that the protection of information was much different from that of material

goods (such as treasures, relics, weapons, etc): the actual value of information does not reside in its physical support, but in its contents. For 4,000 years, man has created codes to protect his secrets, and tried to break codes to learn others' secrets. Codes enable the protection of information, should its physical support be compromised. Codemaking and codebreaking have been at the heart of civilization's secret and intelligence history and have been critical to governments in times of war (for military purposes) as well as in time of peace (for diplomacy and intelligence purposes).

Cryptology – cryptography (codemaking) and cryptanalysis (codebreaking) – has played an essential role, too much overlooked (or on the contrary overemphasized), in human history on many occasions. For instance, the role of cryptology (and cryptanalysis in particular) has been evident in World War II. The abilities of the Allies to break German and Japanese codes indisputably gave them an invaluable edge. A striking example for this is the cryptanalysis of Enigma. The Enigma machine was used by the Germans during World War II to secure their communications, with great effectiveness. However, by the end of war, the Allies' cryptanalysts led by Alan Turing were able to decrypt most of the Enigma traffic, and contributed to the success of the Allies' forces in Europe. The interested reader can find more information about Enigma on the internet [9, 29].

To get a more precise account on the history of cryptology and its influence on human history, we refer the reader to the excellent (and comprehensive) book *The Codebreakers: The Story of Secret Writing* by David Kahn [112].

The machine revolution

Over the centuries, cryptography has become a complex art of designing robust codes, and cryptanalysis has been enriched with ever more sophisticated techniques to break them. However, cryptology – cryptography and cryptanalysis – long relied on the ingenuity of its practitioners.

The advent of machines a century ago dramatically changed the situation. Indeed, machines can perform computations much faster than any human could do, and with many fewer errors. While enabling the use of more complex codes, it has drastically transformed cryptanalysis as well. Cryptology would not be performed by humans any more but by machines.

Cryptology has gone public

Historically, cryptography has been used extensively to protect critical information and secure secret communications for military and intelligence purposes. Until recently, advanced cryptology had been the private domain of military, government and intelligence agencies.

The public presentation of the invention of public-key cryptography by Diffie and Hellman [80] in 1976 and of the new RSA algorithm [141] in 1978 sparked off a controversy with the US National Security Agency (NSA), and the matter of academic (civilian) research on cryptography in the United States would be discussed for years at the highest level. The debate ended when the *Computer Security Act* in 1987 granted to the National Bureau of Standards (NBS) – now the National Institute of Standards and Technology (NIST) – authority regarding standards and guidelines on public (civilian) cryptography over the NSA. Nowadays, much of the cryptographic research in the United States is now done freely by academic institutions and corporate research laboratories. For more information on U.S. crypto policy, the reader may refer to the ACM U.S. Public Committee (USACM)'s report on the subject [157].

Today, in the United States, the Computer Science and Telecommunications Board (CSTB), an operating unit within the National Research Council (NRC), "provide[s] independent advice to the federal government on technical and public policy issues relating to computing and communications". In particular, it issued in 1996 a report on crypto policy [72].

Modern cryptology

The development of academic research on cryptology has dramatically changed the field: the application of scientific methodology and the collaboration with scientists from other fields have contributed to reshape this former art into a new science. Cryptography does no more rely on art or talent than cryptanalysis does on fortunate intuition, but rather on solid scientific grounds. Encryption and decryption are no more esoteric processes imagined by some inspired savant, but well-defined algorithms, which "security" is based on provable mathematic properties and theorems. Also, usage has pushed towards "public" algorithms, whose details are publicly published and which security is well-documented, and the "secret"

now rests with the use of secret keys provided to the algorithms to encrypt and decrypt the documents.

Cryptology has now become a respectable subfield within Theoretical Computer Science on its own. Furthermore, cryptology does not only address encryption and decryption techniques any more, but has grown to also encompass many other fascinating and useful applications, some of which we will review now.

6.1.2 The theoretical foundations of modern cryptology

Digital information and information theory

One of the main contributions of the scientific methodology to the field is the ability to define and quantify more precisely the notion of security.

In his seminal paper *A Mathematical Theory of Communication* [147] in 1948, Claude Shannon introduced for the first time the notion of bit - binary digit - and digital information, and founded the subject of information theory. The following year, in 1949, in his other seminal paper *Communication theory of secrecy systems* [148], he laid the foundations for a more formal treatment of cryptography, based on information theory.

As presented by ETH Zurich's Information Security and Cryptography research group [71]:

> "There are two types of cryptographic security. The security of a cryptographic system can rely either on the computational infeasibility of breaking it (computational security), or on the theoretical impossibility of breaking it, even using infinite computing power (information-theoretic or unconditional security)."

Information-theoretic security

The notion of information-theoretic security or "perfect security" was introduced by Claude Shannon in his seminal paper *Communication theory of secrecy systems* [148]. A system is perfectly secure if an adversary with unlimited time and manpower/computational power cannot break it. In the same paper, he proved the main result in the domain: an information-theoretic encryption scheme needs a key as least as long as the plaintext document.

A practical encryption scheme still in use today that is information-theoretically secure

is the use of one-time pads. To encrypt a message, you "combine" it[1] with a random message (the pad) of the same length to produce the ciphertext; to decrypt the ciphertext, you "de-combine" it [2] with the pad to obtain the plaintext message. This encryption method is perfectly secure provided that each pad is used only once (hence the name).

However, the apparent necessity for a long and new key at each operation makes the scheme quite impractical to use. Therefore, most of the cryptologic research nowadays focuses on computational security, which provides imperfect yet adequate practical security. Nonetheless, there is still ongoing research on various aspects of information-theoretic cryptology at ETH Zurich [71], and Ueli Maurer [119] gave in 1999 an overview of the current knowledge in the field.

Computational security

The nature of computational security is to make the breaking of a system too costly so that no practical computer system could actually break it. Most cryptographic systems have actually a customizable security level: their security depends on a security parameter that can be set according to the application needs. For example, an extreme case would be to set the computational hardness to break the system so high as to require the use of all computers in use today for the next century. Yet, higher security may lead to a lower efficiency for the system. Thus, there is a tradeoff between high security and high efficiency that needs to be decided upon the needs of the application concerned.

Also, we would like to stress that computational security is probabilistic by nature. The security of a typical system could translate as follows: it is as unlikely to break the system in less than 100 years by using all the computer resources now available as to win the grand prize at the lottery 10 times in a row. Although this is not 100 % sure, once again, it is more than sufficient for all applications.

Finally, the reader should know that the computational security of most cryptographic schemes and systems relies on some standard intractability assumption, believed to be true by the research community, yet unproven so far. The most common assumptions used in cryptology include:

- $\mathcal{P} \neq \mathcal{NP}$. We refer the reader to Michael Garey and David Johnson's seminal paper

[1]The operation used here is the bitwise XOR (exclusive or).
[2]The operation used here is also the bitwise XOR (exclusive or).

Computers and Intractability: A Guide to the Theory of NP-Completeness [94] for a more detailed treatment of this problem.

- the existence of one-way functions[3].
- the existence of one-way trapdoor functions[3].
- the one-wayness[3] (or some other property) of the modular exponentiation and RSA.

The role of assumptions in Science and Technology

In this section, we will try to convince the reader that relying on unproven intractability assumptions does not undermine the practicality of computational cryptography.

The use of assumptions has been playing a central role in the development of Science. Prior to any proof of a result or theorem, lies the assumption of its veracity. Discoveries in Science have always preceded their formal proofs – sometimes by a long period of time.

Assumptions are even at the heart of many sciences. The very nature of Physics for instance is to establish (unproven, thus assumed) laws to describe the reality of our environment. A law in Physics is believed to be true if we can positively verify its consequences experimentally.

Let us look at the History of Mechanics. In the late 17^{th} century, Newton established his laws of motion, which led to the theory of universal gravitation, and gave birth to Classical Mechanics. Centuries later, Classical Mechanics was discovered to be only a good approximation of "real" Mechanics, more generally described by Quantum Mechanics. Although Classical Mechanics has been shown to be inexact, scientists and engineers throughout the world have been using it and still use it nowadays. For instance, the robustness of our buildings rely on the (inexact) laws of Classical Mechanics. Yet, we go to bed every night without the single concern about it: although not perfectly exact, these laws are sufficient for this domain of application.

Buildings have been built for millenniums, long before Newton's theory of classical Mechanics. Architecture and civil engineering have long preceded the theory of their underlying science (Mechanics) in the same way cryptology has developed long before the establishment of its theoretical foundations. It is in the very essence of scientific discovery to make

[2]We refer the reader to some general reference on cryptography for an explanation of these concepts. A list of good such references can be found in section 6.4.

plausible assumptions, and there would have been much fewer technological inventions if man did not rely on unproven hypotheses or other heuristics.

As of today, our current technology and many aspects of our lives still rest upon unproven "facts". Do objects fall to the ground when thrown in the air? We all take this as granted, yet it has never been "proved". Is the Earth round and does the Earth rotate around the Sun? This is a fact nobody disputes, but only a few centuries ago, anybody would have laughed at you if you pretended it, including the majority of the most learned scholars.

The assumptions at the heart of modern cryptography are not much different from the assumed laws of Physics. Everybody in the research community believes in these intractability assumptions, and most of our security technology already relies on them. The bottom line is that they are adequate for the security applications in use today.

The reader should be prudent though. Not all cryptographic results are based on the same assumptions. Some assumptions are stronger than others, and quite a few results rely on non-standard assumptions. The results presented here however are based on the more standard assumptions mentioned above, or on some common variants.

6.2 Numerous aspects of modern cryptology

6.2.1 Public-key cryptography

For a long time, encryption and decryption have been using encryption and decryption keys that needed to remain secret between the two parties who wish to communicate securely. In 1976, in their seminal paper *New Directions in Cryptography* [80], Diffie and Hellman introduced the concept of public key cryptography, which constituted a major breakthrough in the field.

Until then, to exchange encrypted messages with your friends for instance, you first needed to agree with each and every one of them on a common secret key *before* being able to send the first message. Besides, you needed to agree on a different secret key with each one of them, unless you want your conversation(s) with one friend to be decipherable by another. One can see that this scheme becomes highly impractical in concrete situations as soon as the number of parties involved grows.

Diffie and Hellman proposed a system in which each individual participating in the system would be issued a pair of keys: one public and one secret. To send a message

to your friend Bob, you would encrypt it using Bob's public key; upon reception of the message, Bob would decrypt it using his own secret key. Note that a fundamental property of such systems is the inability for an adversary of finding out the secret key corresponding to a public key. In this system, a trusted authority would maintain a directory of all public keys.

Not only do you not need to keep track of the keys of all your partners any more, but you could send an encrypted message to a person whom you have never been in contact before by looking up his/her public key – provided that you trust the authority that is depository of the public keys.

The RSA algorithm introduced in 1978 by Rivest, Shamir and Adleman [141] would give to Diffie-Hellman's vision the first practical implementation. Many public-key cryptosystem nowadays are based on the RSA algorithm or the discrete logarithm problem introduced by Diffie and Hellman [80].

Although public-key cryptography seems more attractive than secret-key cryptography, the latter is still widely used as it is much cheaper in resources. In many practical encryption/decryption applications, public-key cryptography is only used to securely exchange a shared secret key, through a key exchange protocol, which will be used for the actual encryption/decryption.

6.2.2 Probabilistic Encryption

Shafi Goldwasser and Silvio Micali [102] introduced the concept of probabilistic encryption in 1984, along with a practical example of such a scheme. Given a message and an encryption key, instead of always encrypting the message into the same ciphertext (which a deterministic encryption scheme does), it produces a "random" ciphertext, which decryption yields the original plaintext message.

This notion is fundamental in many practical applications. Suppose the encryption algorithm (and key) is public, and that the message to be encrypted belongs to a small list known messages. Then, if using a deterministic encryption scheme, in order to decode a given ciphertext, any adversary could compute the ciphertexts corresponding to all the messages, and compare them to the target ciphertext to determine the correct plaintext message. Therefore, randomness is an essential ingredient in any secure public-key encryption scheme.

6.2.3 Digital Signatures

The information contained in a national identification system need not only be confidential, but also authentic. One needs to be convinced that the information he/she gets has been certified by some trusted authority. The natural solution for this issue would be a signature of the authority on the data.

Along with the invention of public-key cryptography, Diffie and Hellman introduced in their seminal paper *New Directions in Cryptography* [80] the notion of digital signatures. Informally, a digital signature is the electronic analogue of a paper-based handwritten signature: only the designated signer can produce a valid signature, while everyone can verify it. In the electronic setting, the signer will compute its signature on a document using his/her secret key while everyone can verify the signature using the signer's public key.

Note that, unlike handwritten signatures, digital signatures depend on the document. This very property is essential in the digital context: since the duplication of electronic data (and thus of a digital signature) is a mere formality, the dependency of the document "guarantees" that only the valid signer could have produced the signature on a given document (and that this signature was not a duplicate of a signature on another document).

Digital signatures have been the object of extensive study, and most of the current research on cryptographic signature schemes today focus on some additional properties these schemes may have.

Blind signatures

The concept of blind signatures was introduced by David Chaum [50, 51] in 1983. As presented by Chaum, "the concept of blind signature can be illustrated by an example taken from the familiar world of paper documents. The paper analog of a blind signature can be implemented with carbon paper lined envelopes. Writing a signature on the outside of an envelope leaves a carbon copy of the signature on a slip of paper within the envelope."

Its main advantage is the following: the signer can sign a document "blindly" (i.e. without knowing its exact contents) while having some assurance about the nature of the document. More precisely, a blind signature protocol would allow the signer to sign only documents satisfying some properties – and thus prevent him/her from signing blindly just any document – while preventing him/her from seeing its exact contents, thus preserving

some level of privacy for the party requesting the signature.

The invention of blind signatures has sprung new research areas in cryptography, dealing with privacy and anonymity related issues. Among the many popular applications are electronic payments and electronic voting. It also enabled the development of privacy-enhancing techniques that are of more direct interest to our purpose here, which we will present in section 6.3.

Group signatures

The concept of group signatures was introduced by David Chaum and Eugène Van Heyst [56] in 1991. Group signatures allow any member to sign on behalf of a group while keeping the identity of the member secret. They are of particular interest in applications where what matters is not the actual identity of the signer, but rather its membership of a group, or its social/legal function. For instance, a person working at the identity authority would not certify some data by signing as John Smith, but as an accredited signer on behalf of the identity authority.

Some recent group signature schemes also allow "a designated group manager [to] revoke the anonymity and identify the originator of a signature" [45], which enables the audit of the signature operations. This could act as a deterrent for not abusing one's signing privileges, insofar as the actual identity of the signer can be revealed in the case of some later investigation for instance.

6.2.4 Threshold cryptography

Adi Shamir [145] first addressed the problem of sharing secrets in 1979 , and gave birth to threshold cryptography: a secret is divided into n different pieces (shares) such that k of them ($1 \leq k \leq n$) are required (and sufficient) to recover the original secret.

Nowadays, the protection of secrets is achieved by using encryption, which in turn "pushes" the secret a level higher: the key. While a secret sharing protocol would be useful for sharing a one-time key such as a one-time pad, it would not be practical in a public-key infrastructure where the same secret key is used many times (and sometimes for many purposes): the party performing the reconstitution of the key could keep it for later use. Yvo Desmedt and Yair Frankel [79] presented in 1990 "practical non-interactive public key systems" that allow the reuse of the shared secret key. They also suggest the use of

pseudonyms for the shareholders to keep secret their actual identities.

Threshold cryptography has also an interesting domain of application in threshold digital signatures: a secret key used for signatures is divided into n different pieces such that k of them ($1 \leq k \leq n$) are required (and sufficient) to perform the signature. Many implementations have been proposed for this notion, introduced in 1988 by Yvo Desmedt [78]. For instance, Victor Shoup [149] proposed in 1999 an efficient yet simple implementation based on RSA.

The utility of threshold for a national identification scheme is evident: high-security functionalities such as the identity creation or the authentication of new personal information (with a signature) are critical privileges that should be shared among multiple persons or authorities to minimize the risk of insider corruption.

6.2.5 Secure multi-party computation

The problem of secure multi-party computation arises in the following context: two or many parties want to determine some property about the pieces of information they hold while keeping them private. This was first addressed by Andrew Yao [161] in 1982 with his famous Millionaire problem: Alice and Bob both hold an integer number; each one wants to know which one is greater without having to disclose to the other party his/her number.

The main results in the domain were achieved by Andrew Yao [162] in 1986 for the two-party case and Oded Goldreich and Silvio Micali and Avi Wigderson [101] in 1987 for the multi-party case. For a more complete exposition of the subject, we refer the reader to Oded Goldreich's treatment of the subject [99], which material will be included in a soon-to-be-published book [97].

For our national identification system application, the utility of secure multi-party computation is manifest: for instance, in the profile revelation protocol, both the individual and examiner may be distrustful of each other. Using secure two-party computation, each party can be assured to complete the protocol if the other party is authentic and honest, while not disclosing any sensitive information if the other party is not.

Note that although the problems of threshold cryptography can be solved within the more general setting of secure multi-party computation, the solutions brought by threshold cryptography are better tailored at the above-mentioned needs and above all more efficient.

6.2.6 Zero-Knowledge

Shafi Goldwasser, Silvio Micali and Charles Rackoff [103] introduced in 1985 the notion of interactive proofs and zero-knowledge (and knowledge). Unlike "normal" proofs, an interactive proof involves two parties: a prover and verifier. Instead of proving to any verifier the veracity of a result, in an interactive proof, a prover only answers specific questions of the verifier.

They also consider the amount of knowledge communicated in a proof. Imagine you want to prove that you are over age in order to register for voting. Possible options include proving your date of birth, or proving your age. While the former clearly meets the proof requirements, it leaks more knowledge than the latter. Goldwasser, Micali and Rackoff define a proof to be zero-knowledge, if it does not provide more information than what is to be proven. In our age example, a zero-knowledge proof will leave the verifier with the conviction that the prover is over age while not having any clue to his/her exact age.

In 1988, Manuel Blum and Paul Feldman and Silvio Micali [38] first showed that a zero-knowledge proof need not be interactive, basing their proof on a non-standard assumption. Later that same year, Alfredo De Santis and Silvio Micali and Giuseppe Persiano [143] prove the same result using a weaker and well-known assumption (the hardness of quadratic residuosity). A non-interactive proof is one that does not need the participation of the verifier: the prover generates a proof by himself, which the verifier verifies at a later time. Non-interactive zero-knowledge proofs are essential for the issuance of "zero-knowledge credentials" for instance: for example, a person may request a series of credentials proving that he/she is over age, which he/she would use at some indefinite later time.

Also, Guilles Brassard and David Chaum and Claude Crépeau [42] introduced in 1988 the related notion of minimum disclosure proofs of knowledge, or zero-knowledge argument: unlike in zero knowledge, the same result – not disclosing more information than that contained in the result to prove – applies here for a prover and verifier with bounded computing power.

Nowadays, current research on zero-knowledge focus on efficiency, especially the round complexity (the number of messages exchanged between the prover and verifier), the composition of zero-knowledge proofs – sequentially, in parallel or concurrently. For the interested reader, we refer to Oded Goldreich's survey [100] of the field in 2002.

6.2.7 Steganography and information hiding

Steganography has as long a history as cryptology: one can trace its origins back to antiquity. Indeed, an appealing alternative to keeping a secret by encrypting it is to hide its very existence within an apparently innocuous document, which is the purpose of steganography. Fabien Petitcolas, Ross Anderson and Markus Kuhn [134] published a survey in 1999 of the information hiding field – which also includes fingerprinting and digital watermarking as well as steganography.

Early steganographic techniques include the use of invisible inks, the use of paper masks with holes to apply to the carrying document to reveal the secret, or hiding the message in an inner layer of the physical support.

Although this field long focused on the actual techniques to achieve these different goals – techniques often based on signals and communications theory – there has been a recent movement towards a formalization of the notion of security, and the introduction of notions similar to those used in cryptography. Gustavus Simmons [150] first addressed the issue by presenting the prisoner's problem in 1984: Alice and Bob, two prisoners, want to engage in a secret communication; however, any communication between them is monitored by Willy the warden. The goal is for them to disguise their actual communication into an innocent one without being detected. More recently, Christian Cachin [44] published an information-theoretic model for steganography in 1998.

The awareness of current information hiding techniques is crucial for instance if one decides to use biometrics data. Indeed, an information authority could embed some steganographic data (which could be some tracking information) within a picture for instance, a possibility that one may want to prevent.

6.3 The use of cryptography for identification

6.3.1 The early approach to identification

Adi Shamir [146] first introduced the idea of identity-based cryptosystems in 1984, where he also gave a concrete implementation for identity-based signature schemes. Later, in 1987, Amos Fiat and Adi Shamir [88] presented the first identification scheme based on the former [146] and zero-knowledge interactive proofs [103]. The idea is then presented more formally by Uriel Feige and Amos Fiat and Adi Shamir [86, 87] in 1987 .

The notion of "identity" used here is some kind of a name. More exactly, it is a public key assigned by an identity authority to the user representing his/her "name and network address [...]. Any combination of name, social security number, street address, office number or telephone number can be used (depending on the context) provided that it uniquely identifies the user in a way he cannot later deny, and that it is readily available to the other party" [146]. After proper "real" identification by the identity authority, the user is given the secret key corresponding to the public key: identification/authentication then consists in proving knowledge of the secret key associated to the claimed identity – represented by the public key.

Although the later version of the scheme proposed uses zero-knowledge proofs to only prove knowledge of the secret key (and no more), it is still vulnerable to a number of attacks: "instant replication" attack (where the adversary would "replicate", in another location and at the same time, the identification protocol performed by Alice to an examiner who is accomplice with the attacker), identity sharing/lending, multiple identity fraud. The main reason is that the identity is not tied to the individual: it only uses for the authentication two "sources of information" as defined in section 3.8 – what the user has, and what the user knows – and does not take into account the most fundamental one – who the user is.

In 1988, George Davida and Yvo Desmedt [75] presented a more general model for identification, for use for passports and visas. This work is interesting for multiple reasons:

- it first introduces the notion of updatable identity: in the scheme proposed, the passport contains "an area (special memory) where data can be appended and read by everybody" [75]. This area is intended in the model for stamps given by the visiting countries' customs and immigration services.

- it then separates the actual identification (identity determination) function performed by the passport from the authorization to visit a foreign country based on the profile enclosed in the visa.

These schemes however lack a fundamental aspect of our model for identification: they do not consider privacy issues, but only address a total identification.

6.3.2 The current notions of identity in cryptography

Most cryptographic schemes addressing identification consider a digital identity to be a public key. This public key would be a unique identifier for the party (person, computer machine, computer process, etc), and thus "represents" its digital identity. A widely used standard endorsing this approach is the X.509 public-key infrastructure [106, 92]. One major difficulty in this approach is how one assigns the public keys to the parties; in particular, one usually wants a party's public key to be somehow tied to its identity: rather than being just a number, it would be more convenient if it included the name of party for instance.

SPKI/SDSI (Simple Public Key Infrastructure/Simple Distributed Security Infrastructure) [25] however uses a new approach to identity by introducing the notion of principal: "a principal is an individual, process or entity whose messages are distinctively recognizable because they are digitally signed by the public key that represents them. It is convenient to say that the principal *is* its public key" [58]. In this approach, rather than worrying about how to assign public keys, a party is identified through its possession of a secret key. A provable use of a party's secret key (signature of a document, zero-knowledge proof of the knowledge of the key, etc) identifies the party: the digital identity is defined as the ability for a party to perform some cryptographic action (through the knowledge of a secret key).

Yet another approach to identification is illustrated by the problem of Identity-Based Encryption (IBE) introduced by Adi Shamir [146] in 1984: an IBE scheme is a public-key cryptosystem where any string is a valid public key. A public key in this model is some "natural" representation of a person's digital identity (typically an email address). There is therefore no need for a trusted authority to certify public keys; rather, the secret key corresponding to a given public key can only be computed by a central trusted authority, upon proper authentication of the person "owning" the public key. Dan Boneh and Matt Franklin [39] proposed in 2001 an efficient scheme to solve this problem, and implemented it in an IBE Secure email system [13].

6.3.3 Achieving Electronic Privacy: credential systems

In 1985, in his seminal paper *Security Without Identification: Transaction Systems to Make Big Brother Obsolete* [52], David Chaum first addresses the fundamental matter of privacy in computer systems. In this paper, he laid the foundations for many a privacy

technique by introducing the notions of credential systems using pseudonyms. He later analyzed in an article published by the Scientific American [53] the impact of his cryptographic inventions, blind signatures especially, on the pursuit of achieving electronic privacy. We present here the fundamental properties of the new notion of credential system, for which David Chaum and Jan-Hendrik Evertse [54] proposed the first practical solution in 1987.

Credential systems

Chaum introduced the matter of credential systems as follows [52]:

> "There are legitimate needs for individuals to show credentials in relationships with many organizations. Problems arise when unnecessary data are revealed in the process. As used here, credentials are statements based on an individuals relationship with organizations that are, in general, provided to other organizations."

In common credentials, such as passports or driver's licences, the disclosure of the credential implies the revelation of irrelevant or unnecessary data, that can later be used to link this disclosure with other transactions, thus seriously endangering the privacy of the individual.

Chaum introduced a new approach to address this problem, which he calls a credential system. In a credential system, each person is known to different organizations through different pseudonyms. In this system, a person can request a credential from organization A relative to personal information under his/her A-pseudonym, and transform it into a credential pertaining to his/her B-pseudonym in order to show it to organization B. This procedure is called a credential transfer (insofar as the credential is transferred from the A-pseudonym to the B-pseudonym).

Let us illustrate this with an example. Say for example, that you get sick and go to the doctor. He/she does not need to know who you are and where you live, but only your personal profile as a patient: age, known allergies, known problems (e.g. asthma, diabetes, or other hypertension), history of past visits and surgery, etc. Therefore, instead of using your actual name, you could use a pseudonym for your medical record. Then, when you turn to your insurance company to get your health care reimbursed, it need not know about your actual health issues, but only that the care the doctor gave you is entitled to

reimbursement. So you transform the credential given by your doctor pertaining to your medical pseudonym, to a credential for reimbursement pertaining to the pseudonym (which you keep distinct) you use for dealing with the insurance company.

This credential transfer and the credential disclosure feature a set of useful properties to maintain the privacy of the individual, which we describe now.

Selective disclosure of information

During the credential disclosure protocol, the person need not reveal all the information contained in the credential, but only that this credential contains information that meets some criteria. We have already seen many examples of how this could be used in a national identification scheme. For instance, in many situations (purchase of alcohol or tobacco, voting registration, etc), you need to prove that your age meets some legal requirements. However, you need not reveal your date of birth or even your actual age: most of the time, all you need is to prove that your age falls in a certain range (greater than 21, greater than 18, etc).

Optimal unlinkability of pseudonyms

The credential transfer from a person's A-pseudonym to his/her B-pseudonym is robust yet, does not allow for anyone but the person from linking both pseudonyms, even if both organizations A and B collude. This unlinkability of pseudonym is not total however, but optimal: while organization B learns nothing more from the person's relationship with organization A than the information disclosed by the individual, nothing prevents organization B from recording this information with the person's B-pseudonym and start collecting information about the person's A-pseudonym. To avoid the constitution of such dossiers, Chaum recommends the periodic change of a person's pseudonym with an organization, yearly for instance.

Limited-show credentials and conditional anonymity revocation

A very useful feature of Chaum's system is the ability to create credentials that are to be shown only a limited number of times. Such a credential issued by organization A preserves the confidentiality of the person's A-pseudonym if used less than the limit, but reveals it otherwise, thus enabling the tracking of the person in case of abuse. This type of credentials

ideally suits the need for temporary or limited authorization: a one-show credential could allow for instance a person to enter a building only once.

One-show credentials were initially studied for their application to electronic cash. David Chaum, Amos Fiat and Moni Naor [55] proposed in 1988 an untraceable electronic cash system, in which the anonymity of a spender is revealed, when the coin is deposited it to the bank, if he/she spends it twice. Stefan Brands [40] proposed in 1994 an efficient solution to the revocation of anonymity in the case of double spending, which allows in addition the prevention of double spending if the spender uses a tamper-resistant device from the bank.

6.3.4 Pseudonym systems

The practical solution proposed by Chaum and Evertse [54] to Chaum's credential system [52], as well as some later schemes [73, 57], effectively preserves the individual's privacy against colluding organizations, but does not protect the organizations against colluding individuals to share their credentials. The pseudonym system[3] introduced by Anna Lysyanskaya, Ronald Rivest, Amit Sahai and Stefan Wolf [117] in 1999 effectively addresses the problem, and adds some additional properties to Chaum's credential system. Jan Camenish and Anna Lysyanskaya [46, 47] improved this pseudonym system in 2001/2002, by constructing in 2001/2002 a multiple-show credential system and mechanisms for credential revocation. These various advances are summarized in Anna Lysyanskaya's PhD thesis [116] in 2002.

Preventing credential sharing

The problem of credential sharing can be illustrated by the following medical example. Chaum's credential system does not prevent Alice and Bob to collude so that Bob would get a reimbursement through his insurance company for Alice's medical visit: Alice and Bob can share some of their pseudonyms (and associated knowledge such as secret keys) to perform the transfer of a credential issued to one of Alice's pseudonym to one of Bob's pseudonym.

Anna Lysyanskaya, Ronald Rivest, Amit Sahai and Stefan Wolf [117] remedy to this problem of identity sharing by presenting a pseudonym system in which "each user has a

[3] A pseudonym system is fundamentally no different from Chaum's credential system. We use here the term pseudonym system for this scheme as it is the term widely used in the literature, and in particular by the authors themselves.

master public key whose corresponding secret key the user is highly motivated to keep secret. This master key might be registered as his legal digital signature key, so that disclosure of his master secret key would allow others to forge signatures on important legal or financial documents in his name. [The] proposed scheme then has the property that a user can not share a credential with a friend without sharing his master secret key with the friend, that is, without *identity sharing.*"

Multiple-show credentials

Chaum's credential system distinguishes between one-show and multiple-show (for unlimited use) credentials. Unlike one-show credentials, multiple-show credentials require another unlinkability property: the uses of the same credentials need to be unlinkable to one another. Let us take the example of an age credential, proving that you meet the legal requirements for purchasing alcohol or tobacco. The possibility for colluding organizations to link your disclosures of this credential would seriously undermine your privacy: since this credential would be mainly used for purchasing alcohol or tobacco, this would enable them to determine your drinking or smoking habits.

While earlier credential systems allow for the use of multiple-show credentials, Jan Camenish and Anna Lysyanskaya [46] constructed in 2001 unlinkable multiple-show credentials: a disclosure of a multiple-show credential cannot be linked to any other disclosure of the same credential.

Anonymity revocation

The pseudonym system proposed by Jan Camenish and Anna Lysyanskaya [46] also allows for anonymity revocation in the case a user tries to perform illegal transactions. In a system where the individual is in control of the personal information he/she discloses, this feature is of great help to enforce the security of a pseudonym system against various frauds and abuses by providing a strong deterrent for the malicious user.

Anonymous credential revocation

Jan Camenish and Anna Lysyanskaya [47] introduced the notion of dynamic accumulators in 2002, which can be used (among other things) for enabling the revocation of anonymous credentials.

6.3.5 Attribute certificates

In his PhD thesis [41] in 1999, Stefan Brands builds on the notion of credential systems introduced by Chaum, but adopts another approach to protect the privacy of private individuals. Instead of relying on the use of different unlinkable pseudonyms, he designed attribute certificates in which the certificate authority encodes attributes along with the individual's public key.

A novel approach to privacy-enhancing cryptography

In Brands' model, the credential takes the form of an attribute certificate issued by a certificate authority, which can be shown to any organization trusting the certificate authority. In particular, the individual need not register a pseudonym with an organization to which he/she wishes to show a credential.

Rather than achieving privacy through the use of pseudonyms, the system relies on a restrictive blinding certificate issuing protocol. As mentioned above, the certificates issued in this model not only includes the authentication of the user's public key, but also attributes. An attribute certificate issued by a DMV could include for instance an individual's name, address, driving authorization and some other identifying information (such as height, eye color, etc) along with his/her public key. The restrictive blinding property of the certificate issuing protocol could be presented as follows: the certificate authority (CA) "blindly" issues a certificate to the individual while getting the assurance it encodes the attributes in the certificate. While the certificate contains a valid signature of the CA on the public key, the CA itself does not know what public key is issued to the individual, and thus cannot link that public key to the attributes it encoded. This successfully achieves the unlinkability of the attributes to the public key after certificate issuance.

Brands' scheme also enables the selective disclosure of information. In that matter, it actually allows for the proof of any algebraic relation on the attributes. For instance, one can prove that "(I am older than 21 and live in Massachusetts) OR (I am older than 18 and was born in California)".

Using multiple certificates

Brands' scheme allows for a person to combine certificates issued by different CAs, provided that they operate in the same discrete logarithm group, or even for multiple persons to combine their certificates to prove a property about the combined set of attributes present in the certificates.

Such a feature could be used to benefit from group or family privileges. For instance, a group of friends can then prove that they are all students to benefit from a student group discount, and a family can take advantage of some promotional offer reserved to families having at least 2 children under 16.

Pseudonyms and public key

Brands' scheme falls short of Lysyanskaya's scheme with respect to the following matter: the disclosures of a multiple-show credential may be linked to one another. However, the absence of a pseudonym in the scheme allows for the issuance of multiple certificates encoding the same attributes but using different key pairs, thus preventing their linkage, should the issuances of the certificates be independent.

Although public keys have been often used as some form of pseudonym (or as a name or identifier) in identification applications, in Brands' scheme, the individual public key need not be tied to the "identity" of the individual, or to any of his/her attributes. In fact, it could be chosen at random by the individual.

Anonymous attribute certificates

We present in this section a model that is underlying in Brands' work.

Insofar as the CA only certifies the attributes it encodes in the attribute certificates it issues, it need not fully identify the individual for issuing certificates. While in the credential/pseudonym systems the CA issues credentials relative to a pseudonym registered with the CA, in Brands' system, the CA needs only to verify the authenticity of the attributes it encodes.

If the purpose of one CA is to certify only physical attributes for instance, all you ever need to do to get certificates from this CA is to request the certificates in person: it will never need you to disclose any personal information, but only needs to measure you height

or assess your eye color to certify these attributes.

The ability to use Brands' scheme for anonymous attribute certificates is utmost interest when considering a national identification scheme. If the identity authority and the different information authorities run independent systems, this allows for an effective partition of the information between the information authorities, which prevents the linkability of the different pieces.

6.4 Cryptology resources

For more information on cryptography, a good starting point is Ronald Rivest's collection of links on cryptography and security [140]. Although a little out of date now, it still lists, organized by categories, a significant number of relevant references. Also, a very comprehensive compilation of links is Helger Lipmaa's *Cryptology Pointers* [115], organized by themes.

For the reader more interested in scientific and/or technical issues, we recommend the following books:

- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein's *Introduction to Algorithms* (2001) [69] is the reference textbook for whomever wishes to get a start in Computer Science, and more specifically algorithms, which cryptography is based on.

- Oded Goldreich's *Foundations of Cryptography* is a three-volume work under progress. The first volume *Foundations of Cryptography – Basic Tools* (2001) [98] offers a theoretical treatment of the mathematical tools on which modern cryptography is built on. The second volume, in preparation, *Foundations of Cryptography – Basic Applications* has preliminary drafts available online [97].

- Douglas R. Stinson's *Cryptography: Theory and Practice* (2002) [152] is an excellent general textbook, which covers the essential core areas of cryptography, as well as a selection of more advanced topics.

- Alfred J. Menezes and Paul C. van Oorschot and Scott A. Vanstone's *Handbook of Applied Cryptography* (1997) [120] is a very comprehensive book focusing on the practical aspects of cryptography, yet with a mathematical presentation.

- Bruce Schneier's *Applied Cryptography* (1996) [144] represents an excellent introduction to applied security and cryptography. This book is much less formal than the previous ones, but rather focuses on the practical aspects of cryptography and security. It also contains a really extensive bibliography.

- Nigel Smart's new *Cryptography, An Introduction* (2002) [151] "provides the rigorous detail required for advanced cryptographic studies, yet approaches the subject matter in an accessible style in order to gently guide new students through difficult mathematical topics"[4].

[4]Description taken from the book's official webpage http://www.mcgraw-hill.co.uk/html/0077099877.html.

Chapter 7

Summary

In this section, we summarize our discussion by recalling here to what extent a national identification scheme using current state-of-the-art technology may address the most pre-eminent concerns raised in the public opinion regarding that matter: enhancing national security, preserving personal privacy, preventing pervasive surveillance, preventing feature creep, and reconciling biometrics and privacy. Finally, we conclude by analyzing the adoption (and possibly deployment) process that should be followed by a government considering seriously the matter of a national identification scheme.

7.1 Enhancing national security

Although the recent focus on national identification systems has been motivated by the need to fight terrorism following the tragic events of September 11, 2001, such systems may actually do little towards this goal. However, a technology-enabled national identification scheme can possibly enhance the security of identification. By using current available technology presented in this thesis (including biometrics, cryptology), one can significantly reduce many risks, including forgery, identify sharing and lending, and identity theft.

7.2 Preserving personal privacy

Even though the general skepticism of the public opinion about national ID cards often focuses on the ID card thrown out a as monitoring tool, the main risks regarding the individual's personal privacy do not reside in the card itself, but actually in the databases

of the national identification system.

While chapter 6 presented many a cryptographic techniques to enable various aspects of the preservation of the privacy of personal information, technology alone cannot win the fight for privacy. Indeed, since nothing but appropriate policy can prevent organizations from maintaining databases with information gathered from identification processes, and sharing it with others, and possibly using the combined aggregated personal information for purposes other than those for which the identification processes took place in the first place.

Technology may limit the amount of information revealed during the legitimate use of the system, but only apt policy can regulate the possible misuses and abuses of the databases containing personal information, especially those internal to the national identification system.

7.3 Preventing pervasive surveillance

The matter of pervasive surveillance is closely related to the aforementioned preservation of personal privacy. Yet, here the use of technology can prove to be really helpful. Indeed, we have seen throughout this thesis many cryptographic techniques preventing the tracking of the individuals, as well as the unlinkability of an individual's pieces of information if he/she wishes so. Also, surprisingly, technology enables the anonymous disclosure of credentials.

7.4 Preventing feature creep

When considering a national identification scheme, with the goal of a somewhat universal identification, one should remember the following: if adopted, such a scheme can potentially be in force for decades. A well-designed system is then less vulnerable to a technical weakness than to a potential misuse or abuse. In particular, it is very appealing for organizations to use the personal information gathered from identification processes for other purposes than those intended.

7.5 Reconciling biometrics and privacy

As a tool for providing reliable automated physical identification, biometrics allows for higher standards of security in identification systems. However, while biometrics by nature aims at providing a better differentiation of human beings, privacy-enhancing cryptology aims at blurring this very differentiation. Therefore ,it remains unclear whether one can reconcile the benefits of both biometrics and privacy-enhancing cryptology.

7.6 The adoption process

Should a government want to initiate the process of evaluating whether or not to adopt a national identification system (and then that of deploying it if adopted), it should proceed to at least the following:

- First and foremost, it should clearly identify the intended purpose(s) of the system and define the desired goals, projected budget and timeline of the project.

- Based on the above, it should take position on the essential policy issues. A good starting point for this would be for instance to give specific answers to the policy questions raised by the Computer Science and Telecommunications Board (CSTB) of the US National Research Council (NRC) in its report on nationwide identification systems [113].

- Then only, a project committee would be able to work on recommendations, and eventually a design proposal. Since the security and quality of the system depend on both appropriate policy directives and sound technological choices, this committee needs to include both policy and technical experts.

- The design proposal should definitely be debated before the Parliament (or equivalent legal structure), which may call for changes. This can go back and forth many times between the project committee working phases and the Parliament discussions.

- The adoption of a national identification system requires eventually the approval of the Parliament (or equivalent legal structure).

- All along the long political, technical and legal process, provisions should be made to allow for public consultation and debate, as well as public comments and suggestions on actual recommendations and proposals.

- As regards the choice of the technologies to be used, the state of currently available technology needs to be evaluated. To that end, the government can either invite tenders or order its own study.

- If adopted, technical specifications, as well as possibly the definition of certain standards, will be needed prior to the actual implementation and deployment of the system. This may involve a standardization agency, such as the National Institute of Standards (NIST) in the US.

Bibliography

[1] *Against TCPA.* http://www.againsttcpa.com/.

[2] *Auto-ID Center.* http://www.autoidcenter.org/.

[3] *Avanti Biometric Reference.* http://homepage.ntlworld.com/avanti/home.htm.

[4] *Biomet - The Biometric Resource Center.* http://www.biomet.org/.

[5] *Biometric Consortium.* http://www.biometrics.org/.

[6] *Biometric Digest.* http://www.biodigest.com/.

[7] *Biometrics Institute.* http://www.biometricsinstitute.org/.

[8] *The CAPTCHA Project.* http://www.captcha.net/.

[9] *Crypto History: Enigma.* http://www.tcs.hut.fi/~helger/crypto/link/history/enigma.html.

[10] *DARPA's Information Awareness Office (IAO) and Total Information Awareness (TIA) Program – Frequently Asked Questions.* http://www.darpa.mil/iao/TIA_FAQs.pdf.

[11] *Defense Advanced Research Projects Agencys Information Awareness Office and Total Information Awareness Project.* http://www.darpa.mil/iao/iaotia.pdf.

[12] *The History of the International Movement To Standardize Passports.* http://www.icao.int/icao/en/atb/fal/mrtd/guide.htm#history.

[13] *Identity-Based Encryption.* http://crypto.stanford.edu/ibe/.

[14] *International Biometric Industry Association (IBIA).* http://www.ibia.org/.

[15] *International Civil Aviation Organization.* http://www.icao.int/.

[16] *Liberty Alliance Project.* http://www.projectliberty.org.

[17] *Machine Readable Travel Documents.* http://www.icao.int/icao/en/atb/fal/mrtd/.

[18] *Microsoft .NET Passport.* http://www.microsoft.com/netservices/passport/.

[19] *Microsoft .NET Passport Privacy Statement.* http://www.passport.net/Consumer/PrivacyPolicy.asp.

[20] *Microsoft .NET Passport Review Guide.* http://www.microsoft.com/net/downloads/passport_review_guide.doc.

[21] *Microsoft Palladium: A Business Overview.* http://www.microsoft.com/presspass/features/2002/jul02/0724palladiumwp.asp.

[22] *Natural-Person Home Page.* http://www.natural-person.ca/.

[23] *No TCPA!* http://www.notcpa.org/.

[24] *Shrouds of Time - The History of RFID.* http://www.aimglobal.org/technologies/rfid/resources/shrouds_of_time.pdf.

[25] *A Simple Distributed Security Infrastructure (SDSI).* http://theory.lcs.mit.edu/ cis/sdsi.html.

[26] *Social Security Online - The Official Website of the Social Security Administration.* http://www.ssa.gov/.

[27] *TCPA - Trusted Computing Platform Alliance.* http://www.trustedpc.org.

[28] *U.S. National Biometric Test Center.* http://www.engr.sjsu.edu/biometrics/.

[29] *World War II Codes and Ciphers.* http://www.codesandciphers.org.uk/.

[30] Social Security Administration. *Your Number And Card.* http://www.ssa.gov/pubs/10002.html.

[31] Philip E. Agre. *Your Face Is Not a Bar Code: Arguments Against Automatic Face Recognition in Public Places*, 2003. http://dlis.gseis.ucla.edu/people/pagre/barcode.html.

[32] AIM. *Bar code main page.* http://www.aimglobal.org/technologies/barcode/.

[33] AIM. *Radio Frequency Identification (RFID) home page.* http://www.aimglobal.org/technologies/rfid/.

[34] Ross Anderson. *TCPA / Palladium Frequently Asked Questions.* http://www.cl.cam.ac.uk/rja14/tcpa-faq.html.

[35] Biometric Research at MSU (Michigan State University). *Publications.* http://biometrics.cse.msu.edu/publications.html.

[36] Avanti. *The Biometric White Paper.* http://homepage.ntlworld.com/avanti/home.htm.

[37] Avanti. *The Distinction Between Authentication and Identification.* http://homepage.ntlworld.com/avanti/home.htm.

[38] Manuel Blum, Paul Feldman, and Silvio Micali. Proving security against chosen cyphertext attacks. In S. Goldwasser, editor, *Proc. CRYPTO 88*, pages 256–268. Springer-Verlag, 1988. Lecture Notes in Computer Science No. 403.

[39] Dan Boneh and Matt Franklin. Identity-based encryption from the Weil pairing. In Joe Killian, editor, *Proc. CRYPTO 01*, pages 213–229, 2001.

[40] Stefan Brands. Untraceable off-line cash in wallets with observers (extended abstract). In Douglas R. Stinson, editor, *Proc. CRYPTO 93*, pages 302–318. Springer, 1994. Lecture Notes in Computer Science No. 773.

[41] Stefan A. Brands. *Rethinking public key infrastructures and digital certificates – Building in privacy.* PhD thesis, Eindhoven Institute of Technology, Eindhoven, The Netherlands, 1999.

[42] Guilles Brassard, David Chaum, and Claude Crépeau. Minimum disclosure proofs of knowledge. *JCSS*, 37(2):156–189, 1988.

[43] Dr. Manfred Bromba. *Biometrics FAQ.* http://home.t-online.de/home/manfred.bromba/biofaqe.htm.

[44] Christian Cachin. An information-theoretic model for steganography. *Lecture Notes in Computer Science*, 1525:306–318, 1998.

[45] Jan Camenisch. Efficient and generalized group signatures. In Walter Fumy, editor, *Advances in Cryptology – Eurocrypt '97*, pages 465–479, Berlin, 1997. Springer. Lecture Notes in Computer Science 1233.

[46] Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In Birgit Pfitzmann, editor, *Advances in Cryptology - EUROCRYPT 2001*, pages 93–118. Springer-Verlag, 2001.

[47] Jan Camenisch and Anna Lysyanskaya. Dynamic accumulators and application to efficient revocation of anonymous credentials. In Moti Yung, editor, *Advances in Cryptology - CRYPTO 2002*, pages 61–76. Springer-Verlag, 2002.

[48] Electronic Privacy Information Center and Privacy International, editors. *Privacy and Human Rights - An International Survey of Privacy Laws and Developments.* Electronic Privacy Information Center and Privacy International, first edition, 2001.

[49] U.S. National Biometric Research Center. *Collected Works*, 1997–2000. http://www.engr.sjsu.edu/biometrics/nbtccw.pdf.

[50] David Chaum. Blind signatures for untraceable payments. In R. L. Rivest, A. Sherman, and D. Chaum, editors, *Proc. CRYPTO 82*, pages 199–203, New York, 1983. Plenum Press.

[51] David Chaum. Blind signature system. In D. Chaum, editor, *Proc. CRYPTO 83*, pages 153–153, New York, 1984. Plenum Press.

[52] David Chaum. Security without identification: Transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10):1030–1044, October 1985.

[53] David Chaum. Achieving electronic privacy. *Scientific American*, 267(2):96–101, August 1992.

[54] David Chaum and Jan-Hendrik Evertse. A secure and privacy-protecting protocol for transmitting personal information between organizations. In A. M. Odlyzko, editor, *Proc. CRYPTO 86*, pages 118–167. Springer-Verlag, 1987. Lecture Notes in Computer Science No. 263.

[55] David Chaum, Amos Fiat, and Moni Naor. Untraceable electronic cash (extended abstract). In Shafi Goldwasser, editor, *Proc. CRYPTO 88*, pages 319–327. Springer-Verlag, 1988. Lecture Notes in Computer Science No. 403.

[56] David Chaum and Eugène Van Heyst. Group signatures. In D. W. Davies, editor, *Advances in Cryptology — Eurocrypt '91*, pages 257–265, Berlin, 1991. Springer-Verlag. Lecture Notes in Computer Science No. 547.

[57] Lidong Chen. Access with pseudonyms. In Ed Dawson and Jovan Golic, editors, *Cryptography: Policy and Algorithms*, pages 232–243. Springer-Verlag, 1995. Lecture Notes in Computer Science No. 1029.

[58] Dwaine Clarke, Jean-Emile Elien, Carl Ellison, Matt Fredette, Alexander Morcos, and Ronald L. Rivest. Certificate chain discovery in SPKI/SDSI. *Journal of Computer Science*, 9(4), 2001.

[59] Roger Clarke. *Human Identification in Information Systems: Management Challenges and Public Policy Issues*, 1994. http://www.anu.edu.au/people/Roger.Clarke/DV/HumanID.

[60] Roger Clarke. *Chip-Based ID: Promise and Peril*, 1997. http://www.anu.edu.au/people/Roger.Clarke/DV/IDCards97.html.

[61] Roger Clarke. *Introduction to Dataveillance and Information Privacy, and Definitions of Terms*, 1999. http://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html.

[62] Roger Clarke. *Biometrics and Privacy*, 2001. http://www.anu.edu.au/people/Roger.Clarke/ DV/Biometrics.html.

[63] Privacy Rights Clearinghouse. *Your Social Security Number: How Secure Is It?* http://www.privacyrights.org/fs/fs10-ssn.htm.

[64] CNN. *Japan ID system raises Big Brother fears*, August 2002. http://www.cnn.com/2002/WORLD/asiapcf/east/08/04/japan.idcard/.

[65] CNN.com. *Ellison offers free software for national ID card.* http://www.cnn.com/2001/TECH/industry/09/25/ellison.software.idg/.

[66] Air Force Material Command. *Biometrics Identification.* http://www.afmc.wpafb.af.mil/organizations/HQ-AFMC/LG/LSO/LOA/bio.htm.

[67] Federal Trade Commission. *Ads for International Drivers' Licenses or Permits Could Be a Dead End.* http://www.ftc.gov/bcp/conline/pubs/alerts/driveralrt.htm.

[68] Biometric Consortium. *Publications and Peridicals.* http://www.biometrics.org/html/publications.html.

[69] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms.* MIT Press/McGraw-Hill, second edition, 2001.

[70] Population council. *Population Statistics, the Holocaust, and the Nuremberg Trials.* http://www.popcouncil.org/mediacenter/newsreleases/pdr24_3_holocaust.html.

[71] Cryptography and Information Security Research Group at ETH Zurich. *Information-Theoretic Cryptography.* http://www.crypto.ethz.ch/research/itc/.

[72] Kenneth W. Dam and Herbert S. Lin, editors. *Cryptography's Role in Securing the Information Society.* Computer Science and Telecommunications Board (CSTB), The National Academies Press, 1996. This report is avalaible online at http://books.nap.edu/html/crisis/.

[73] Ivan B. Damgård. Payment systems and credential mechanisms with provable security against abuse by individuals. In S. Goldwasser, editor, *Proc. CRYPTO 88*, pages 328–335. Springer-Verlag, 1988. Lecture Notes in Computer Science No. 403.

[74] Defense Advanced Research Projects Agency (DARPA). *Total Information Awareness (TIA) Systems.* http://www.darpa.mil/iao/TIASystems.htm.

[75] George I. Davida and Yvo Desmedt. Passports and visas versus ids (extended abstract). In Christoph G. Gnther, editor, *Proc. EUROCRYPT 88*, pages 183–188. Springer-Verlag, 1988. Lecture Notes in Computer Science No. 330.

[76] Commission Nationale de l'Informatique et des Libertés (French National Commission for Information Technology and Liberties). *Loi No. 78-17 relative l'informatique, aux fichiers et aux libertés (Law No. 78-17 relative to information technology, files and liberties)*, 1978. http://www.cnil.fr/textes/text02.htm.

[77] Dorothy E. Denning and Peter F. MacDoran. Location-based authentication: Grounding cyberspace for better security. In Dorothy E. Denning and Peter J. Denning, editors, *Internet Besieged: Countering Cyberspace Scofflaws*, pages 167–174, New York, 1988. ACM Press / Addison-Wesley.

[78] Yvo Desmedt. Society and group oriented cryptography: A new concept. In Carl Pomerance, editor, *Proc. CRYPTO 87*, pages 120–127. Springer-Verlag, 1988. Lecture Notes in Computer Science No. 293.

[79] Yvo Desmedt and Yair Frankel. Threshold cryptosystems. In G. Brassard, editor, *Proc. CRYPTO 89*, pages 307–315. Springer-Verlag, 1990. Lecture Notes in Computer Science No. 435.

[80] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Trans. Inform. Theory*, IT-22:644–654, November 1976.

[81] EPIC. *Biometrics identifiers.* http://www.epic.org/privacy/biometrics/.

[82] EPIC. *ID Cards Archive.* http://www.epic.org/privacy/id_cards/.

[83] EPIC. *Social Security Numbers and Privacy.* http://www.epic.org/privacy/ssn/.

[84] EPIC. *Total Information Awareness (TIA).* http://www.epic.org/privacy/profiling/tia/.

[85] EPIC. *Your Papers, Please: From the State Drivers License to a National Identification System*, February 2002. http://www.epic.org/privacy/id_cards/yourpapersplease.pdf.

[86] Uriel Feige, Amos Fiat, and Adi Shamir. Zero knowledge proofs of identity. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 210–217, May 1987.

[87] Uriel Feige, Amos Fiat, and Adi Shamir. Zero knowledge proofs of identity. *Journal of Cryptology*, 1(2):77–94, 1988.

[88] Amos Fiat and Adi Shamir. How to prove yourself: practical solutions to identification and signature problems. In A. M. Odlyzko, editor, *Proc. CRYPTO 86*, pages 186–194. Springer, 1987. Lecture Notes in Computer Science No. 263.

[89] Organization for Economic Cooperation and Development (OECD). *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, 1980. http://www1.oecd.org/publications/e-book/9302011E.PDF.

[90] Computer Professionals for Social Responsibility. *Frequently Asked Questions on SSNs and Privacy.* http://www.cpsr.org/cpsr/privacy/ssn/ssn.faq.html.

[91] Computer Professionals for Social Responsibility. *National Identification Schemes (NIDS) and the Fight against Terrorism: Frequently Asked Questions.* http://www.cpsr.org/program/natlID/natlIDfaq.html.

[92] Warwick Ford and Michael S. Baum. *Secure Electronic Commerce: Building the Infrastructure for Digital Signatures and Encryption.* Prentice Hall PTR, second edition, 2000.

[93] Federation Trade Commission (FTC). *Identity Theft: When Bad Things Happen To Your Good Name.* http://www.ftc.gov/bcp/conline/pubs/credit/idtheft.pdf.

[94] Michael Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, 1979.

[95] Simson Garfinkel. *Database Nation - The Death of Privacy in the* 21^{th} *Century.* O'Reilly, 2000.

[96] Privacy Commissioner of Canada George Radwanski. *Annual Report to Parliament 2001-2002*, January 2003. http://www.privcom.gc.ca/information/ar/02_04_10_e.asp.

[97] Oded Goldreich. *Foundations of Cryptography - Basic Applications.* Preliminary draft of this book are available online at http://www.wisdom.weizmann.ac.il/ oded/foc-vol2.html.

[98] Oded Goldreich. *Foundations of Cryptography - Basic Tools.* Cambridge University Press, 2001.

[99] Oded Goldreich. *Secure multi-party computation (Final (incomplete) draft)*, October 2002. Available from http://www.wisdom.weizmann.ac.il/~oded/PS/prot.ps.

[100] Oded Goldreich. Zero-knowledge twenty years after its invention. Technical Report 2002/186, U.S.C. Computer Science Department, December 2002.

[101] Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 218–229, New York City, May 1987. ACM.

[102] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *JCSS*, 28(2):270–299, April 1984.

[103] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems. In *Proc. 17th ACM Symp. on Theory of Computing*, pages 291–304, Providence, 1985. ACM.

[104] International Biometric Industry Association (IBIA). *Frequently Asked Questions About Biometric Technology.* http://www.ibia.org/faqs.htm.

[105] International Civil Aviation Organization (ICAO). *Executive Summary of the Technical Report on ICAO Work on Selection and Testing of a Biometric Technology for Identity Confirmation with Machine Readable Travel Documents (MRTDS).* http://www.icao.int/icao/en/atb/fal/mrtd/biometric_tech.htm.

[106] Internet Engineering Task Force (IETF). *Public-Key Infrastructure (X.509) (pkix).* http://www.ietf.org/html.charters/pkix-charter.html.

[107] The Internet Engineering Task Force (IETF). *Transport Layer Security (tls) Charter.* http://www.ietf.org/html.charters/tls-charter.html.

[108] Privacy International. *National ID Cards.* http://www.privacyinternational.org/issues/idcard/.

[109] Privacy International. *UK National ID Cards.* http://www.privacyinternational.org/issues/idcard/uk/.

[110] Privacy International. *National ID Cards - Frequently Asked Questions*, August 1996. http://www.privacyinternational.org/issues/idcard/idcard_faq.html.

[111] ITworld.com. *Japan national ID system raises privacy concerns*, August 2002. http://www.itworld.com/Man/2688/020806japanid/.

[112] David Kahn. *The Codebreakers: The Story of Secret Writing.* Macmillian, New York, revised edition, 1996.

[113] Stephen T. Kent and Lynette I. Millett, editors. *IDs – Not That Easy.* Computer Science and Telecommunications Board (CSTB), The National Academies Press, 2002.

[114] Ray Kurzweil. *The Age of Spritual Machines.* Penguin Books, 1999.

[115] Helger Lipmaa. *Cryptology Pointers.* http://www.tcs.hut.fi/~helger/crypto/.

[116] Anna Lysyanskaya. *Signature Schemes and Applications to Cryptographic Protocol Design.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, September 2002.

[117] Anna Lysyanskaya, Ronald L. Rivest, Amit Sahai, and Stefan Wolf. Pseudonym systems (extended abstract). In Howard Heys and Carlisle Adams, editors, *Selected Areas in Cryptography*, pages 184–199. Springer-Verlag, 1999. Lecture Notes in Computer Science No. 1758.

[118] Tsutomu Matsumoto, Hiroyuki Matsumoto, Koji Yamada, and Satoshi Hoshino. *Impact of Artificial "Gummy" Fingers on Fingerprint Systems.* http://cryptome.org/gummy.htm.

[119] Ueli Maurer. Information-theoretic cryptography (extended abstract). *Lecture Notes in Computer Science*, 1666:47–64, 1999.

[120] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography.* CRC Press, 1997.

[121] Kevin D. Mitnick and William L. Simon. *The Art of Deception: Controlling the Human Element of Security.* John Wiley & Sons, 2002.

[122] Netscape. *Introduction to SSL.* http://developer.netscape.com/docs/manuals/security/sslin/contents.htm.

[123] Wired News. *Call It Super Bowl Face Scan.* http://www.wired.com/news/politics/0,1283,41571,00.html.

[124] Wired News. *The Oracle of National ID Cards.* http://www.wired.com/news/conflict/0,2100,47788,00.html.

[125] Council of Europe. *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108)*, 1981. http://www.coe.int/T/E/Legal_affairs/Legal_co-operation/Data_protection/.

[126] US Department of Justice. *The Freedom of Information Act of 1966 (amended in 2002)*. http://www.usdoj.gov/04foia/foiastat.htm.

[127] US Department of Justice. *The Privacy Act of 1974*. http://www.usdoj.gov/foia/privstat.htm.

[128] US Department of Justice. *Overview of the Privacy Act of 1974*, May 2002. http://www.usdoj.gov/04foia/04_7_1.html.

[129] US Department of Justice. *Freedom of Information Act Guide*, May 2002. http://www.usdoj.gov/oip/foi-act.htm.

[130] American Association of Motor Vehicle Administrators (AAMVA). *AAMVA National Standard for the Driver License/Identification Card*. http://www.aamva.org/Documents/stdAAMVADLIDStandrd000630.pdf.

[131] US Department of the Treasury. *The Use and Counterfeiting of United States Currency Abroad*, March 2003. http://www.treasury.gov/press/releases/docs/counterfeit.pdf.

[132] Europa: The European Union On-line. *Protection of the euro: Pericles programme*, January 2002. http://europa.eu.int/scadplus/leg/en/lvb/l33151.htm.

[133] European Parliament. *Directive 95/46/EC of the European Parliament on the protection of individuals with regard to the processing of personal data and on the free movement of such data*, 1985. http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=EN&numdoc=31995L0046&model=guicheti.

[134] Fabien A. P. Petitcolas, Ross J. Anderson, and Markus G. Kuhn. Information hiding – A survey. *Proceedings of the IEEE*, 87(7):1062–1078, July 1999.

[135] Dr. Despina Polemi. *Review and evaluation of Biometric Techniques for Identification and Authentication*, 1997. http://www.cordis.lu/infosec/src/study5.htm. This report

has been prepared for the European Commission by their authors and are placed on this web site to ensure the widest possible dissemination of the results of our work on ETS. However, please note that the European Commission does not necessarily endorse the content or conclusions of these reports.

[136] Cato Institute Report. *A National ID System: Big Brother's Solution to Illegal Immigration*, September 1995. http://www.cato.org/pubs/pas/pa237es.html.

[137] Eric Rescorla. *SSL and TLS: Designing and Building Secure Systems.* Addison Wesley Professional, 2000.

[138] Deutsche Bank Research. *Biometrics-Hype and Reality*, 2002. http://www.dbresearch.com/PROD/999/PROD0000000000043270.pdf.

[139] Consumer.gov Your resource for consumer information from the federal governement. *Identity Theft.* http://www.consumer.gov/idtheft/.

[140] Ronald L. Rivest. *Cryptography and Security.* http://theory.lcs.mit.edu/~rivest/crypto-security.html.

[141] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[142] Marc Rotenberg. *The Privacy Law Sourcebook 2001 - United States Law, International Law, and Recent Developments.* Electronic Privacy Information Center, first edition, 2001.

[143] Alfredo De Santis, Silvio Micali, and Giuseppe Persiano. Non-interactive zero-knowledge proof systems. In Carl Pomerance, editor, *Proc. CRYPTO 87*, pages 52–72. Springer-Verlag, 1988. Lecture Notes in Computer Science No. 293.

[144] Bruce Schneier. *Applied Cryptography: Protocols, Algorithms, and Source Code in C.* John Wiley & Sons, New York, second edition, 1996.

[145] Adi Shamir. How to share a secret. *Communications of the ACM*, 22:612–613, November 1979.

[146] Adi Shamir. Identity-based cryptosystems and signature schemes. In G. R. Blakley and D. C. Chaum, editors, *Proc. CRYPTO 84*, pages 47–53. Springer, 1984. Lecture Notes in Computer Science No. 196.

[147] Claude E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.*, 27:623–656, 1948.

[148] Claude E. Shannon. Communication theory of secrecy systems. *Bell Sys. Tech. J.*, 28:657–715, 1949.

[149] Victor Shoup. Practical threshold signatures. In B. Preneel, editor, *Proc. EUROCRYPT 00*, pages 207–220. Springer-Verlag, 2000. Lecture Notes in Computer Science No. 1807.

[150] Gustavus J. Simmons. The prisoners' problem and the subliminal channel. In D. Chaum, editor, *Proc. CRYPTO 83*, pages 51–67, New York, 1984. Plenum Press.

[151] Nigel Smart. *Cryptography, An Introduction.* McGraw-Hill, 2002.

[152] Douglas R. Stinson. *Cryptography: Theory and Practice.* CRC Press, second edition, 2002.

[153] The New York Times. *Finding Pay Dirt in Scannable Driver's Licenses.* http://www.nytimes.com/2002/03/21/technology/circuits/21DRIV.html.

[154] United Nations (UN). *General Assembly Resolution 45/95: Guidelines for the Regulation of Computerized Personnel Data Files*, 1980. http://www.un.org/documents/ga/res/45/a45r095.htm.

[155] American Civil Liberties Union. *National Identification Cards: Why Does the ACLU Oppose a National I.D. Card System?*, 1996. http://www.aclu.org/library/aaidcard.html.

[156] Network USA. *Social Security Number FAQ.* http://www.networkusa.org/fingerprint/ page6/fp-ssnfaq.htm.

[157] ACM U.S. Public Policy Committee (USACM). Codes, keys and conflicts: Issues in u.s. crypto policy. Technical report, ACM, June 1994. This report is available online in

html version at http://info.acm.org/reports/acm_crypto_study.html and in postscript version at http://info.acm.org/reports/acm_crypto_study/acm_crypto_study.ps.

[158] James L. Wayman. *Biometric Technology: Testing, Evaluation, Results.* http://www.engr.sjsu.edu/biometrics/publications_technology.html.

[159] James L. Wayman. *Biometric Identification Standards Research*, 1997. http://www.engr.sjsu.edu/biometrics/publications_fhwa.html. This report has been conducted by the U.S. National Biometric Research Center for the U.S. Federal Highway Administration.

[160] Alan F. Westin and Michael A. Baker. *Databanks in a Free Society: Computers, Record Keeping and Privacy.* Times Books, 1972.

[161] Andrew C. Yao. Protocols for secure computations. In *Proc. 23rd IEEE Symp. on Foundations of Comp. Science*, pages 160–164, Chicago, 1982. IEEE.

[162] Andrew C. Yao. How to generate and exchange secrets. In *Proc. 27th IEEE Symp. on Foundations of Comp. Science*, pages 162–167, Toronto, 1986. IEEE.