Privacy-Enhancing Technologies: The Path to Anonymity

Volume 1

Information and Privacy Commissioner/Ontario
Canada

80 Bloor Street West, Suite 1700
Toronto, Ontario, Canada M5S 2V1
416-326-3333
1-800-387-0073
Fax: 416-325-9195
TTY (Teletypewriter): 416-325-7539

Registratiekamer
The Netherlands

Sir Winston Churchilllaan 362
P.O. Box 3011
2280 GA Rijswijk, Netherlands
Tel. 011 (31) 70-3190190
Fax 011 (31) 70-3940460

August 1995

Table of Contents

Foreword

Enormous advances in information and communications technology around the world have meant an equally enormous growth in the amount of personal data accumulated by organizations in every jurisdiction. This development has increasingly jeopardized the privacy of those whose information is being collected. We believe this report will demonstrate the merits of minimizing personally identifying data, in a way that will help to restore individual privacy.

This is the first joint study ever undertaken by two organizations charged with the mandate of privacy protection in their own jurisdictions. Not only does it demonstrate the benefits of international cooperation on a subject that touches the lives of citizens in both Ontario and the Netherlands, it also shows clearly that issues of privacy protection are not bound by national borders.

This has been an opportunity to explore an exciting new area of study, to shed some light on an important issue where the future of privacy protection may lie. We hope the discussion and conclusions herein will be of benefit to organizations beyond our two jurisdictions, where measures to protect privacy are also being studied and debated.

Tom Wright,
Information and Privacy Commissioner
Ontario, Canada

Peter Hustinx, President
Registratiekamer,
The Netherlands

Back to Table of Contents

1.0 Introduction

At the present time, you are almost always required to reveal your identity when engaging in a wide range of activities. Every time you use a credit card, make a telephone call, pay your taxes, subscribe to a magazine, or buy something at the grocery store using a credit or debit card, an identifiable record of each transaction is created and recorded in a computer database somewhere. In order to obtain a service or make a purchase (using something other than cash), organizations require that you identify yourself. This practice is so widespread that it is simply treated as a given — an individual's identity must be collected and recorded in association with services rendered or purchases made. But must this always be the case? Are there no situations where transactions may be conducted anonymously, yet securely? We believe that there are and will outline a number of methods and technologies by which anonymous yet authentic transactions may be conducted.

Back to Table of Contents

1.1 Joint International Report: The Netherlands and Ontario, Canada

The Dutch Data Protection Authority (the "Registratiekamer" or RGK) and the Information and Privacy Commissioner for the Province of Ontario, Canada (IPC) are both privacy protection agencies that oversee compliance with their respective jurisdiction's privacy laws. The RGK and IPC decided to pool their resources and collaborate in the production of a report exploring privacy technologies that permit transactions to be conducted anonymously. The first international paper of its type includes a survey of companies that might be expected to offer such technologies, and organizations that might use them. In addition to anonymous transactions, the range of security features commercially available for use and the types of services actually being used by various organizations were also examined (see 2.1, Methodology). The RGK and IPC felt that a joint report outlining the practices followed in their respective jurisdictions would shed some light on this little-studied but extremely important area where the future of privacy-protection in an electronic world may lie.

Back to Table of Contents

1.2 Theoretical Basis for the Joint Report

Prior to this joint report with the IPC, the Registratiekamer, within its legally vested scope of powers and duties, conducted a study on the possibilities offered by conventional information systems and communications technologies for curbing the use of identifying data, particularly within information systems. This study, conducted in collaboration with the TNO Physics and Electronics Laboratory of the Netherlands Institute for Applied Scientific Research (TNO-FEL), formed the theoretical basis for the international study. The results of the Registratiekamer/TNO-FEL study are detailed in the companion volume to this report (Volume II).

Back to Table of Contents

1.3 Background

Consumer polls have repeatedly shown that individuals value their privacy and are concerned with its potential loss when so much of their personal information is routinely stored in computer databases, over which they have no control. Protecting one's identity goes hand in hand with preserving one's ability to remain anonymous — a key component of privacy. While advances in information and communications technology have fuelled the ability of organizations to keep massive amounts of personal data, this has increasingly jeopardized the privacy of those whose information is being collected. Minimizing **identifying** data would restore privacy considerably, but would still permit the collection of needed information.

When assessing the need for identifiable data during the course of a transaction, the key question one must start with is: how much personal information/data is truly required for the proper functioning of the information system involving this transaction? This question must be asked at the outset — prior to the design and development of any new system. But this is not the case today. This question is rarely asked at all since there is such a clear preference in favour of collecting identifiable data, "the more the better." However, with the growth of networked communications and the ability to link large numbers of diverse databases electronically, individuals will become more and more reluctant to leave behind a trail of identifiable data. What is needed is a paradigm shift away from a "more is better" mindset to a minimalist one. Is it possible to minimize the amount of identifiable data presently collected and stored in information systems, but still meet the needs of those collecting the information? We believe that it is.

The technology needed to achieve this goal exists today. We will describe some of the privacy technologies that permit one to engage in transactions without revealing one's identity by introducing the concept of an "identity protector." The notion of "pseudonymity" will also be introduced as an integral part of protecting one's identity. These technologies are available now and are within our reach; what is needed is the **will** to implement privacy technologies over the tracking technologies that are in use today.

When organizations are asked what measures they have in place to protect privacy, they usually point to their efforts at keeping information secure. While the use of security measures to prevent unauthorized access to personal data is a very important component of privacy, it does not equal privacy protection. The latter is a much broader concept which starts with the questioning of the initial collection of the information to ensure there is a good reason for doing so and that its uses will be restricted to legitimate ones that the data subject has been advised of. Once the data have been collected, security and confidentiality become paramount. Effective security and confidentiality will depend on the implementation of measures to create a secure environment.

Alternatively, instead of restricting the focus to security alone, a more comprehensive approach would be to seek out ways in which technology may be used to enhance the protection of informational privacy or data protection. We use the term "privacy technologies" to refer to a variety of technologies that safeguard personal privacy by minimizing or eliminating the collection of identifiable data.

Not only are measures that safeguard privacy becoming an important mark of quality, but increasingly, consumers are demanding that organizations pay attention to their privacy concerns. Social acceptance of demands for one's personal information, without adequate assurances of protection, appears to be on the decline. Not only do consumers wish to maintain control over their personal data and be informed of its uses, but insufficient protection will be reason enough for consumers to take their business elsewhere — to companies that follow privacy-protective practices.

Back to Table of Contents

1.4 Privacy Laws and Codes of Conduct

Respect for individuals' privacy, particularly with respect to the computer processing of personal data concerning one's self, is a fundamental principle underlying data protection. In Europe, data protection principles may be found in several instruments such as the Council of Europe's Convention 108 (Treaty for the protection of persons with regard to automated processing of personal data, Council of Europe, January 1981 (1988 Official Journal of Treaties, 7). One of the objectives of these principles is to ensure that personal privacy is safeguarded when new information technology applications are developed. The principles are reflected in various

European laws and regulations such as the Dutch Data Protection Act (WPR) and the draft EU-directive SYN 287. In addition, the OECD's Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (September, 1980) is internationally acclaimed as a "code of fair information practices" with respect to the treatment of personal information.

One of the basic principles in both the OECD guidelines and Convention 108 is the principle of "purpose specification." The quantity and nature of personal data that an organization is permitted to collect is limited by the purpose of the collection. The primary rule is that the data be relevant and sufficient, but not excessive for the stated purpose. In other words, the personal information to be collected must be needed to carry out the stated purpose.

This principle also seeks to ensure that restraint is exercised when personal data are collected. In accordance with this principle, one may question when identifying data is being sought from individuals where it is not necessary to do so. This is associated with the "use limitation principle," where the purpose specified to the data subject at the time of the collection restricts the use of the information collected. Thus, the information collected may only be used for the specified purpose (unless consent has been obtained for additional uses).

Another important data protection principle is "transparency" or "openness." People have the right to know what data about them have been collected, who has access to that data, and what the data are being used for. The principle of transparency simply means that people must be made aware of the conditions under which their information is being kept and used.

The principle of transparency may also be used to explain the logic behind the data processing underlying a collection — asking for identifying information in a situation that does not strictly require it, must be questioned. Indeed, the collection and use of personal data for identification purposes when not truly necessary (where alternatives are available), cannot be supported in relation to the principles noted above. Since these data protection principles are incorporated into most privacy laws such as the Ontario Freedom of Information and Protection of Privacy Act and the Dutch Privacy Act (Wet persoonsregistraties), or EU-directive SYN 287, in some situations, the unnecessary collection of identifiable data may have a direct bearing on compliance with these statutes.

Back to Table of Contents

1.5 Information Systems

An information system is broadly defined as a system which provides organizations with the information required to conduct various activities. There are generally three types of information systems: transaction-processing systems, programmed decision-making systems, and decision-support systems. Transaction-processing systems collect and keep track of information relating to a transaction. Examples range from direct marketing systems, mail order catalogue purchasing systems, telephone records systems, and so forth.

Programmed decision-making systems process data in accordance with formal, structured procedures. The system is programmed, on its own, to complete the entire order from the time it is received through its entire processing, without any human intervention. Examples include hotel reservation systems, payroll accounting systems, money transaction systems for automatic teller machines, flight reservations systems, etc.

Decision-support systems assist in the decision-making process by using the information collected to either generate potential solutions or additional information to assist in the decision-making process. Examples include systems for calculating mortgages, management information systems, recommended itinerary systems, etc.

The one common feature of all these systems is that their use entails the collection and processing of personal information. Whenever an individual (the user) comes into contact with an information system, the service provider usually requires that they identify themselves through some means.

1.5.1 The Structure of an Information System

The elements of an information system consist of the following: user representation (containing a means of identification), service provider representation, and a database(s) containing the data required for the information system to function. The database usually consists of two files, the privileges file and the audit file. The privileges file contains the user's privileges (which the service provider would check to see whether he/she was eligible for the various services offered). The audit file records the use of the information system and can charge the user of a service or track what services were used by whom, at what times. Using the example of a health club, the privileges file would contain a record of the user's entitlements, i.e., that a particular user was entitled to use certain services such as the use of the tennis facilities (five times a month), the squash courts (four times a month), but not the golf course (which the user had not paid the required additional fee). The audit file would keep track of the actual uses of the various privileges and charge the user a per-use fee for any additional services that the user was not entitled to (i.e., playing golf).

A user representation can take the form of an account number, a membership card or a smart card. A service provider representation represents the interests of the organization and controls access to the organization's various resources (through passwords; tiered levels of authorization to increasingly sensitive information, etc.).

1.5.2 The Processes in an Information System

The use of an information system entails the following processes: authorization; identification and authentication; access control; auditing and accounting. We refer to a process as an exchange of information between two or more elements within the information system. In conventional systems, the user's identity is usually viewed as being essential to the performance of all the above processes. For example, one's identity is used within the authorization process to identify and record instances involving a user's privileges. Once the user's identity has been collected, it will travel throughout the various processes involved in the information system. We will suggest that this need not be the case. One must examine whether the user's identity is truly required for the operation of each of these processes.

We will propose that a user's identity is only necessary during the processes of authorization and accounting. For the processes of identification and authentication, access control, and audit, a user's identity may be sheltered through some type of "identity protector." We will describe how technologies of privacy may be used to separate one's true identity from the details of one's transactions through the use of "pseudo-identities."

Back to Table of Contents

1.6 The Identity Protector

An identity protector may be viewed as an element of the system that controls the release of an individual's true identity to various processes within the information system. Its effect is to cordon off certain areas of the system which do not require access to true identity. The identity protector works in such a way as to protect the interests of the user. One of its most important functions is to convert a user's actual identity into a pseudo-identity — an alternate (digital) identity that the user may adopt when using the system.

Alternate identities also exist in conventional systems such as bank account numbers, social insurance/social security numbers, health insurance numbers, etc. But these cannot be viewed as pseudo-identities since they may easily be linked to one's true identity. In the privacy-protective systems of the future, the identity protector would most likely take the form of a smart card controlled by the user, which could generate pseudo-identities as desired.

An identity protector performs the following functions:

- generates pseudo-identities as needed;
- converts pseudo-identities into actual identities (as desired);
- combats fraud and misuse of the system.

Since the identity protector is under the control of the user, he/she can set it to perform a variety of functions such as revealing one's

actual identity to certain service providers but not to others. When an identity protector is integrated into an information system, the user may use the services or engage in transactions anonymously, thereby elevating privacy to an all-time high.

When an identity protector is introduced into an information system, two domains are created: an identity domain and a pseudo domain — one in which the user's actual identity is known and accessible, and one in which it is not. The identity protector functions so as to separate the two domains and may be applied anywhere in the system where personal data can be accessed. A simple guideline for designers of new information systems is to minimize the identity domain wherever possible and maximize the pseudo domain.

The identity protector permits the designer of a system to minimize the personal data stored in a database. In effect, the service provider would not record the user's privileges or activities under their true identity but rather, under their pseudo-identity. While the service provider must be able to determine what the user is authorized to do, this may be accomplished without learning the user's true identity. Since the identity protector acts somewhat as an intermediary between the user and the service provider, it must be trusted by both parties. However, there is no disadvantage to service providers since their ability to verify the user's privileges/eligibility for services remains intact. Indeed, the identity protector is designed in a way which prevents fraud and improper use. The latter can take various forms ranging from prevention, detection, and correction. It can prevent the user from using his/her anonymity as a shield to commit fraud, and, in appropriate circumstances, can lead to having the true identity of the user being revealed to the service provider and/or the authorities. For example, cryptographic techniques may be used to prevent a sum of money (digital cash) from being used anonymously more than once, or a service being used but not charged to the user.

Back to Table of Contents

1.7 Implementation Techniques

Thus far we have discussed a theoretical approach using the concept of an identity protector in the design of systems that would permit individuals to interact anonymously with service providers. Below, we will outline several specific techniques for introducing an identity protector into an information system. Specifically, encryption techniques involving digital signatures, blind signatures, digital pseudonyms and trusted third parties, will be described. Additional readings are provided in the companion text (Volume II) for those wishing to explore these techniques in greater detail.

1.7.1 Digital Signatures

A digital signature is the electronic equivalent of a handwritten signature. Just as a signature or sealing wax on a document is proof of its authenticity, a digital signature provides the same, if not better, authentication. It provides the necessary assurance that only the individual who created the signature could have done so, and it permits all others to verify its authenticity. A particular type of encryption, "public key encryption," considered to be the most reliable and secure form of encryption ever developed, forms the basis for digital signatures.

In a public key system, two keys are created for each individual — one private one public. The private key is known only to the individual while the public key is made widely available. When an individual encrypts a document with his or her private key, this is the equivalent of signing it by hand since the private key is unique to that individual alone. Any third party may decrypt the message using the individual's public key, which corresponds only to his/her private key. If the document is successfully decrypted, then one has the necessary assurance that it could only have been created by that individual. Otherwise, one would not have been able to decode it. Digital signatures thus provide proof of a document's authenticity — that the document originated from the sender. For a more detailed description of this cryptographic technique, please refer to Volume II.

1.7.2 Blind Signatures

The blind signature, created by David Chaum of Digicash, is an extension of the digital signature, but with one critical feature added: it ensures the anonymity of the sender. While digital signatures are intended to be identifiable and to serve as proof that a particular

individual signed a particular document, blind signatures provide the same authentication but do so in a non-identifiable manner. The recipient will be assured of the fact that the transmission is authentic and reliable, but will not know who sent it. One application involving blind signatures is the use of "digital cash" which may be used as an electronic form of payment that can be transmitted over computer networks. Just as cash is anonymous, digital cash is anonymous in that it cannot be traced to a particular individual — it is considered to be "unconditionally untraceable." However, the service provider is assured of its authenticity; all that is missing is the ability to link the transaction to a particular person. In describing his system of blind signatures, Chaum adds that it also provides much-needed protections against fraud and abuse of the system. For a detailed description of blind signatures, we refer you to Volume II.

1.7.3 Digital Pseudonyms

A digital pseudonym is a method of identifying an individual through an alternate digital pseudo-identity, created for a particular purpose. It permits users to preserve their anonymity by concealing their true identities. While users are not "known" to service providers in the conventional sense, they are, nonetheless, known by their pseudonyms, for the purposes of conducting transactions.

Digital pseudonyms build upon the blind signature technique. However, in this instance, it is the service provider who assigns privileges to a given pseudonym (user) by creating a blind signature. The user keeps the allotted privileges (for example, five uses of the tennis courts per month), and uses them as desired. Again, we refer you to Volume II for a more detailed discussion of this subject.

1.7.4 Trusted Third Parties

A trusted third party is the term used for an independent third party who is trusted by both the user and service provider alike (comparable to a "digital attorney"). This party can be entrusted with keeping such things as the master key linking digital pseudonyms with the true identities of their users. The trusted party knows that the relationship between an user's true identities and pseudo-identities must be kept completely secret. However, if certain conditions require it, the trusted party will be permitted to reveal the user's identity (under previously agreed upon terms) to a service provider. The conditions under which an individual's identity would be revealed must be known to both the user and service provider prior to entering into an agreement with the trusted party.

1.7.5 Moving from Conventional Technologies to Privacy Technologies

The most important prerequisite to moving in the direction of privacy technologies is to start by asking whether identifiable information is truly needed when a new information system is being contemplated, or an existing one upgraded. If the client, the systems designer and supplier ask this question right from the start, privacy is sure to be addressed. The creation of some form of identity protector within the system must also be a crucial part of the design phase. To recap, the identity protector is a term for all those functions within an information system that protect the user's true identity, such as the creation of pseudo-identities. A pseudo-identity is a pseudonym that the user may assume for the purpose of engaging in a particular transaction or service. The guiding principle should always be to keep the identity domain as small as possible, thereby maintaining the absolute minimum amount of identifiable information. The actual implementation of an identity protector may be done in a number of ways, usually involving advanced encryption techniques (best left to systems designers and technical staff).

The point to emphasize is that it is indeed possible to collect less identifiable data, or unlink the data from an individual's true identity through the use of pseudo-identities. It is only the application of privacy technologies that is lacking, not the technologies themselves.

Back to Table of Contents

2.0 IPC–RGK Joint Survey

The primary purpose of the joint survey was to assess the types of privacy technologies commercially available for keeping anonymous the identity of individuals during the rendering of services. We also wanted to establish an estimate of the companies offering such technologies, as well as assessing the types of technologies available for protecting the security and confidentiality of information transmitted electronically. These technologies could be offered through a variety of means such as hardware, software, network communications, encryption products or application systems design practices. Our only criteria was that they not be at a theoretical stage, in other words, they had progressed beyond the lab test phase; we wished to survey applications and products that were already built and being offered in the marketplace.

Another element of the survey involved the potential users of these products, in an effort to determine the extent to which these technologies were actually being used and the degree of awareness or interest in their availability.

The specific objectives of the survey were:

- to establish an estimate of the number of information technology (IT) providers that were making privacy technologies available to their customers;
- to determine whether organizations falling under the jurisdictions of the RGK and IPC were using or considering privacy technologies to keep the identity of individuals anonymous during service delivery.

Back to Table of Contents

2.1 Methodology

The sample consisted of 100 IT providers; 50 companies from the Netherlands and 50 from Ontario, Canada. The following types of IT providers were sampled: product vendors; service providers (application and systems software developers, consulting firms, online service providers, security associations). A small sample of technology users was also included; these consisted of organizations falling under the respective jurisdictions of the RGK and IPC.

The IT providers and user organizations selected were contacted by telephone and asked to participate in a short survey. A copy of the questionnaire may be found in Appendix A. Detailed product brochures and technical specifications were also requested.

Back to Table of Contents

2.2 Findings

2.2.1 IT Providers

Tables 1 to 3 present the responses obtained from the IT providers sampled.

Table 1

Privacy Technologies to keep Identity Anonymous

Information Technology:	Yes	No

	IPC	RGK	IPC	RGK
	%	%	%	%
Available now	10.0	33.3	90.0	66.7
May be available in the future	8.0	28.6	92.0	71.4
Asked to develop	6.0	33.3	94.0	66.7

Table 1 presents the percentage of IT providers presently offering or developing privacy technologies that permit an individual's identity to be kept anonymous during service delivery. In Ontario, only 10% of the providers sampled offered products capable of keeping identities anonymous. This contrasts with a much larger number of Dutch companies, 33.3%, who presently offer anonymous technologies. In Ontario, only four providers contacted had any thoughts of developing such technologies in future. We should add, however, that the types of anonymous technologies referred to here are not the kind that use encrypted techniques such as digital pseudonyms.

In the Netherlands, however, one company, Digicash, is known for its work in developing anonymous technologies based on public key encryption and blind signatures. In Ontario, the main technology involved in maintaining anonymity was the prepaid or stored-value smart card — a card that is "loaded up" with cash and used anonymously. Such cards ensure anonymity because, like cash, they carry no personal identifiers. Another way in which anonymity was preserved was through the use of a "remailer" service wherein an individual's true identity was stripped from the data prior to being forwarded electronically to a third party. In the Netherlands, privacy technology tends to be more application driven as in the anonymous recording and processing of data in tax files, personal information systems, and payroll processing systems.

While the incidence of companies offering anonymous technologies was low, most respondents, when asked why they were not devoting more time and effort to such product development, said there was presently no demand for it from their customers. If there was, they would reconsider. Therefore, if customers' awareness of such technologies and their benefits (more privacy and anonymity) increased, so would the likelihood that suppliers would develop products offering these features.

Table 2 presents the percentage of IT providers who had developed or were considering developing products that protected the **security** of information transmitted electronically.

Table 2

Availability of Products that Protect the Security of Information

Information Technology:	Yes	No

	IPC	RGK	IPC	RGK
	%	%	%	%
Available now	56.0	81.0	44.0	19.0
May be available in the future	12.0	66.7	88.0	33.3
Asked to develop	32.0	76.2	68.0	23.8

Over half (56%) of the Ontario IT providers sampled had developed products that protected information security. Twelve per cent of the providers sampled were considering developing such products in the near future. A much higher percentage of the Dutch providers sampled — 81%, presently offer such technologies. Thus, the majority of companies contacted offered products or applications that protected the security of information transmitted electronically.

Security examples involving financial transactions such as interactive banking were identified as being in special need of stringent security measures. Large multinationals, banks and other financial institutions were regarded as being the driving force behind security-enhancing technologies. This was due to increases predicted in the electronic transmission of sensitive financial transactions and the use of electronic currency. It was said that the highest degree of confidence needed to be associated with the authenticity and integrity of financial transactions, requiring the development of "fail-safe" information systems, to boost consumer trust and confidence.

There was a substantial difference in the responses of the Ontario and Netherlands IT providers regarding the offering of such products in the future. Two-thirds of the Netherlands companies sampled expected to be developing encryption products for information security in the near future. In Ontario, however, only 12% expected to be doing so. This difference could in part relate to how often these companies had been approached to develop such products. The companies in the Netherlands had more than twice as often been asked to develop such products than their counterparts in Ontario. This may have to do with a major public debate that took place in the Netherlands in 1994, regarding the use of encryption techniques, triggered by proposed legislation for the use of encryption in telecommunications services.

When asked specifically what types of technologies they used to protect security, cryptographic techniques such as encryption, both public key and symmetric systems, were named most frequently (DES; RSA; DES in conjunction with RSA or IDEA). Several references were also made to the use of PGP (pretty good privacy). The use of digital cash was also noted several times as a way to preserve anonymity during electronic transactions requiring electronic forms of payment.

In addition, each of the following was named at least once as a means of protecting information security: unique one-time transaction numbering systems, frequently changing encryption keys, and the use of smart cards for PIN code protection.

Not only was the need for security perceived to be critical during electronic transmissions, one respondent pointed to the growing need for secure transmissions during mobile (wireless) telephony. Since wireless communications are no longer restricted to telephones, extending well beyond to wireless data transmissions and networked communications, there will be a strong demand for the development of secure wireless networks (one example being the Mobitex Network which makes use of data encryption techniques).

Unlike privacy technologies designed to preserve anonymity, products relating to the protection of information security were far more widely available. Several Dutch companies felt that politically, the demand for information security would increase, perhaps making legislation in this area a necessity. Information highway applications such as e-mail and electronic service delivery were especially viewed as being in need of stringent security controls.

One reason for the low degree of awareness of **anonymous** technologies, contrasted against the high awareness of technologies protecting information security, is the fact that technology has commonly been associated with keeping information secure. Protecting privacy, however, in the form of keeping an individual's identity anonymous, is a relatively new concept — one that has rarely been considered by vendors and suppliers, or requested by customers. The much-needed awareness of the existence of such technologies among users is virtually non-existent at the present time.

Table 3 presents the types of technologies used in the products offered by IT providers.

Table 3

Types of Information Technologies Offered

Information Technology	To Preserve Anonymity		To Protect Information Security	
	IPC	RGK	IPC	RGK
	%	%	%	%
Encryption	—	12.5	100.0	90.0
Smart Cards	60.0	25.0	—	—
Application Design	20.0	50.0	—	10.0
Anonymous Remailing	20.0	12.5	—	—
Total	100.0	100.0	100.0	100.0

In Ontario, smart cards, or specifically prepaid "stored-value cards" appear to be the most popular method of preserving anonymity, while encryption was the key technology (100%) used to protect information security. In the Netherlands, the use of encryption was also comparably high (90%) for security purposes. But the use of smart cards in the Netherlands (25%), was lower than in Canada (60%). The RGK expects to see similar developments in the Netherlands in the near future. A greater use of applications design (50%) was made in the Netherlands to preserve anonymity. By "application design" we mean the privacy protective design features of applications developed in the creation of new information systems.

All providers in the area of systems development indicated that they could easily design features into any system to keep an

individual's identity anonymous (through such techniques as encryption and smart cards), if their clients asked for these features. Again, however, what is lacking is the demand for such features, which in turn is related to the lack of awareness of their benefits among users.

2.2.2 IT Users

Due to an extremely low response rate from the organizations contacted (government, trade and industry), no quantitative reporting of users' responses will be presented. Instead, a narrative description of respondents' views will be presented below. It should be noted that only one of the government organizations contacted in Ontario was making any effort to protect anonymity: an internal procedure had been developed that allowed the identities of staff to be kept anonymous while communicating with the organization's human resources department. Of the companies contacted in the Netherlands, one bank was using a software application to produce internal reports where data about clients was kept anonymous; another bank used pseudo-identities (numbers) instead of the true names of individuals; and another company separated the flow of personal from other data — maintaining a link between the two through a numerical system.

2.2.3 Respondents' Views

Most respondents (providers and users alike) felt that there was increasing pressure on both the public and private sectors to provide additional security for electronic transmissions of information, especially in the areas of electronic commerce and electronic payment systems. While overall awareness in this area appears to be growing, the same cannot be said for technologies that preserve anonymity. Since these technologies are relatively new and difficult to grasp conceptually, this should come as no surprise. In addition, one could speculate that most organizations would not wish to have anonymity maintained — quite the contrary, they would like to collect more, not less, identifiable information (that would permit the development of personal profiles and the tracking of purchasing patterns and activities). The public must be made aware of the availability and benefits of privacy-enhancing technologies so that they will be better positioned to make informed choices.

Back to Table of Contents

2.3 General Observations

- There was a general lack of understanding as to the difference between privacy and confidentiality; survey respondents tended to use the two words interchangeably. While confidentiality (keeping information secure and inaccessible to unauthorized parties) is an important component of privacy, it is just that — one component. The areas covered by privacy are much broader, extending from limitations on the initial collection of personal data (whether it is truly needed), to restrictions on its use to the purpose specified, to prohibitions on any secondary uses (without the express consent of the data subject).
- When asked what products/technologies were offered to protect privacy, most IT providers identified security-related products that served to keep personal data confidential (secure from third party interception). The importance of maintaining security and confidentiality were clearly understood. But this was not the case for the importance of maintaining privacy through anonymity. While the great majority of IT providers expected to see dramatic growth in the area of security-related technologies (firewalls, encryption products, etc.), they had no similar expectations for the growth of privacy-enhancing technologies.
- It was clear that IT providers would respond to market forces: if the demand is there, they will build it. Even a minimal growth in awareness of privacy technologies and acceptance of anonymous services could yield a demand significant enough to warrant the attention of IT providers. Public awareness and education are key.
- Information technology will be one of the primary ways to ensure the protection of privacy in electronic transactions and networked communications.

Back to Table of Contents

2.4 Discussion of Findings

When embarking upon this joint project, the Registratiekamer and IPC wished to outline a variety of technologies that protected privacy by preserving an individual's anonymity during service delivery. Another wish was to provide some empirical evidence of what was generally believed to be true — that privacy technologies known to be in existence were not widely known or used by providers or the public. Thus, the findings outlined in this report came as no surprise since this proved to be the case. Awareness of these technologies tended to be somewhat higher in the Netherlands, but not as high as one might hope in light of the public discussions about encryption which had taken place, and the presence of a prominent company (Digicash) that specializes in the development of such technologies.

While one may be disheartened by the overwhelming lack of awareness (particularly in Ontario), one may also view this as a great opportunity — awareness can increase dramatically if the word gets out. The challenge that must be faced is how to "get the word out there," and get it out to enough people to make a difference. Due to the present lack of demand for such technologies from the public, there is little reason for IT providers to develop them. The need for public education cannot be over-emphasized: this is one area which is unlikely to be discovered (or understood) without some assistance. Nor can we expect service providers or organizations to encourage the development of such awareness — the task will fall upon privacy commissioners and privacy advocates. We hope that efforts such as this joint report will contribute to increasing awareness, and to begin going down the road to greater understanding.

Back to Table of Contents

3.0 Conclusions and Recommendations

In this report, it has been submitted that an individual's identity is only truly necessary for certain parts of an information system, namely, during the processes of authorization and accounting. We have introduced the concept of an "identity protector" and described how technologies of privacy involving encryption and digital pseudonyms may be used to separate one's true identity from the details of one's transactions and communications. Such practices would lead to far fewer collections of identifiable information and would thus greatly enhance the protection of privacy.

Among the challenges that lie ahead will be the reluctance of both public and private sector organizations that wish to collect **more**, not less, identifiable information (until the benefits of collecting less are understood and organizations come to realize that identifiable information is not always necessary for their activities). Add to this a public which generally lacks awareness of the benefits to be had through the use of anonymous technologies (never having been exposed to them), and the challenge grows even larger. To help meet these challenges, we make the following recommendations.

3.1 Recommendations

1. International information systems design standards should be developed incorporating the need to examine whether an individual's identity is truly required for the operation of various processes within the system. Attention should be explicitly directed to the introduction of an identity protector, which functions to separate the identity domain from the remaining pseudo-domains.
2. At the design stage of any new information system, or when revising an existing one, the collection and retention of identifiable personal information should be kept to an absolute minimum. Only that which is truly needed should be collected and maintained in an identifiable form (as opposed to a pseudonymous form).
3. Consistent with the privacy principle that information systems should be transparent and open to view to data subjects, they should also provide users with the ability to control the disclosure of their personal information. Data subjects must be placed in a position to decide for themselves whether or not their identity should be revealed or maintained in an information system.
4. Data Protection Commissioners, Privacy Commissioners and their staff should make every effort to educate the public and raise levels of awareness in the area of privacy-enhancing technologies. The use of privacy technologies by public and private sector organizations should be also encouraged. The message that it is now possible to preserve an individual's

anonymity during service delivery should be included in all public outreach efforts. The benefits to be derived to individuals from the use of anonymous technologies are far-reaching and will ensure the continuation of privacy protection in a fully networked world.

5. Data Protection and Privacy Commissioners should ask the parties involved to review the use of identifiable data in light of privacy protection principles and make use of privacy-enhancing technologies wherever possible. The unnecessary collection of identifiable data, where appropriate, should give rise to further action to promote compliance with existing statutes.

Back to Table of Contents

Appendix A: Survey Questionnaire

To IT Providers:

1. Do you presently supply/sell any technologies that allow for an individual's identity to remain anonymous during the rendering of services?

2. Are you developing/planning to offer for sale such technologies in the near future?

3. Have you ever been asked for or approached to develop such a technology?

4. Do you presently supply/sell any technologies that protect the security/confidentiality of information transmitted electronically (from interception by unauthorized third parties)?

5. Are you developing/planning to offer for sale such technologies (to protect the security/confidentiality of information transmitted electronically) in the near future?

6. Have you ever been asked for or approached to develop such a technology?

To Government Organizations/Users:

7. Are you presently using any technologies that allow an individual's identity to remain anonymous during the rendering of services?

8. Are you contemplating offering such services to your clients in the near future?

9. If such a technology was made available to you at a reasonable cost, would you recommend its use?

10. Are you presently using any technologies to protect the security/confidentiality of information transmitted electronically (from interception by unauthorized third parties)?

11. Are you contemplating offering such services to your clients in the near future?

12. If such a technology was made available to you at a reasonable cost, would you recommend its use?

IF YES TO ANY OF THE ABOVE: Ask the following questions to all groups

13. Please describe the technologies referred to above (get exact details re. **type** of encryption, ie. symmetric or public key encryption,

etc.)

14. What are the specific privacy-enhancing features offered through these technologies?

15. How can these features be used to:

a) keep the identity of an individual anonymous? or

b) protect in some way an individual's identity from being revealed?

c) protect the security/confidentiality of the information transmitted?

16. Do you have any examples of existing or proposed applications of these technologies?

17. Could we have more information about these technologies/applications? If possible, could you please provide us with:

- any brochures or promotional material;
- technical specifications;
- automated demonstrations.

18. What are your views as to the extent to which the use of these technologies will spread? Could you explain the reasons why?

19. If you had a crystal ball, what projections would you make for the use of privacy-enhancing technologies in the next five years?

20. Are you aware of Ontario's privacy protection laws?

21. Do you have any comments regarding the potential effects of technology on privacy?

Back to Table of Contents

Appendix B: Table of Contents, Volume II

Contact the IPC for details on how to obtain a copy of Volume II.

Back to Table of Contents

Appendix C: Project Team

Registratiekamer, Rijswijk, The Netherlands:

John Borking

Vice President

Huib Gardeniers

Legal Advisor

Henk Van Rossum

Staff Member Technology

Information and Privacy Commissioner, Ontario, Canada:

Ann Cavoukian

Assistant Commissioner

John Brans

Manager of Compliance

Noel Muttupulle

Compliance Supervisor

Nick Magistrale

Compliance Officer

TNO Physics and Electronics Laboratory, Telematics and Information Security Group, The Hague, Netherlands:

Joost Meijers

Paul Overbeek

Paul Verhaar

Back to Table of Contents

Privacy-Enhancing Technologies: The Path to Anonymity

Volume II

Registratiekamer
The Netherlands

August 1995

Registratiekamer

Sir Winston Churchilllaan 362
P.O. Box 3011
2280 GA Rijswijk, Netherlands
Tel. 011 (31) 70-3190190
Fax 011 (31) 70-3940460

Table of Contents

1.0 Introduction

In this part of the report we present the theoretical study conducted by the Registratiekamer in collaboration with the TNO Physics and Electronics Laboratory (TNO–FEL).

1.1 Methodology

This study is based on two central questions:

• What conditions must be kept in mind when engineering an information system in order to guarantee that the system can be used effectively and efficiently without revealing the user's identity?

• What types of information and communication technology can contribute towards achieving this goal?

TNO–FEL's role in the study was to make an inventory of the information and communication technological (ICT) possibilities to separate the use of the information system from the identity of the user. A few models are presented to serve as examples to designers, developers and marketers when setting up information systems. The Registratiekamer outlined the general framework and guidelines of this study and provided assistance.

1.2 Overview of Volume II

Chapter 2 defines the concept of information systems. There is a great diversity in information systems, and the system used generally depends on the environment in which it functions. Each information system has certain basic elements and processes in common. These elements and processes can be used to construct a model of an information system, which can then be used to examine whether the various information system processes contain identifying personal data.

Chapter 3 takes a closer look at the privacy enchancing technology concepts introduced in chapter 2. The information system model is expanded in several places to include identity protectors to safeguard users' privacy. Examples illustrate how these models with integrated identity protectors are used.

Chapter 4 explores a number of potential techniques for the implementation of privacy enchancing technology in information systems. The end of the chapter introduces a flow diagram for the design of new information systems.

2.0 Information Systems and Identity Use

The current generation of information systems make use of the user's identity at various points in the system. In this report, a "user" is defined as someone who uses the information system, in whatever capacity. The central question is whether it is necessary for the system to know the user's identity. A model is presented to examine how an information system functions. The model developed serves as a basis for further elaboration on the privacy enchancing technology concept. It is essential when developing a model to know what the term "information system" actually entails. What is the purpose of such information systems, how do they work and what are they made of? The next section will address these questions. Subsequently, the difference is explained between the current generation of information systems and information systems based on privacy technology.

2.1 What is an information system?

Information systems serve to provide people with information required for performing goal-oriented activities [42]. "Performing" can be understood in the broadest sense of the word meaning the planning, conducting and monitoring of specific activities. The scope and nature of information systems display a great degree of diversity, however. They may support a process only involving a few people. Such information systems are generally limited in structure and fairly transparent. On the other hand, there are also information systems utilized by people, who do not necessarily belong to the same organization. Nor does the information system have to be limited to one organization. An information system for internal use can also be used for interorganizational and international data flow. Information exchange has thus been greatly simplified and intensified. The developments surrounding the "information superhighway" [43] will swell the flow of interorganizational and international data even further.

The different information systems can be divided into three types: transaction-processing systems, programmed decision-making systems and decision-support systems [42]. The transaction-processing systems register a transaction. Examples include:

- entrance registration systems
- mail registration systems
- order registration systems
- telephone records
- pharmacists' systems.

Programmed decision-making systems process data according to formalized, structured procedures. The system completes the entire order, from the time of its receipt to its processing, often without any human involvement. Examples include:

- hotel booking systems
- wage accounting
- money transaction systems for automatic teller and payment machines
- financial aid systems
- (international) flight reservation systems
- hospital information systems
- ticket systems
- voting machines.

As the name suggests, decision-support systems assist decision-makers in making decisions. These systems use the information entered to generate potential solutions or other information on the basis of which the decision can be made. Examples include:

- systems for calculating mortgages
- direct marketing systems
- address systems
- recommended itinerary systems
- Management Information Systems.

The list of information systems could be expanded to include many more examples. Although the systems have widely diverging purposes, they have one thing in common: their use entails personal data processing. Obviously, each information system operates within a certain environment, and thus has a relationship with that environment, such as links with other automated or non-automated information systems as well as the person using the systems and internal and external organizations.

Information systems consist of four components: organization, personnel, procedures and technology. All of these components are crucial to the proper functioning of the information system. This study focuses on the technical set-up of information systems, which determines the degree of protection of the user's privacy. Where necessary, attention will also be paid to the other components.

2.2 Conventional and privacy information system

The terms "conventional information systems" and "privacy information systems" are used to denote the information systems mentioned in the preceding section and those which protect user's privacy. Conventional information systems thus generally record a high amount of information with a high identification content. This means, of course, that it is easy to link the data to a private individual. Privacy information systems are systems which only reveal the user's identity to combat fraud.

There are two options for privacy information systems. The first is not to generate or record data at all. The second option is to not record data unique to an individual — identifying data. The absence of identifying data renders it nearly or completely impossible to link existing data to a private individual. A combination of the two options offers a third alternative.

By applying the potential forms of privacy enchancing technology, a conventional information system can be transformed into a privacy information system. The study focuses on the second possibility offered by privacy enchancing technology: omitting data linked to a person, i.e. identifying data.

2.3 Identity in the information systems

To determine whether a user's identity is, in fact, required for the adequate working of an information system, its functions must be evaluated and the following questions answered: Which elements of an information system is identity used for? For which processes? The following sections will first define the elements and then the processes of an information system. Each time individual processes are discussed, the following question will be asked: Is the user's identity required for the information system to function properly?

2.3.1 Elements of the information system

A (technical) model of an information system contains four separate elements: user representation, service-provider representation, database, and services (see Figure 2.1).

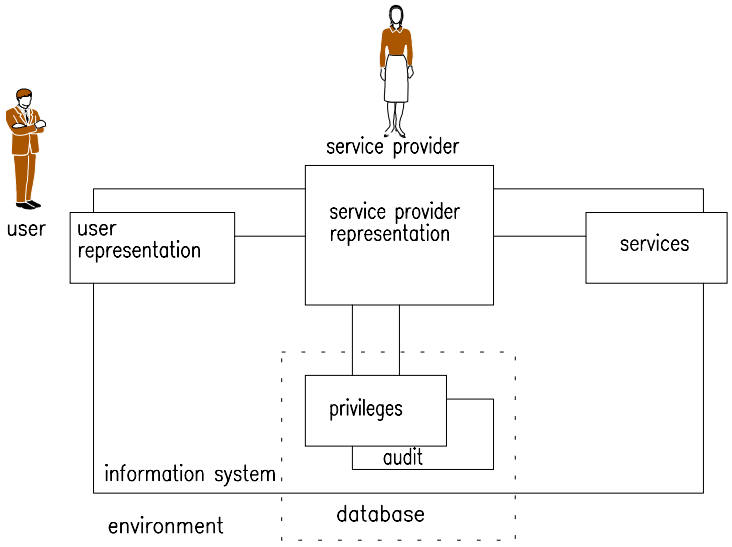

Figure 2.1: A (technical) model of an information system

The user representation is the representation of the user — a private individual — within the information system. A user representation will generally be a process that performs certain functions at the user's request, and consists of a technical interface between the information system and the user. Via this interface, the user can control the user representation.

The service-provider representation is the internal representation of the agency or business from whom the user procures a service. The service-provider representation within the information system represents the person responsible for the system (e.g. the owner) and promotes the interests of the organization it represents. A key functionality of the service-provider representation is to control access to services. A service-provider representation can also collectively represent several businesses or organizations.

Services should be understood in the broadest sense. In many cases, these services will consist of information or information processing. Examples of services are: teletext and other databases for information collection, reading and writing of documents on a computer network, communication services, payments, etc. A service can also be a link to another (external) information system.

A database is the information system's internal (electronic) administration and contains the data required for the information system to function. The database controls the information system and is therefore not considered a service. Simple information systems do not even require a database: an example of such a service is teletext.

The database consists of two files: a privileges file and an audit file. The privileges file contains the users' privileges (equivalent to those of the user representation). The service-provider representation checks in the privileges file whether or not the user is authorized to access the various services of the information system. The audit file records the use of the information system and can be used to charge the user for the use of an information system, or, for instance, to check when, why and by whom an information system has been used.

Each element of the model may be partially outside of the (computerized) information system. All elements of the information system can interface with the system's environment, as outlined in section 2.1. An audit file could be printed on paper. A user representation could take the form of a smart-card.

Each line connecting two elements of the model is an interaction line. Adjacent elements can generate an interaction across that line, e.g. data exchange. Thus each interaction line poses a potential threat to user's privacy since identifying data can be spread through the system by each of these lines. The elements will generally interact as part of a process initiated when the information system is used. In order to determine whether the person's identity is required for these processes, the processes carried out within an information system and their functions within the system as a whole must be clarified.

2.3.2 Processes in the information system

Use of an information system entails a number of processes: authorization, identification and authentication, access control, auditing and accounting. A process is an exchange of information between two or more elements within the information system, as indicated in the preceding section. Interaction lines connecting the elements are used for data exchange. The processes can take place independently of each other, with one process utilizing data generated by another process. Figure 2.2 shows the relationship between these processes. The processes of identification and authentication, access control and auditing take place entirely within the information system. The authorization and accounting processes have an interface with the environment.

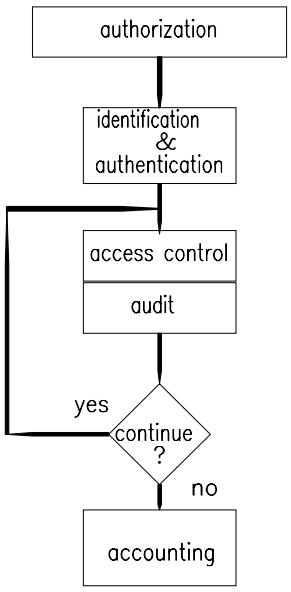

Figure 2.2: A possible order of processes in an information system.

Authorization is the allotment of privileges to the user. Before a user can use an information system for the first time, the service-provider determines the user's privileges and files this information in a database. User privileges are determined on the basis of user characteristics. The user is subsequently assigned a user representation within the information system. The service-provider representation links the user's privileges with his internal representation. A bank account number is a well-known example of internal representation.

The process of identification and authentication of a user representation is carried out when a user wishes to gain access to the information system via a user representation. In most information systems, the user introduces himself to the service-provider (identification), and then the service-provider checks the user's identity (authentication). The user uses the interface that is part of the user representation for identification and authentication. A common method of identification is to enter a user ID, although even the possession of a bank card can be considered identification. Authentication then takes place when a password or, in the case of the bank card, personal identification number (PIN) is entered.

Access control is a continuous process. The service-provider representation checks whether the user representation is authorized for each service provided. In this way, the service-provider representation prevents unauthorized use of services.

Auditing is also a continuous process. The service-provider representation can keep track of data pertaining to a service provided to a user's representation, registering, for example, which services have been used and for how long. This information, called audit data, is saved in the database's audit file. The service-provider decides which data the audit file is to record. Telephone units used to determine the cost of a call is one example of audit data.

In the accounting process, the service-provider charges the user for (trans)actions. Say the user has to pay for a service. The service-provider charges for use on the basis of audit data. Accounting generally takes place after the service has been used. However, accounting can also take place while a service is being used. The information system can, for instance, undertake direct action once the audit process sets off an alarm. An example is when a person trying to make an electronic payment types in the wrong PIN representation several times and the system cuts off the transaction or even "swallows" the card.

2.3.3 Need for identification within the information system

In the conventional information system, the user's identity is often needed to perform the processes outlined in the preceding section. Identity is used within the authorization process, for instance, to identify and record a user's privileges and duties. The user's identity is thus introduced into the information system. Since all of the various elements of the information system are involved in the five processes (in conventional information systems), the user's identity travels throughout the information system. Figure 2.3 illustrates which elements are involved in the various processes.

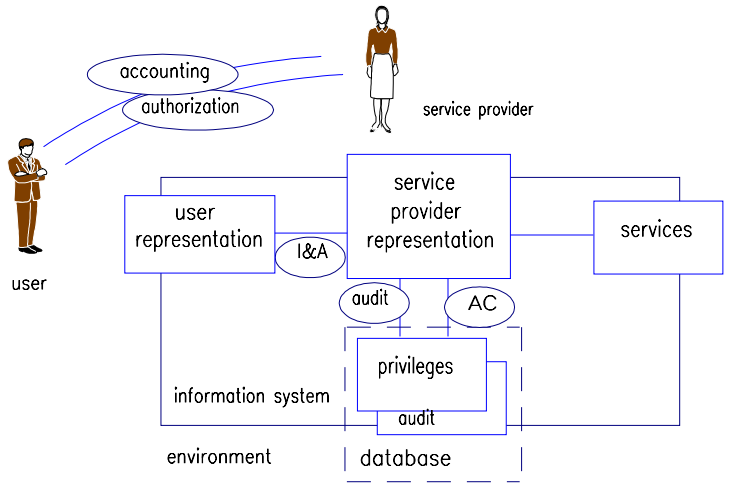

Figure 2.3: The relationship between processes and elements. (I&A: identification and authentication, AC: access control)

For each of the mentioned processes, the question can be asked whether the user's identity is really required.

Is identity necessary for authorization?

In the authorization process, the service-provider assigns privileges to a (future) user. Whether identity is required for authorization depends on the manner in which the service-provider determines the user's privileges. If the service-provider wants to assign privileges on the basis of individual characteristics, then the user is required to demonstrate those characteristics. If privileges are given on the basis of a group characteristic, demonstrating this one characteristic suffices. A few characteristics include:

1. The user (known to the service-provider by a pseudo-identity) begins with limited privileges and accrues more over the course of time (depending on his behaviour). Take the no-claims bonus system for automobile insurance, for example. For each year the driver does not submit any insurance claims, he receives a discount on his premium. The accrual of no-claim benefits is comparable with the accrual of rights.

2. The user receives privileges through being a member of a group. The user must be able to demonstrate that he belongs to the group, club or association. Hotels guest gain access to hotel facilities like swimming pools, weight-rooms and parking places when they show their key.

3. Someone or something serves as a guarantor (trusted third party). Based on pledges made by this trusted third party, the service-provider can grant privileges on the basis of specific (individual) characteristics. One example is parking permits for the handicapped — a hospital can state that the patient, known by a pseudo-identity, does in fact have a handicap.

4. Privileges based on those obtained elsewhere, for example, transfer of privileges from another pseudo-identity. Employees can register for their employer's pension fund under a pseudo-identity. If the employee switches employers, the employee's rights — in the form of the premium paid — can be carried over to the new pension fund. The employee can then adopt a new pseudo-identity for these pension rights.

5. Privileges based on personal characteristics, for instance, age. All people 65 or older can travel for half price. The local authorities can issue a statement to this effect.

It is possible, however, that another information system must be used in order to verify certain characteristics required by a privacy information system. If this information system is a conventional one, i.e. one which

uses the user's identity, the identity of the user will in effect be known to the privacy information system as well. A case in point is when a person requesting a visa has to show his passport as proof of nationality. Conclusion: In most cases, it is not necessary to know the user's identity in order to grant privileges. However, there are some situations in which the user must reveal his identity to allow verification of certain required characteristics.

Is identity necessary for identification and authentication?

In many cases, the authorized user receives an internal representation he will go by when using the information system. The user can then identify himself with his internal representation. Depending on the choice of internal representation and how well-known the representation is, the user's identity may or may not be known. By constantly changing the user representation, it becomes more difficult to link representation and user.

Conclusion: The user's identity is not necessary for identification and authentication.

Is identity necessary for access control?

The access control process checks whether the user representation authorizes the user to perform certain activities. This process takes place within the information system. The internal representation of the user can be used as a reference in lieu of the user's identity.

Conclusion: The user's identity is not necessary for access control.

Is identity necessary for the auditing process?

Internal representation of the user also suffices for the auditing process. After all, it is only necessary to record what a (random) user representation does, so the user's identity is superfluous.

Conclusion: Identity is not necessary for auditing.

Is identity necessary for accounting?

It may be necessary to know a user's identity when he has to be billed for the use of the information system. This can be the case, for instance, if the user misuses or improperly uses the information system and must personally account for it. However, as long as the user follows the rules, his identity need not be revealed.

Conclusion: Identity is necessary for accounting in certain cases.

On the basis of the above analyses, the conclusion can be drawn that it may be necessary, in certain cases, to know the user's identity for accounting and authorization. The necessity depends on the relationships that exist between the privacy information system and the environment. This situation arises if, in the environment of the privacy information system, a conventional information system requests the user's identity. For the processes of identification and authentication, access control and auditing, which take place within the information system, knowledge of the user's identity is unnecessary. Figure 2.4 indicates which processes involve the use of identity, both in conventional and privacy information systems.

Processes	The use of identifiable data in a conventional system	The use of identifiable data in a privacy system
authorization	yes	sometimes(1)
identification & authentication	yes	no
access control	yes	no
audit	yes	no
accounting	yes	sometimes(2)

Figure 2.4: The use of identity in conventional and privacy information systems. (1) In certain cases a consumer must appeal to conventional information systems, which uses the user's identity. (2) In certain cases the user must personally account for it.

3.0 Identity Domains

This chapter illustrates how privacy techniques can be used to separate the user's identity from the use of the information system. A number of these techniques are given in the literature [27]. Based on the model of the information system presented in Chapter 3, a description will be given of how the information systems can be structured in order to better protect the privacy of the user. Section 3.1 will introduce a new system element designed for this purpose: the identity protector. The technical set-up of this identity protector depends on the specific information system. Appendices A to D describe a number of concrete applications.

3.1 The identity protector

The identity protector can be seen as a system element that controls the exchange of the identity between the various system elements. The identity protector is installed, quite logically, on one of the interaction lines in the information system. This means the user's identity can no longer be spread to the cordoned off area of the information system. The role of the identity protector is comparable to that of the service-provider representation in the information system: whereas this protects the interests of the service-providing organization by e.g. monitoring access of users to the services, the identity protector protects the interests of the user — specifically, it screens dissemination of his identity. Just as the service-provider wishes to protect his services, the user wishes to protect his identity.

An important functionality of the identity protector is conversion of a user's identity into a pseudo-identity. The pseudo-identity is an alternate (digital) identity that the user may adopt when using the system. Examples of pseudo-identities in conventional information systems include account numbers at banks and social security numbers for the tax authorities. In the conventional and future information systems, the identity protector may take the form of, say, a separate functionality within the information system, a separate information system controlled by the user (e.g. smart-card), or another information system that is under the supervision of a third party trusted by the service-provider and the user.

The identity protector offers the following functions:

- reports and controls instances when identity is revealed
- generates pseudo-identities
- translates pseudo-identities into identities and vice versa
- converts pseudo-identities into other pseudo-identities
- combats misuse.

The user can set the identity protector for certain purposes, for instance so that his identity is kept entirely confidential when the system is used legitimately. Another possibility is for the user to set the identity protector to reveal his identity only to certain service-providers.

Integration of an identity protector creates two domains within the information system: one in which the user's identity is known or accessible, and one or more in which it is not. The term "identity domain" denotes the domain in which the user's identity is known, the domains in which the user's identity is secret are termed "pseudo domains," see Figure 3.1.

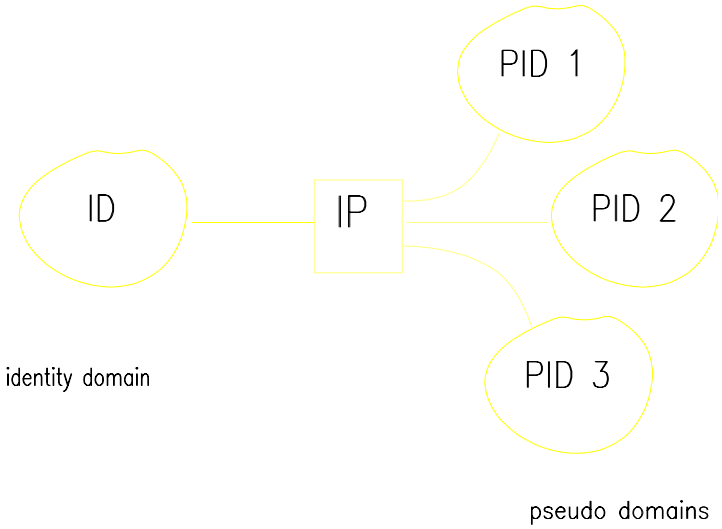

Figure 3.1: The identity protector separates the identity and pseudo-identity domains.

The user must be able to trust the way his personal data is handled in the domain where his identity is known. The identity protector can be placed anywhere in the system where personal data is exchanged. A simple guideline for the designer of a new information system is: minimize the identity domain. Depending on the elements within the information system that can be trusted (in terms of privacy protection), a number of configurations of a privacy information system can be distinguished. The

3.2 Cordoning off areas of services and other users

following section describes a number of these configurations in which the user's identity is unlinked from parts of the information system.

The services element of an information system can be structured in such as a way that the privacy of the user is not adequately protected. By placing identity protectors between the services and the other elements of the information system, privacy protection can be improved. This means services are located in the pseudo domain, while other elements remain in the identity domain (see Figure 3.2).

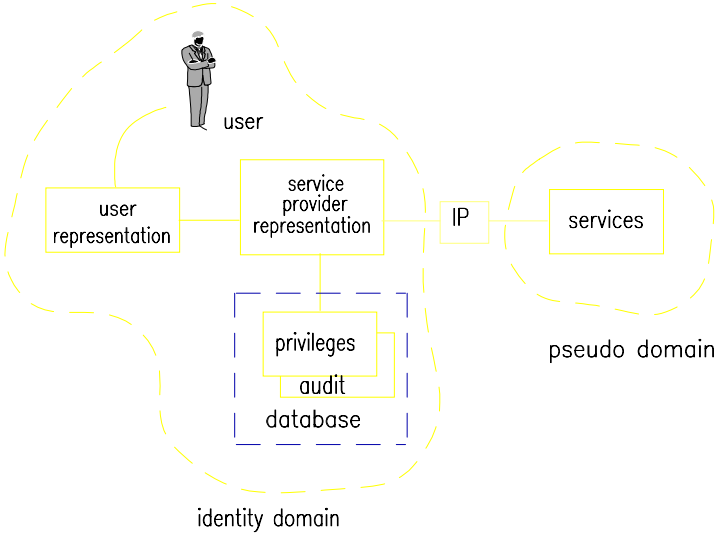

Figure 3.2: An identity protector protects the privacy of a service user

When an identity protector is integrated into a system, the user can use services anonymously, not only increasing privacy in terms of that particular service, but in relation to other users. This last aspect is especially relevant to communication services. Several users can generally use the communication services offered. A communication system

such as a data network is an information system intended for use by many people. In many cases involving an information system with multiple users, the identity of users can easily be kept confidential from fellow users. A precondition is that service-providers take measures, that they furnish the information system with an identity protector, or the functions corresponding with one.

The following two examples illustrate the point. The first example is a direct extension of the communication system and illustrates a situation in which both services and other users are cordoned off. In the second example, only a service is cordoned off.

Example 1. In the regular telephone network, a caller is anonymous to the person receiving the call. The person on the receiving end cannot identify the caller by a telephone number on a display, or the like. The digital telephone networks of the future will enable the receiving telephone to display the number of the person calling. With the help of suitable peripheral equipment, the displayed telephone number can also be saved and used in conjunction with all available data files [37]. The function allowing the caller's number to be displayed is termed "calling line identification." This function offers the caller a number of possibilities for blocking his number so it is not revealed: the calling line number is not sent to the receiving line [36]. Appendix A provides further information on Calling Line Identification.

Example 2. Sometimes users do not have direct access to an (international) network, such as Internet, but need an intermediary information system to gain access to the system and its services. In the case of Internet, this is done via an Internet server. This kind of information not only acts as an intermediary, it can also act as a representative of the user: the users are given a temporary pseudo-identity with which they can use the services the network offers.

3.3 Protection of registration in the database

A service-provider's database consists of a privileges file and an audit file. The privileges file contains the users' privileges and the audit file contains all the other information the service-provider has recorded for provision of his services. Since these two files may register personal data, this system element merits the special attention of the privacy-conscious designer.

The identity protector makes it easy for the designer to minimize the personal data filed in the database. In effect, the service-provider does not register the user's privileges and/or actions under his real identity, but under a pseudo-identity. Figure 3.3 presents a situation in which both the privileges file and the audit file are included in the pseudo domain. It is also possible to cordon off one of the two files.

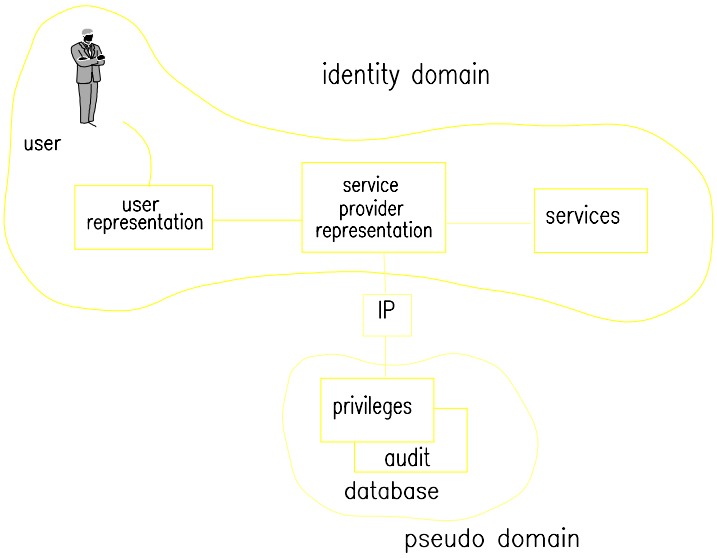

Figure 3.3: An identity protector prevents the registration of the user's real identity in the data-bases (the privileges and the audit file).

In this example, a pseudo domain is included in the audit file.

Example: A large business starts using a call-center, a telephone exchange linked to a computer system, which directs internal and external telephone and data traffic. The telephone numbers of all the calling and receiving lines and the duration of calls are registered for all outgoing telephone calls and external data services (for internal charging and capacity and waiting time statistics). Not the name of the caller or employee making an outgoing telephone call is recorded, but a code that changes daily. This daily representation is generated by a reliable network function: the identity protector. This does not detract from the possibilities of making statistical calculations of capacity and waiting times. Costs can be charged internally because the system keeps records of the cumulative data per department.

3.4 Cordoning off the entire information system

By placing the identity protector between the user representation and that of the service-provider, a pseudo domain emerges which envelops the services, service-provider's database, and the service-provider representation itself (see Figure 3.4).

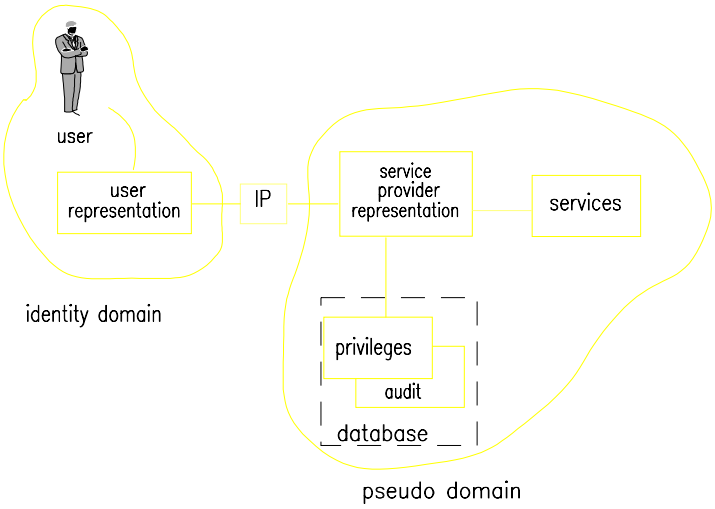

Figure 3.4: Cordoning off the entire information system.

In this situation, the identity domain only contains the user representation. This is also the only part of the information system that the user must trust. Less stringent privacy protection requirements can be set for the other system elements in the pseudo domain. When installing an identity protector, it is important that the way in which communication between the user representation and the service-provider representation be clearly defined and sufficiently secured against intrusion from third parties. User trust in the user representation can be won if the service-provider takes very stringent security measures, or if users have access to and control over a user representation that they can set themselves. This can be a portable computer or a smart-card.

An important aspect of this configuration is that the service-provider must be able to determine what the user is authorized to do, without learning the user's identity. There are various different possibilities for authorizing the user. Section 2.3.3 describes a few situations.

Within the configuration, the identity protector acts as a sort of intermediary for the processes both the user and service-provider go through. So both parties must be able to trust the identity protector. Techniques that are suitable for use with a trusted third party (what could be called a digital attorney) are also suitable for an identity protector in this situation.

Example 1: A new employee of a large organization must be given access to the corporate network. The systems manager has to set up a directory and the authorizations in accordance with the employee's access profile, which is strictly confidential. The access profile is drawn up by the head of the department on the basis of the required access level. The profile, not containing data that can be associated with the new employee, is sent to the systems manager, who checks the profile for authenticity, implements the authorizations and then returns the request form to the department head. The systems manager has added a user ID number and password to the form. The new employee now has access to the network without the systems manager knowing who the employee is. If the employee does something he is unauthorized to do, he can be identified through the department head. It is important that both the systems manager and the employee trust the department head.

Example 2: Membership to a different organization gives a person access to certain benefits. For example, membership to a staff association entitles one to buy goods at a considerable discount.

3.5 Situations with several service-providers

In many cases, several service-providers are involved in the provision of services: it is only possible to pay with a bank card, for instance, if the bank and shopkeeper work together and construct their information systems to accommodate it. Situations involving several service-providers can be complex, and adding an identity protector to a common or linked information system can create specific problems.

A common situation is when two service-providers, let us say A and B, both provide a service to a user, whereby service-provider A supplies a primary service and B a secondary service. Take the bank card example: the shopkeeper supplies a primary service and the bank a secondary service. In this case, the user's privileges are recorded at the secondary service-provider. Figure 3.5 presents a diagram of this situation.

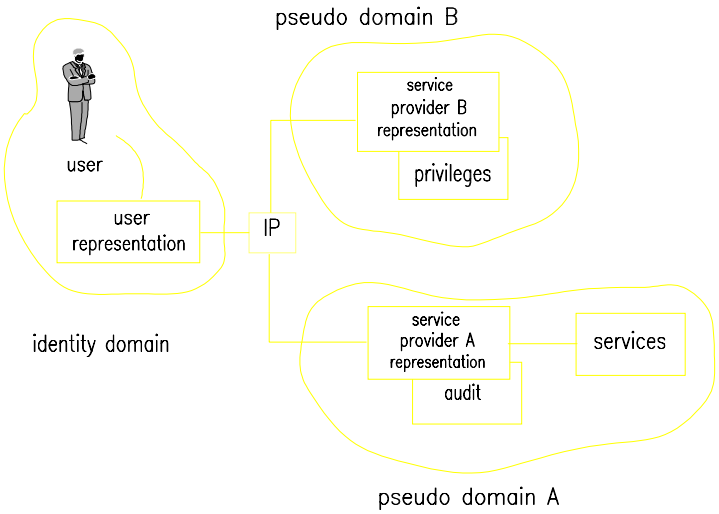

Figure 3.5: Two service providers in different pseudo domains

Service-provider A verifies the user's privileges at/through service-provider B. The identity protector can be installed as two separate functions: one function for each separate service-provider, or as one function for both. This function can be integrated into a smart-card, for instance, that the user carries in his pocket.

It is even possible to integrate a service-provider with the user's representation. Service-provider B can mark an electronic document and give it to the user. Then service-provider A can determine what the user's privileges are by verifying service-provider B's mark on the electronic document. Figure 3.6 shows this situation. In this situation, too, service-provider A determines the user's privileges by checking with service-provider B.

An information system arranged in such a way that the user carries his privileges with him is comparable to an ambassador carrying a Letter of Credence. In the literature, privileges granted in this manner are termed "credentials" [18, 27]. Credentials can be compared with certificates issued by one agency and valid when presented to other agencies. The term "credentials" will be used throughout the rest of this report to denote privileges that the user carries on his person.

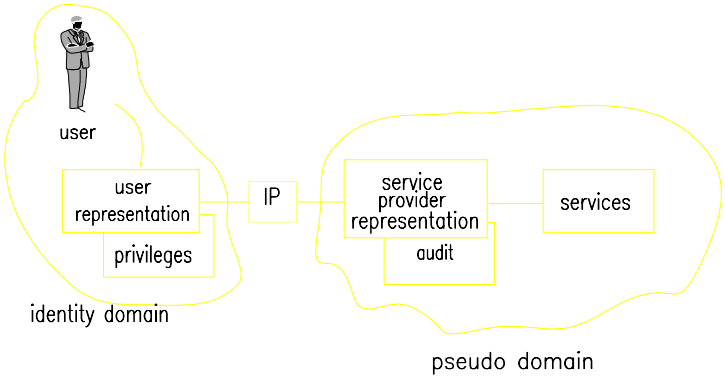

Figure 3.6: The user carries his privileges with him.

The following examples demonstrate that the service-provider does not need to know a user's identity in order to provide services. The first two examples illustrate anonymous payment. The third example describes an interaction between a hospital and an insurance company. When the patient comes in for a certain treatment, his privileges (e.g. insurance policy) are checked without the patient's name being revealed to either the hospital or insurance company.

Example 1. Payment transaction, in which the account number serves as the pseudo-identity and a trusted third party is the only one, besides the user himself, who knows the relationship between the account numbers and the identity of the account-holder. In this case, the identity of the account-holder corresponds with his name, address and town of residence. The trusted third party must also send mail for the bank, after all, the bank does not have any addresses.

The trusted third party could be an independent agency or a part or department of the bank itself: in that case, the service-provider enters a small part of the identity domain, i.e. that part in which mail is sent.

Example 2. Users have an electronic wallet, provided at no charge by the bank, with digital cash. Users can deposit a maximum amount of money in the electronic wallet, for instance by depositing real cash. The digital cash is actually a number representing an amount, which is sealed with a bank identification mark. The shop-keeper also has a digital wallet. The bank can transfer the digital cash from one wallet to another by calculating the new total amounts and sealing these with the bank's digital mark.

Example 3. The service-provider, say a hospital or doctor, wants to check whether a patient is insured for a particular treatment. The hospital and the insurance company know the patient by different pseudo-identities. Via the identity protector, which can translate pseudo-identities, the hospital can determine what coverage the patient has for which treatments.

3.6 Fraud prevention

The identity protector should also prevent fraud or improper use by the user. This can take various forms, such as prevention, detection and correction. One possibility is for the identity protector to prevent the user from being able to use his anonymity to commit fraud. Another approach is based on a combination of detection and correction. The identity protector can determine which measures can be taken "against" the user, such as revealing his identity to the service-provider involved or to the authorities (e.g. police). The set-up of the identity protector should make it possible to also inform the user that his identity is to be revealed.

Examples of preventive methods to keep people from taking improper advantage of their anonymity include hospital insurance cards and (digital) cash. Authentication through entrance representations or biometric data (e.g. fingerprints) renders it impossible for someone else to use a health insurance card. Paper bills are generally made difficult to counterfeit through the use of water-marks and special types of paper and ink in the production process. The same principles hold for digital cash. Cryptographic techniques can be used to prevent one sum of cash from being spent anonymously more than once.

In this example, an identity protector detects a user trying to take unfair advantage of his anonymity and corrects the user: a user receives access to a certain service through the mediation of a go-between (such as a "digital" attorney) which acts as an identity protector. The service-provider wants to charge the user for the service provided and sends the bill to the intermediary, who, in turn, sends the bill to the user. If the user does not pay, the service-provider will eventually ask the intermediary for payment again. There are now several ways in which the intermediary can approach the user. He can use cryptographic techniques to reveal the user's identity to the service-provider. The service-provider can then contact the user directly or through a collection agency. Another option is for the intermediary to seek contact directly or through a collection agency in order to secure the user's payment. However, the user should always be given the chance to prove he has been falsely accused of misconduct before his identity is revealed. Maybe the user never received the first bill at all.

4.0 Implementation Techniques

Are the models we presented in the previous chapter feasible? This chapter begins with an explanation of some specific techniques for integrating an identity protector into a system and concludes with some guidelines for the development of privacy-protecting information systems.

4.1 Setting up an identity protector

So far this report has presented the identity protector as an abstract functionality, or black box as it were, which places the designer in a position to construct the information system so that the user's identity is cordoned off and only revealed in certain situations. The designer is not limited in his choice of special techniques for the creation and implementation of the identity protector. Some techniques, such as digital signatures and trusted third parties, merit special attention [38, 39].

4.1.1 Digital signatures

A signature or wax seal on a document is proof of its authenticity. A digital signature is an electronic version of a hand-written signature. The key aspects of both types of signatures are that only one person or service-provider is capable of producing the signature, and all others are capable of verifying it (see Figure 4.1).

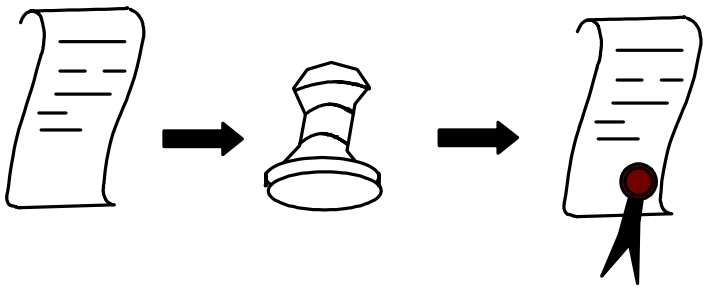

Figure 4.1: A digital signature corresponds with a written signature or a wax seal. A signature on a document is proof of its authenticity.

How is a digital signature made? In most cases, digital signatures are created by means of an irreversible process within the electronic document that calculates a digital value. This value is called the hash or "compaction" value ("to hash" means to chop fine). The purpose of the hash value is to convert a random electronic document into a digital value of a fixed length (in bits). This simplifies the application of cryptographic

techniques, used to encipher the hash value into numbers. The result is a digital signature, which can be distributed together with the electronic document.

The signature, i.e. proof of a document's authenticity, can be validated as follows. The sender and recipient make agreements concerning the enciphering method and irreversible process, which enables the recipient to calculate the document's hash value. The digital signature is deciphered cryptographically. The recipient now has two values he can compare. If the values match, the file received is authentic, if they differ, the file has been altered in transit. This could be due to tampering or because of a transmission error.

Everyone who has an agreement with the person compiling the document (sender) can verify that the electronic document is authentic by checking the corresponding signature. Digital signatures are only valid for the electronic document for which they were created. Each electronic document has its own (unique) digital signature.

A potential application of digital signatures is digital driver's licenses. The Dutch Central Division of Motor Vehicles (CBR) could attach a digital signature to an electronic document which holds the class of the permit. Other organizations like car rental companies and the police can then check the driver's credentials by screening the digital signature on the electronic driver's license.

4.1.2 Blind digital signature

A *blind* digital signature is a special kind of digital signature [18]. The difference does not lie in the signature itself, but in the document to which it is attached. When a person places a regular digital signature on a document, he is familiar with the contents of that document. A person placing a blind digital signature, on the other hand, has no or only partial knowledge of the document's contents. The signer often has a certain authority or represents a certain agency, such as a notary, and is not accountable for the document's contents.

A blind signature works like this: a user brings a document to a notary. The user does not want anyone, including the notary, to know the contents of the document. The user seals the document in an envelope. A portion of the document is visible through the envelope. The notary places a wax seal on the visible portion. The seal is proof of the document's authenticity. When a blind digital signature is used, cryptographic techniques replace the envelope and wax seal. The user enciphers the digital document, which is comparable to putting the document in an envelope. The notary places a digital signature on the

document in the envelope (see 4.1.1). When the document must be checked for authenticity, the signature is validated.

The document can be represented as an electronic letter and envelope (see Figure 4.2). Figure 4.2 schematically illustrates the cryptographic process.

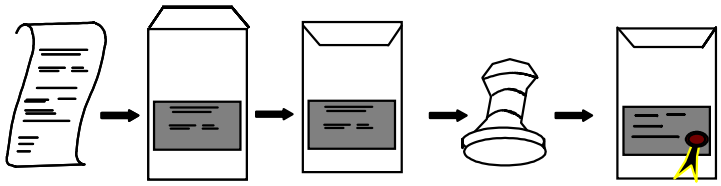

Figure 4.2: A blind digital signature: The digital envelope protects the contents of the digital letter. The digital signature on the letter is proof of its authenticity.

An application involving blind digital signatures is "digital cash" [27]. A user takes an envelope to the bank. The envelope states the user's account number and contains a piece of carbon paper and a bill. The user asks the bank to assign a value of 10 dollars to the bill. The bank places an official stamp on the envelope to give it the value of 10 dollars (blind digital signature). The bank uses a different stamp for every value. The stamp is copied onto the bill through the carbon paper. Now the user can remove the bill from the envelope and he has a 10-dollar bill. The bank cannot link the bill to the user's account number and thus to his identity. When the user spends the bill, neither the bank nor the service-provider receiving the bill as payment can draw a connection between the bill and the user. The service-provider can tell from the stamp whether the bill is real.

4.1.3 Digital pseudonym

A digital pseudonym can be represented by a completely random selection of characters (letters, numbers and punctuation marks). The user is not known to a service-provider by his identity (name, address, city), but by this series of characters. He can select a different pseudonym for every service-provider. Consequently, service-providers cannot exchange information about individual users. A different pseudonym can also be used for each service or individual time a service is used.

If there are n service-providers, the user chooses n pseudonyms: PID-*1*, PID-*2*, up to PID-n. The "ith" service-provider knows the user by the pseudonym PID-i. The service-provider assigns privileges to this pseudonym by furnishing a blind digital signature. The user keeps the assigned privileges and can use these privileges with other service-providers under a different pseudonym.

Users have a special "envelope" with a transparent window for each service-provider, which enables them to communicate with service-providers. The user — or a third party in whom he trusts — collates all these pseudonyms in one digital letter. Service-providers can give users new privileges by adding blind digital signatures, a signature corresponds with a specific privilege. The user can present other service-providers with proof that he is (properly) authorized (see Figure 4.3).

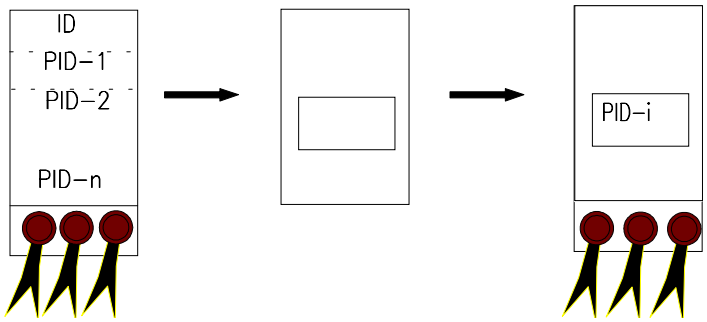

Figure 4.3: Digital pseudonyms offers a user the possibility to present proof of his privileges under different pseudo-identities. The user has for this purpose a number of digital envelopes with a transparent window.

The user can also use obtain services from service-providers without a pseudo domain provided he reveals his identity. The user then presents proof of his identity and the digital signatures he has obtained.

Digital pseudonyms can also be used for the digital driver's license mentioned in Section 4.1.1 above. Here, the Central Division of Motor Vehicles (CBR) has given the driver a blind signature which corresponds with a pseudonym. The CBR uses a digital signature for each class of license. The driver can use a different pseudonym to prove (e.g. to a car rental company) that he is authorized to drive certain vehicles, by presenting the digital signature(s) he received from the CBR.

4.1.4 Trusted third parties

A trusted third party is a term for a service-provider who is trusted by both users and service-providers (a sort of electronic attorney). The trusted third party can, for instance, keep track of the digital pseudonyms a user uses in his relationships with a number of service-providers (see Figure 4.4).

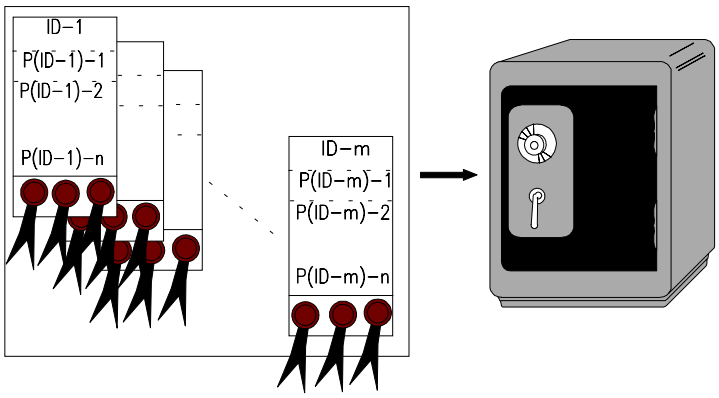

Figure 4.4: A trusted third party can keep track of the digital pseudonyms a user uses in his relation with a number of service-providers.

The user's trust is founded on the discretion the trusted third party observes with respect to the user's identity: the trusted third party must keep the relationship between the identity and pseudo-identities secret. The service-provider's trust, on the other hand, is based on the assumption that — if conditions require — the trusted third party will reveal the user's identity. A service-provider may need the identity of a user in order to hold the user accountable for wrongful or improper use. After the user has accounted for his actions, he can initiate a new relationship, under a different pseudo-identity, with the service-provider.

In the above example of digital driver's licenses, a trusted third party can register and keep track of the relationship between the driver's identity and the pseudo-identities stated on his license. In certain cases, a driver's identity can still be determined on the basis of his pseudo-identities. These powers should be reserved for organizations like law enforcement agencies.

4.2 From conventional to privacy information system

When a new information system is being engineered, or a conventional information system is being upgraded, the client, designer, developer or supplier of an information system can ask himself how the user's privacy can be better protected.

In the analysis phase, the question should be asked of how much and which personal data is in fact required for the information system to function properly. An attempt must be made to minimize the amount of

information, particularly identifying data, filed by an information system. Minimization of data has implications for information system processes of input and output and the ways in which a system records information.

The position of the identity protector — or an equivalent functionality — within the information system is a crucial part of the design phase. A decision has to be made about which elements are to belong to the pseudo domain and which to the identity domain. This is also the phase in which to determine how the user is to exert control over release of his personal data. This is a matter of how the identity protector is to be set up. What are the identity protector's functions to be?

Questions concerning specific techniques for creating the identity protector arise in the implementation phase. The issue of concern is that the information system must not allow data to circumvent the identity protector and thus leak from the identity domain into the pseudo domain. Special attention must be paid to what could be unique serial or production numbers generated "automatically" by hard- and software.

Figure 4.5 indicates how the designer can take the user's privacy into account during the different phases of the design process.

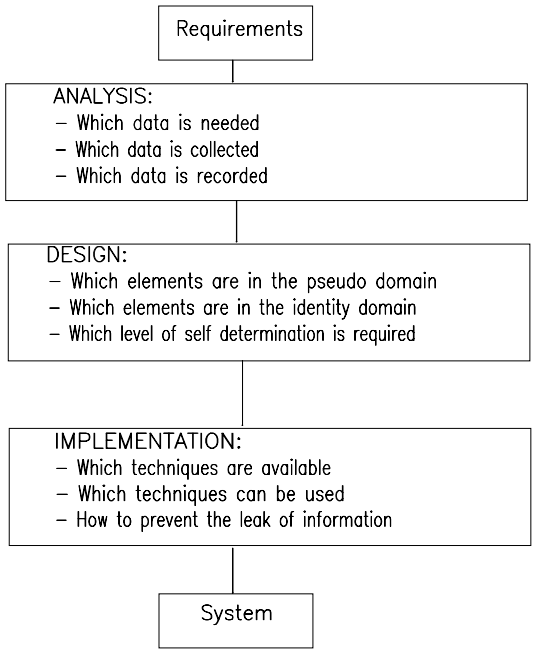

Figure 4.5: Aspects to take into account during the different phases of the design process of a privacy information system.

Literature

[1] Anonymous one-time signatures and flexible untraceable electronic cash. B. Hayes. Advances in cryptology — AUSCRYPT '90, January 1990 p. 294–305.

[2] Blind signatures and perfect crimes, On. S. von Solms and D. Naccache. Computers & security, 11 (1992) p. 581–583.

[3] Caller display and call return. W. Dangerfield, S. Garrett and M. Bond. British Telecommunications Engineering, Vol. 12, October 1993 p. 176–182.

[4] Chip-card als blackbox van de burger. Y. Cramer. Technisch weekblad, 21 September 1994 p. 15.

[5] Communication protocols with untraceability of sender and receiver. M. Mambo, H. Kinoshita and S. Tsujii. Systems and computers in Japan, Vol. 23, no. 9, 1992 p. 11–18.

[6] Demonstrating possession without revealing factors and its application. H. Shizuya, K. Koyama and T. Itoh. Advances in cryptology — AUSCRYPT '90, January 1990 p. 273–293.

[7] Development of a road pricing system in the Netherlands, the. H.J. Stoelhorst and A.J. Zandbergen. Traffic Engineering + Control, February 1990 p. 66–71.

[8] Distributed consensus mechanisms for self-interested heterogeneous agents. E. Ephrati and J.S. Rosenschein. 1993.

[9] Electronic voting scheme allowing open objection to the tally. K. Sako. IEICE Trans. Fundamentals, Vol. E77–A, no. 1 January 1994 p. 24–30.

[10] Improved privacy in wallets with observers. R.J.F. Cramer and T.P. Pedersen. 1994 p. 329–343.

[11] Oz card is dead! — long live the oz card! the. Computer fraud & security bulletin, Vol. 10, No.2 p. 13–14.

[12] Peeping Sam: Uncle is watching us. G.B. Trubow. Computer security journal, Vol. IV, number 1 p. 15–20.

[13] Privacy, anonymity and interpersonal competition issues identified during participatory design of project management groupeware. M.J. Muller, J.G. Smith, J.Zachary Shoher and H. Goldberg. SIGCHI Bulletin, Vol.23 number 1, January 1991 p. 82–87.

[14] Privacy, Cryptographic Pseudonyms, and the state of health. S.Fr. Mjolsnes. Advances in cryptology — ASIACRYPT '91. November 11–14, 1991 p. 493–494.

[15] Privacy without authentication. J.M. Galvin. Message Handling Systems and distributed applications, 1989 p. 187–202.

[16] Providing Location information in a ubiquitous computing environment. M. Spreitzer and M. Theimer. SIGOPS '93, 12–1993, N.C., USA p. 270–283.

[17] Security features in the GSM-information system, The. U. Michel. 1991 p. 385–389.

[18] Showing credentials without identification transferring signatures between unconditionally unlinkable pseudonyms. D. Chaum. Advances in cryptology —AUSCRYPT '90, January 1990 p. 246–264.

[19] Sicherheit ohne identifizierung. D. Chaum. Informatik-Spectrum (1987) 10: p. 262–277.

[20] Three-pass identification protocol using coding theory, A (non-practical). M. Girault. Advances in cryptology — AUSCRYPT '90, January 1990 p. 265–272.

[21] Untraceable off-line cash in wallet with observers. S. Brands. Advances in cryptology — CRYPTO '93, August 1993, p. 302–318.

[22] Videotex without "big brother". H.A. Maurer, N. Rozsenich and I. Sebestyén. Electronic publishing review, Vol. 4, no. 3, 1984 p. 201–214.

[23] Wallet databases with observers. D. Chaum and T.P. Pedersen. 1994 p. 89–105.

[24] Optimierte computergestützte zufallszahlengenerierung zur anonymisierung patientenbezorgener Informationen. A.J.W. Goldschmidt und L. Ga l. Software Kurier für Mediziner und Psychologen 1991 nr. 4 p. 145–150.

[25] Anonymous and verifiable registration in databases. J. Brandt, I.B. Damgard and P. Landrock, Advances in Cryptology — EUROCRYPT 88, pp 167–176.

[26] Untraceable electronic cash. D. Chaum, A. Fiat and M. Naor, Advances inCryptology, CRYPTO 88, pp 319–327.

[27] Achieving electronic privacy, a cryptographic invention known as a blind signature permits numbers to serve as electronic cash or to replace conventional identification. The author hopes it may return control of personal information to the individual. D. Chaum. Scientific American August 1992, page 96–101

[28] Dining cryptographers problem, the: unconditional sender and recipient untraceability. D. Chaum. Journal of cryptology 1988 nr. 1 p. 65–75.

[29] Networks without user observability. A. Pfitzmann and M. Waidner. Computer and security nr. 6 1987 p. 158–166.

[30] Security without identification: card computers to make big brother obsolete. D. Chaum.

[31] Privacy Protection, M. Rotenberg, Government Information Quarterly, Volume 11, number 3, 1994.

[32] Privacy Protection and the Increasing Vulnerability of the Public, P. Hernon, Government Information Quarterly, Volume 11, number 3, 1994.

[33] Privacy Protection and Data Dissemination at the Cencus Bureau, H.A. Scarr, Government Information Quarterly, Volume 11, number 3, 1994.

[34] Pseudo-identities in health registers?, E. Boe, The International Privacy Bulletin, Volume 2, Number 3, July 1994.

[35] Beveiliging van persoonsregistraties, Registratiekamer, December 1994.

[36] Euro-ISDN, User-networkaspect, PTT Telecom, 1993.

[37] Nummeridentificatie bij telefoonverkeer, Registratiekamer nr. 93.A.012, 25 February 1994

[38] New Directions in Cryptography, W. Diffie and M.E. Hellman, IEEE Transactions on Information Theory pp. 644–654, vol. 22, no. 6, November 1976.

[39] Security Without Identification: Transaction Systems to Make Big Brother Obsolete, D. Chaum, Communications of the ACM, pp 1030–1044, vol. 24, no. 2, February 1981.

[40] Smart cards, T. Wright, Canada.

[41] Rothfeder, Privacy for Sales: how computerization has made everyone's private life an open secret,

[42] Bemelmans, Bestuurlijke informatiesystemen en automatisering, Stenfort Kroese, 1987.

[43] Actieplan elektronische snelwegen, 'Van metafoor naar actie'. Minsterie van Economische Zaken, 1994.

Appendix A: Automatic Number Identification (ANI)

A digital telephone network enables the receiver to identify the caller via the telephone number the network communicates to the telephone or other peripheral equipment of the receiver. This number can then be directly displayed or used as a search key within a database so that data pertaining to the caller is displayed directly. The function enabling the caller's number to be conveyed to the receiver is termed Automatic Number Identification (ANI). In ANI, the caller has a say about whether his telephone number is to be revealed. Put in this perspective, ANI offers the functionality of an identity protector. Here, the identity protector is located between the service-provider and the services (see Figure A.1).

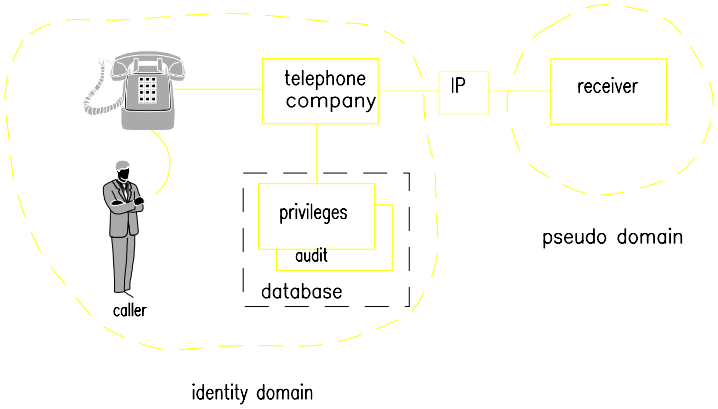

Figure A.1: The functionality of the Automatic Number Identification can be compared to the functionality of an identity protector.

To date, the service-provider in Figure A.1 (the telephone company in this case) still requires the caller's identity in order to charge him for the services provided. This means it is not (yet) possible for the caller to remain anonymous to the service-provider. The person receiving the call in Figure A.1 is another user of the information system who can be approached via the service of "phoning." The caller can keep his identity secret from the receiving party through the use of an identity protector, which consists of a number of blocking options integrated in the functionality of the Automatic Number Identification.

A.1 Blocking options offered by ANI

The caller has the option to block his number so the number of the calling line is not passed on to the receiving line [36]. The different blocking possibilities offered by ANI include:

1. blocking ANI per call
2. total blocking of ANI.

Re 1. Blocking ANI per call

By pressing a code before dialing the receiver's telephone number, the caller's telephone number is not displayed to the receiving party. This code is checked by the identity protector. When the code is typed in, the identity protector does not pass on the telephone number of the caller to the receiver. The identity protector works in this case as a user-controlled filter for identifying information (telephone number).

Re 2. Total blocking of ANI

It is arranged with the telephone company that the telephone number of the caller is never to be given to those on the receiving end. Here, the identity protector works as a pre-set fixed filter for the identifying information: the telephone number.

In addition to the caller's options to block display of his telephone number, there are ways to "protect" the one receiving a call from the caller. After all, the caller could be invading the privacy of the person he is calling. There are two possibilities:

• the receiver can decide that anonymous callers are not to be given access to his peripheral equipment. In this case, the caller does not know whether the receiver is out or just not taking his call.

• certain (governmental) agencies (such as those providing assistance) have the option to overrule the caller's choice to block his number for each call or all calls. This allows the receiver to receive the number of the caller. The caller then receives a signal that, in spite of blocking, the receiver has been informed of the number.

A.2 Conclusion

Automatic Number Identification and concomitant blocking options exemplify the function of the identity protector in a digital telephone network. The most important aspect of privacy protection with respect to ANI is that the caller can decide whether or not his number is to be given to the person receiving a call. The caller does, however, have to take extra action to block his number from being passed on. If the telephone company makes ANI blocking a standard option, on the other hand, the user need not make any extra arrangements to keep his number private. From the perspective of privacy, this is preferable: the caller should be able to turn the standard blocking feature off with the touch of a button. This option should be possible at the time the connection is being made and while the conversation is underway.

In addition, the person receiving calls can guard himself from unidentified callers by refusing to take calls when the number has been blocked. Sometimes, such as when calls are received by police and emergency hotlines, it may be advisable to overrule blocking. Then the receiver will still be able to see the caller's number.

Appendix B: Provision of Medical Data

Every day, data concerning individuals is stored in databases. The registration of medical data is one example. Medical information is not only important and interesting to the treating physician, but to many others like fellow doctors, nursing staff, pharmacists, insurance companies, scientific researchers, and employers. Databases where this information is filed do not usually have features to protect privacy, meaning that anyone who has access to these databases has access to all data on this individual [35].

Not all involved parties need know the patient's identity. Scientists conducting research into certain illnesses/trends, for example, do not need to know the identity of the person. What is important to them is that they have access to all the data relevant to a study. Not only the illnesses and treatments that a patient has gone through are of interest, but also certain habits, like smoking, exercise, etc. So far, scientists have used patients' identities in order to collate all of the registered information.

B.1 System description

There are a number of methods for protecting the patient's privacy when medical data is stored in a database. This appendix focuses on two options: one in which the patient has one pseudo-identity, and one in which the patient has a different pseudo-identity for every involved party.

Each of these methods is discussed separately, while it is assumed that the doctor knows the identity of the patient, but the other parties do not.

B.1.1 One pseudo-identity per patient

The doctor gives each patient a pseudo-identity. The doctor keeps the relationship between the identity and pseudo-identity of the patient secret. The doctor could, for instance, entrust the identity and corresponding pseudo-identity to a trusted third party. The doctor records the medical data on the patient under his pseudo-identity. Other parties can now have access to the database containing medical information without learning the patient's identity.

B.1.2 Multiple pseudo-identities per patient

A second method is based on multiple pseudo-identities per patient. These pseudo-identities can be stored together with the identity in files that are only accessible to the trusted third party. The pseudo-identity of a patient is different for each party 1, 2, ..., n (see Figure B.1).

The doctor can assign the patient certain characteristics by including a digital signature with the patient's identity (ID). Say the patient is administered a certain medicine — the doctor places the signature corresponding with that medicine under the patient's identity. The other parties (i.e. pharmacy, insurance company and researcher) can now determine whether a patient receives that particular medicine by checking for the corresponding signature under the pseudo-identity PID-1, PID-2, ..., PID-*n*.

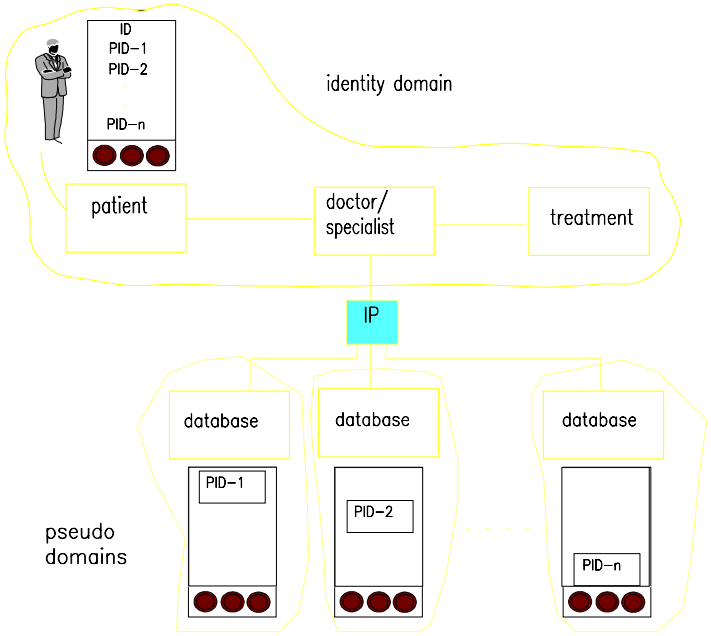

Figure B.1: Multiple pseudo-identities in the database. The different pseudo-identities cannot be associated with each other. So the patient can not be identified without the help of the identity protector.

B.2 Discussion

The first method ensures that organizations have access to all data except the identity. However, all of this information could be used to link the pseudo-identity to the patient's actual identity. There is a chance that a single pseudo-identity can be associated with the patient's identity.

In the second method, the patient uses a different pseudo-identity for each agency. The different pseudo-identities cannot be associated with each other.

B.3 Conclusion

When only one pseudo-identity is used, the risk that the pseudo-identity will be traced to the identity is greater than when multiple pseudo-identities are used. The latter requires sound management of these pseudo-identities when all of the data on a patient is called up. The trusted third party is responsible for this management.

Appendix C: Road-pricing

In the late eighties, the Dutch Ministry of Transport and Public Works considered introducing road-pricing. The purpose of this system was to charge road users for actual road use, as contrasted with customary road tax based on possession of a vehicle. The preferred method for a road-pricing system was one in which road users could pay automatically with a smart-card. It still remains to be seen whether the road-pricing system will ever be implemented [7].

C.1 System description

There are two fundamentally different approaches to road-pricing: the first is a system in which the road user pays *afterwards* and the second in which this occurs *beforehand*. In the literature, these variants are referred to as post-paid and pre-paid systems.

The post-paid system can be simply achieved by requesting the vehicle registration number at the time the vehicle passes a toll point. The registration number is automatically called up. This system offers little or no protection of the road user's privacy — the vehicle registration number is easy to associate with the owner of the vehicle — and will not be discussed in any more detail here.

The other possibility, which is based on the pre-paid model, can be set up as follows. The road user deposits cash on his card — with digital cash — at fixed deposit points along the road, for example gas stations. The deposit points accept cash, which is then added to the value of the card as digital cash. Amounts are deducted from the card at so-called toll collection points. This is completely automatic with the aid of telecommunications. Each vehicle is furnished with what is called a transponder. The smart-card can be linked to the transponder, so that the smart-card can communicate with the toll collector [7]. The card and the deposit points are made available by the toll collector. The above system is what is known as a closed system: the digital cash can only be spent at the toll collector's. Appendix D ("Digital Cash") describes an open system whereby the bank issues and accepts digital cash. In such a case, digital cash can be spent everywhere.

Digital cash consists of electronic documents that the toll collector signs with his digital signature. A road user may select the electronic documents himself. Each signed document represents a fixed value which allows the road user to pass a toll collection point. The road user sends a signed document to the toll collector at each toll collection point (see

Figures C.1 and C.2). The value of the signed document does not depend on the content of the document: it is important that the document be signed by the toll collector and no one else.

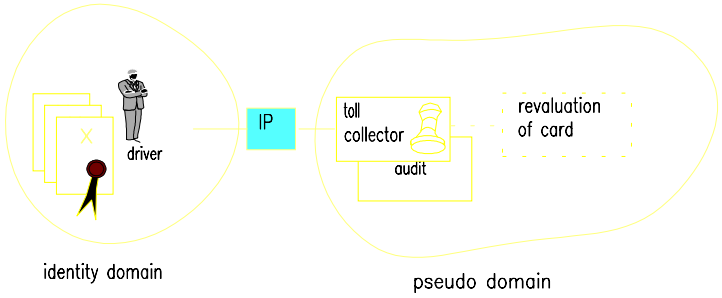

Figure C.1: The toll collector signs with his digital signature the road user's electronic documents. Not the document's content but the toll collector's signature represents a value.

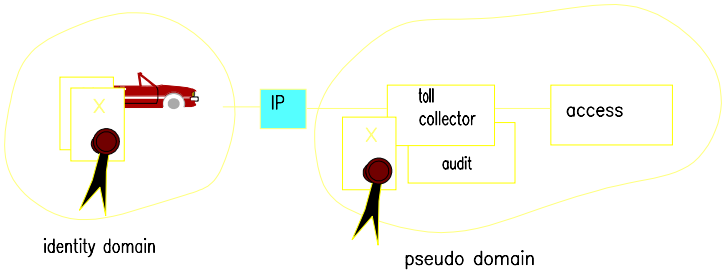

Figure C.2: Each time passing by at a toll collection point the road user sends a signed document to the toll collector. The toll collector's digital signature on the document is proof to its authenticity.

A variation of this system which offers less privacy protection is one in which the user is granted a single pseudo-identity by the toll collector. The road user goes by this pseudo-identity when communicating with the toll collector. However, the privacy of the road user is jeopardized as soon as it becomes possible to link the pseudo-identity with his real identity.

C.2 Summary/Conclusion

The proposal for a road-pricing system [7] does not make any statements about the data to be exchanged between the road user and toll collector. The proposal does indicate that most of the information made available during the various transactions can be deleted by the toll collector afterwards. In that case, the user must be able to trust the toll collector.

A road-pricing system as outlined above would not allow the toll collector to trace the identity of the road user. In that case, the road user need not depend on the toll collector's good will when it comes to protection of his privacy.

Implementation of the system is a whole new ball-game. A well-designed privacy information system can be completely undone if, for example, the transponder can be linked with the vehicle on which it is mounted. If the transponder has a unique identification number (e.g. a factory number stored in the equipment's hardware), then each transponder can be associated with the registration number of the vehicle onto which the transponder is mounted.

There are various possibilities for implementation when designing a road-pricing system. Designers must be aware of the locations in the system where the road user's identity can be tracked down. This part of the system should be minimized.

Appendix D: Digital Cash

Users can pay for articles purchased in a store in a number of ways: with cash, with a bank card, or with a credit card. The last two payment options involve use of data that can easily be linked with the user's identity. The bank statements the shopkeeper receives state highly identifying data, such as the account number and name of the user. If a user wants to remain anonymous, he is currently forced to use the first means of payment — cash.

D.1 System description

There are different ways to improve safeguarding user privacy when making payments. We will discuss three methods: procedural measures taken by the bank, pre-paid cash cards, and transferrable credentials.

D.1.1 Procedural modifications by the bank

The only difference between this solution and customary payments with a bank card is that the shopkeeper does not receive the name and account number on bank statements. In this way, the shopkeeper cannot keep records on users and their spending patterns. The procedural measures at the bank consist of not stating the name/account number of the customer.

D.1.2 Pre-paid cards

Pilot projects are currently underway involving the deposit of cash onto a smart-card. The cards are issued by interested parties, such as a large chain of department stores or a bank. The cards are anonymous: no records contain information enabling the card to be linked with the user's identity. When payment is made, cash is deducted from the smart-card. These cards are also referred to as pre-paid cards. Pre-paid cards could also be used for road-pricing systems (see Appendix C).

Cash or bank card money can be deposited on the card. Machines are required with which money can be deposited on the card. The service-provider has a machine to check whether the user has enough cash on his card. The service-provider can also use this machine to transfer cash from the user's card to his own card, till or account number. When a smart-card is used to deposit cash and make payments, the system must — if it is to protect the user's privacy — make it impossible to draw a link between the account number and the smart-card, which would be possible if the smart-card contained a unique serial or factory number and communicated this to the machine used to deposit cash on the smart-card.

The costs of copying a pre-paid card are not proportionate to the (limited) maximum amount that can be deposited onto it. The card is generally not secured against loss or theft. Someone who comes into possession of a lost or stolen card can use it without problems.

Telephone cards are an example of a pre-paid card. With the card, the user can use the services of the telephone network while remaining anonymous. This card has a certain initial value. Each time a pay-telephone is used, the value is reduced. Cash can never, however, be added to this kind of card. If the card is stolen from or lost by the user, he loses the amount remaining on the card. From the perspective of privacy, it is better to use telephone cards with smaller amounts. A user can buy one telephone card worth twenty-five dollars and create one (big) pseudo domain. Five telephone cards of five dollars each means that he creates five (small) pseudo domains. Five different pseudo domains affords the user more privacy than one large pseudo domain. This example also demonstrates that measures to protect privacy need not entail higher costs.

D.1.3 Payment by credentials

A third way to pay anonymously is based on so-called transferable "credentials." Here, blind digital signatures are used [27, 30]. The bank knows the user's identity, but with this method, the bank cannot find out where the user spends his money. Nor is the shopkeeper able to draw a link between the money and the user's identity.

Figure D.1 shows how a bank places a digital signature on an electronic document, say a bill, belonging to the user. This signature corresponds with a certain amount of cash: the bank uses a different signature for every fixed amount. This sum of money is deducted from the account as soon as the bank signs it. Figure D.2 indicates how the user can pay a shopkeeper under a pseudo-identity. He transfers the digital signature from his identity (ID) to his pseudo-identity (PID).

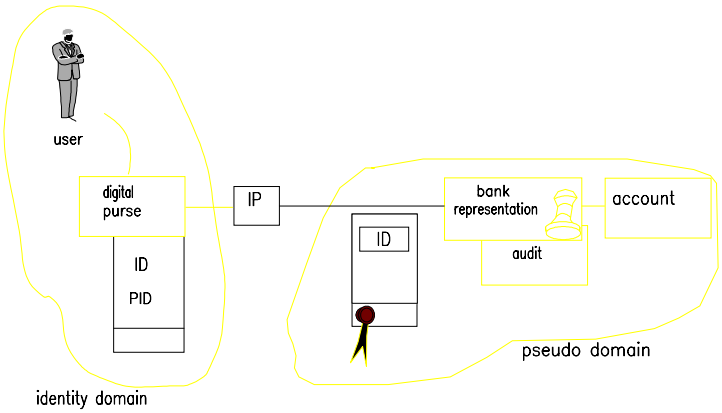

Figure D.1: The user produces a digital document with both his identity and a pseudo-identity on it. Before sending the document to the bank the user covers the pseudo-identity. The bank verifies the document, signs it and debits the user's account. After this transaction the user possesses a document representing a fixed value.

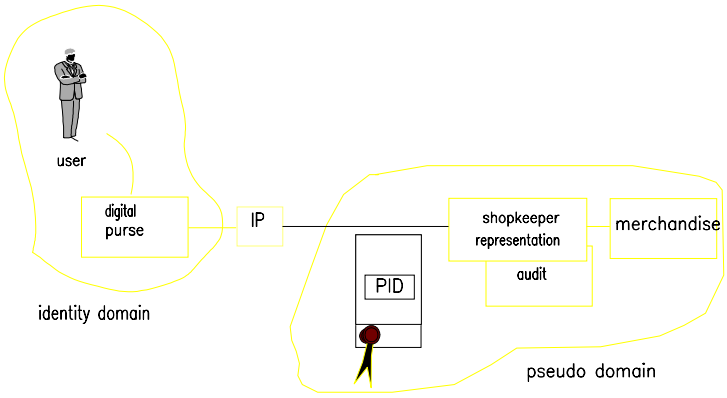

Figure D.2: The user enters a store. Before the signed document is handed to the storekeeper, the user covers his identity. The storekeeper can only read the pseudo-identity and the value. The bank's digital signature on the document is proof of its authenticity. The bank credits the storekeeper's account.

D.2 Conclusion

There are a number of ways to improve protection of the user's privacy when making payments. The options vary from simple procedural adaptations to entirely new systems. When procedural adaptations are made, the shopkeeper no longer receives the names of his customers on his bank statements. In this case, the user is dependent on the bank to protect his privacy. New systems make use of cryptographic techniques such as the digital signature which compel protection of the user's privacy.

Appendix E: Table of Contents Volume I

Appendix F: List of Participants

Registratiekamer, Rijswijk, The Netherlands:

John Borking
Vice President

Huib Gardeniers
Legal Adviser

Henk Van Rossum
Staff Member Technology

TNO Physics and Electronics Laboratory, Telematics and Information Security Group, The Hague, The Netherlands:

Joost Meijers

Paul Overbeek

Paul Verhaar