IJIE
International Journal of Information Ethics Vol. 2 (11/2004)

Thomas B. Hodel-Alma Schütter

Informational Self-Determination Databases in an Intercultural Perspective

Abstract:

An Informational Self-Determination Database System allows to store, manage and query data while at the same time respecting the data subjects' rights of information privacy. We argue that in a world of everincreasing amounts of data that are directly or indirectly related to identifiable individuals and which are being maintained by many organizations, it is of utmost importance to offer strong, effective and reliable concepts and mechanisms – technical, organizational as well as legal – to avoid adverse effects of information processing on people. We present a short motivation for our claims. We then sketch our vision of an Informational Self-Determination Database System and its working. We maintain that our approach offers a realistic, practical and pragmatic solution for enhancing people's privacy, without hindering organizations in getting their business done.

Agenda

ISSN 1614-1687

Authors:

Dr. Thomas B. Hodel:

- University of Zurich, Department of Informatics, Winterthurerstr. 190, CH-8057 Zürich, Switzerland
- ✉ hodel@ifi.unizh.ch, 💻 http://www.ifi.unizh.ch

Alma Schütter

- University of Zurich, Department of Informatics, Winterthurerstr. 190, CH-8057 Zürich, Switzerland
- ✉ alma.schuetter@iew.unizh.ch

Introduction

Information systems and electronic data processing have increasingly become a part of our daily lives. Ever growing amounts of personal data are being stored and processed and the explosive development of privacy-invasive technology such as RFID tags (radio frequency identification), bioimplants or DNA sniffers make informational privacy a growing concern. Although many countries have enacted data protection laws, many people perceive these laws as being inadequate and are concerned about the loss of privacy in the Internet age. Privacy-enhancing technologies have been developed to curb the use of personal data in information systems. However, both technical and legal measures have yet failed to give people control over their personal data. Generally, most people do not know where data about them is stored or how this data is used.

In this paper, we propose a novel approach to build privacy-protecting database systems, so called informational self-determination database system. With our database system we aim to give people better control over their data and heighten transparency in data processing. As a major innovative feature, we propose that data processor and data subject establish a contract before engaging in data processing. This contract clearly specifies for what purposes data may be processed. Through this form of contract, the privacy principle of consent (as stipulated by Alan Westin [15]) can - for the first time - be truly implemented. Furthermore, the proposed system leads to increased transparency, as citizens can view a detailed log file for each data collection that states when and for what purpose their personal data have been accessed. These log files are accessible through an easy to use portal service. This enables the compliance with a major section of the data protection law.

Our approach builds on existing work in the domain of privacy-enhancing technologies. In particular, the approaches made by Karjoth [6] (EPAL) and Agrawal [2] may be cited as related work. However our approach differs in several aspects: we aim at restoring transparency and control over personal data. This is achieved by redesigning database systems in combination with a contract that is established before any data is processed (consent principle). Our solution comprises both legal measures and a new approach to information systems in order to improve informational privacy. Our main goal is to find a realistic and practical solution to return the control and autonomy over personal data to private individuals. Therefore, our approach differs significantly from existing approaches, both in technical and conceptual aspects. We do not intend our proposal as a replacement for existing privacy-enhancing technologies but rather as an additional concept which could be used to complement these technologies. We also recognize that not all data exist in database systems. We thus feel that the approach of autonomic databases promises to yield benefits that cannot be attained by following existing approaches. With this paper, we hope to contribute to the discussion on privacy issues in the information society. We also make a contribution in the technical and conceptual aspects by proposing a new approach to data processing that pertains to the protection of privacy and can be implemented with available technology.

If such an approach should be widely accepted, its impact dare not hinder business and/or national security. Therefore we do not claim that our system guarantees complete privacy but we believe that this concept can influence people's awareness about their personal data. We hope that the a informational self-determination database system will soon come and that our concept will provide additional inducement for personal data to be sent back to where it belongs. If nothing else, our concept of a usercontrolled Personal Data Identification System may provide guidance for similar structures in other types of data repositories.

In this paper, we begin by providing the founding principles for informational self-determination which are based on privacy principles as defined by Westin, and on current privacy legislation and guidelines. After describing these principles, we discuss a design for informational self-determination database systems. We describe the features of the architecture and explain how the consent principle is implemented and how a portal service helps citizens to keep more control over their data. We also discuss changes in data protection legislation which would be necessary to complement our approach. The paper closes with an extend discussion on the intercultural perspective of informational self-determination database systems.

An overview of privacy invasive technology and related privacy and security issues, state-of-the-art in privacy enhancing technology and the concept of privacy revisited is described in the appendix for interested readers.

Founding Principles for Informational Self-Determination Database systems

Privacy enhancement can be understood as an increase in the control which each customer has regarding personal data which is shared with organizations. In this section, we introduce our concept for privacy enhancement and point out the key principles on which our system design is based.

Our founding principles are motivated by the value of privacy itself. These principles are rooted in existing data protection laws. They articulate what it means for a personal data collection system to responsibly manage private information. We argue for the following six 'new' principles, in addition to the several privacy regulations which already exist. In a few aspects some of the principles are related to but not similar to [15] and [2].

- Consent: People know when their personal data are stored and have to consent this storage.
- Purpose: Persons affected (see consent) must have the possibility to specify the purpose and usage of their data.
- Separation: Personal data and any other business data have to be stored separately.
- Audit: Transactions involving personal data must be recorded in transactional logs. Persons affected can then follow executed transactions and retrace usage of their personal data.
- Participation: Persons affected have access to their personal data, its usage and purpose specification. They can choose where and how to manage their personal data.
- Ease of use: Persons affected have the choice to bundle access to personal and audit data through portals and can define automatically applied patterns.

In comparison with [15] and [2], principles such as 'limited collection', 'limited use' and 'limited retention' are not requested within our approach, but each individual can regulate the mentioned principles as they wish. Within our approach, the 'consent' principle is enforced by law and is strictly connected to the 'purpose specification' principle, which is supported by technology. This infrastructure is expanded in such a way that each individual knows all his or her data sources. This makes principles like 'limited retention', 'openness' and 'compliance' traceable, so that mistreatments of the data-protection law can be investigated. Principles such as 'accuracy' or 'safety' are essential requirements, and as such, will not be mentioned again.

Consent

Nowadays almost any transaction, regardless of what it represents, is recorded. As long as no exact identification of a specific person can be made by using these data, no privacy issues are involved and there is no need for us to care about it. As soon these data are linked to personal data, however, privacy could be jeopardized as described in the appendix.

The first principle is that people, whose private data are stored, must give their consent for this storage, and the specified organization is obliged to inform these individuals 'where and what' data are stored. In most cases, people do not remember which companies store their data; they often have no chance to know this because in many cases they are completely unaware of such a data collection.

Personal data can be used for evaluations and for marketing purposes. It may be sold to other companies without the customer's consent or knowledge and as well as that, such data could even be stolen. Generally people do not pay attention to who manages or what happens with their data, but as soon as they are harassed with spam, telemarketing calls or advertising mails they want to know how this problem has arisen. On the other hand, it is important that organizations are not able to refuse services to any individual on the grounds of an eventual risk. Excluding customers from setting up a life insurance policy, denying access to buildings or generally concealing information are just a few examples of this. The importance of giving customers more information about data storage and the necessity of the customer's consent for further usage of that data is evident. At the same time, organizations gain competitiveness while data management transparency is offered to customers.

Purpose

The first principle illustrates the importance of customers being informed where and what personal data is stored. Now we outline why it's important to specify the purpose as to how personal data can be used.

Personal data can be used for different purposes and it is often used against people's intentions. This data-misuse problem can be solved if organizations put the people affected in a position from which they can influence the further data management. Each organization defines its own purposes which determine the intended use of personal data. Individuals are then able to decide how these settings should be applied to their personal data. For example, a purpose specification may be to receive special offers by e-mail. Organizations can distinguish themselves from competitors and at the same time enhance trust and confidence in their services. This method of participation naturally varies from organization to organization. The only exceptions when peoples personal data is passed without their consent are defined by legal regulations or occur during criminal investigations.

Separation

An area which urgently requires more attention with respect to privacy and security, is the stage at which business data is separated from personal data. During such a separation, business data, which contains sensitive information (e.g. about executed transactions), can be used for data mining without any need for the person's consent. Only an identifier indicates that these data belong to a specific person, so the data are anonymous as long as no connection to personal data can be made. As soon as personal data are requested for a specific purpose by linking to these data, this process must be permitted by the person affected and subsequently recorded in the audit trail.

Audit

Both people and organizations must have the possibility to understand and detect unauthorized uses of personal data. This leads us to the need for audit information where all executed transactions which accessed personal data can be traced. Such information should contain all of the following: who had when with which purpose access to what kind of personal data. This knowledge provides more security to individuals and organizations. This audit information simultaneously supports data protection and helps to minimize fraud. Usually these data are stored at the organizational side, but should be readily accessible to the persons affected.

Participation

While discussing the principles above, we saw why it is so important for people to manage and control the usage of their data. On the one hand, customers must be informed about further utilization of personal data, and on the other hand, they must be able to give their consent for any usage purpose.

To fulfill these requirements, customers need access to personal data which is stored on the organizational side. This participation can be realized in different ways, such as per telephone, forms or internet.

Ease of Use

A possibility for accessing personal data is realized via web portals. The central idea is to aggregate the information shared with all the organizations we are dealing with, and to create one personal portal. This provides people with a better overview and ensures that organizations know where users are managing their data and that they are informed of any changes. The resulting benefit for organizations is improved customer contact, enhanced trustworthiness and a higher level of confidence.

This kind of information aggregation results in a possible security gap. Each person can minimize this problem by depositing their personal data on different web portals. Each portal is physically separated, certificated and protected by a password.

This solution encompasses good standards, open interfaces and the possibility for organizations to buy these systems out of the box, its main objective being to enhance the ease of use by offering standardized interfaces and always adhering to the security requirements.

Design

In this part of the paper, we discuss the design aspect. We study a scenario and visualize the idea of purpose specification with the help of two examples. Furthermore, we outline the structure to indicate the direction in which the setup of such databases could be preceded, however it is not a full implementation guide.

A Use Scenario

Avatara and Belios are two online booksellers who want to enhance customers' confidence in their

company by implementing an autonomic database system. The main idea is to provide a service which gives customers the possibility to define what happens after personal data is entrusted to their companies. Basically, customers set purposes for their personal data usage. During the process in which business data is separated from a customer's personal data, this anonymous business data can be used for data mining and data analysis. References from business to personal data always need a customers' consent.

Additionally, customers are able to see and verify all executed transactions in a transactional list (audit trail), which is automatically updated each time the personal data is accessed.

In this section, we look at examples revealing how the two booksellers handle this requirement and what purpose specifications they define.

Purpose Specification Belios

Avantara and Belios must observe legal regulations and inform customers about these exceptions. For example, in the case of criminal investigations, personal data may be handed over to public agencies without the customer's consent.

Avantara and Belios have different opinions about how much information and customer's cooperation is necessary. Belios defines only a few settings for purpose specifications of personal data, and only asks general questions, for example, if the customer would like to receive advertisements.

Purpose Specification Avantara

Avantara, on the other hand, gives customers various possibilities to define purpose specifications regarding the use of their personal information. For instance, Avantara assumes that customers have preferences as to which information should come via which channel. Hence Avantara offers various channels for communication and makes distinctions between private and business phone numbers. Furthermore, customers can classify how they prefer to be contacted. These options are contracted under the tab "Contact". Under "Order", general order properties are defined, such as whether or not customers wish to be informed about their order status. Other companies and individuals are also employed to perform functions on Avantara's behalf. Examples include fulfilling orders and delivering packages, sending postal mail and emails, etc. They require access to personal information which is necessary in order to perform their functions, but they are not permitted to use it for any other purpose. Avantara guarantees that business or personal data is never passed to third parties without the customer's prior agreement, and that customers are always asked if data may be used for purposes other than those defined at the beginning. For customers who don't want to answer each single question under the "Defaults" tab, Avantara defines settings-categories for data usage. The data usage allowance can be set on "Minimum" or "Maximum". Last but not least, Avantara gives customers the chance to define the intensity of advertisement.

Alice and Bob (compare [2]) are looking for a skilled online bookseller, whereby Avantara and Belios are short-listed. Alice is a privacy fundamentalist who normally doesn't want companies to retain any information once her purchase transaction is complete. However she is willing to commit her personal data in order to receive some specific information if she can be certain that her data will be handled confidentially and only for the chosen purposes. For this reason, Alice decides to buy her books at Avantara since there she has the best overview of her personal data usage. Bob, in contrast, is a privacy pragmatist. He appreciates the convenience of only having to provide his email and postal address once when registering with organizations. He likes to receive new recommendations, but does not want to be part of purchase circles. He also chooses Avantara but his reasons are different from Alice's.

Tent is Avantara's privacy officer. He is responsible that the information system complies with the company's privacy policies. Mallory is an employee and he has questionable ethics.

Architecture

Finally, we present the architecture of an autonomic database. Central to the design is the active participation of customers in providing specific information within the organizational systems.

Components

Customers Data Requestor is responsible for opening a communication channel to the Request Handling Agent, which is located on the Customers Data System side.

Request Handling Agent only accepts properly formulated requests from the corresponding Customer Data Requestor.

Privacy Settings Rule Model covers rules which determine for which purposes customers' personal data can be accessed. These rules are constituted in the Privacy Control Settings. Trent designs these privacy definitions with regards to the company's privacy policy. For instance, he determines the purposes as to when a customer's email address can be used.

Rule Compliance Validator examines whether or not a personal data request complies with the Privacy Control Settings of each user.

Access Control takes care of accesses before and during query execution. Access Control is carried out on both the Business and Personal Data Identification System.

Query Intrusion Detection checks the accuracy of accesses after the queries by comparing the access with the usual access patterns for queries with that purpose and by that user. For example, Mallory decides to steal all email addresses of Avantara's registered users and to sell them to Avantara's competitors. Normally customers' email addresses can only be accessed for sending them recommendations or offers, or to enable order status tracking etc., as defined in the Privacy Settings Rule Model. Before the query results are returned, the Query Intrusion Detection matches these queries with the usual access patterns and detects the fraud.

Audit Trail records all possible queries for privacy audits and addresses challenges regarding compliance. Furthermore, this is where the customer's personal preferences as well as any changes to the Privacy Control Settings are maintained. Since customers have access to audit information, they are in a position to view all transactions and to detect any fraud.

Privacy Policy

Fig. 1 illustrates the separation of customers' personal and business data. The privacy policies of the two systems therefore differ in certain aspects, as explained in the following section.

Business Data System

Authorized users and applications of the Business Data System are specified in the privacy policy. These are the set of Avantara's employees and applications who, or respectively which, can access particular information. The anonymous business data is accessible for purposes such as data maintenance, data mining and data analysis. As a result of the data separation, Avantara doesn't require a customer's personal information for most data mining and analysis activities - that is, not until Avantara addresses its customers directly.

Personal Data Identification System

The privacy policy for the Personal Data Identification System is more sophisticated and consists of three main parts.

Authorized users: This is a group of employees, customers and applications. Employees and applications access this data for maintenance purposes only. Customers, in comparison, access the Privacy Control Settings to assign their preferences and restrictions with regard to data usage. Moreover, customers access Audit Trail information to view and verify the suitability of the use of their personal data. Returning to our example case, Mallory is employed by Avantara to maintain customers' business data, therefore he has no authorization to access customers' personal data.

Rule Mechanism: Privacy rules are defined in the Privacy Settings Rule Model. This model covers rules which determine the general purposes for which customers' personal data can be accessed. The Rule Compliance Validator checks customers' Privacy Control Settings to examine if specific accesses should be allowed.

Request / Reply Mechanism: The only way of connecting anonymous business data to customers' personal data is via a communication channel between the Customer Data Requestor and the Request Handling Agent. The Customer Data Requestor asks for information from the Request Handling Agent, which handles these requests and sends back a reply verified by the rule mechanism.

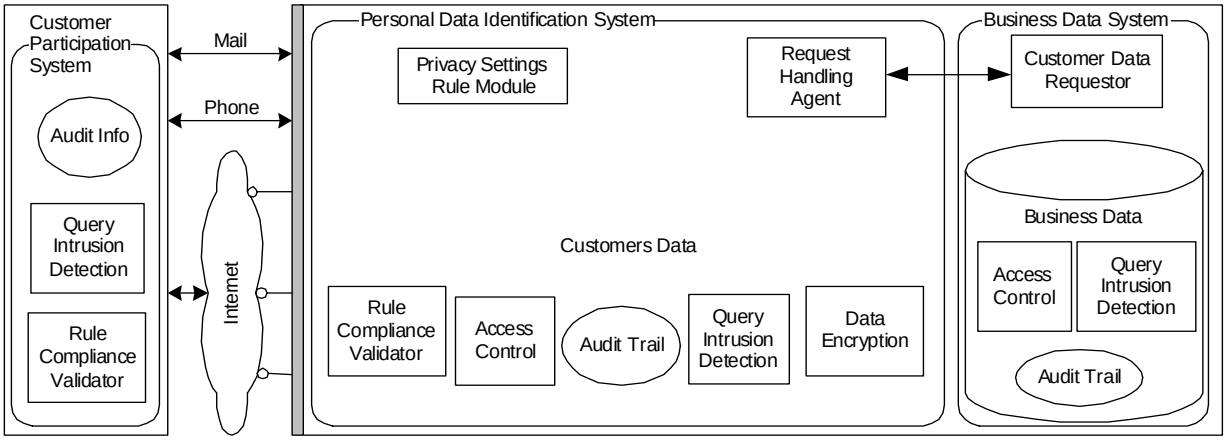

Figure 1 Architecture

Queries

Avantara decides to launch a new marketing promotion, and therefore selects 500 records from the Business Data, with the intention of sending these customers specific recommendations per post or per email. In order to do so, Avantara needs to access the Personal Data Identification System where customers' addresses are stored. The access from the Business Data System to the Customers Data System is only possible via a controlled channel. All queries for customers' personal data are first sent to the Customer Data Requestor. The Customers Data Requestor forwards these queries to the Request Handling Agent, which is located in the Personal Data Identification System. The Request Handling Agent passes all properly formulated queries it receives to the Rule Compliance Validator. The query for customers' postal or email addresses with the purpose "recommendation" was sent by an authorized employee at Avantara. The Rule Compliance Validator now checks, in accordance with the Privacy Settings Rule Model, if this query can be accepted. After the commit, customers' Privacy Control Settings are checked. Alice stipulated in her Privacy Control Settings that she doesn't want to receive any recommendation whilst Bob would like to be sent recommendations per email. Therefore only Bob's email address is sent back to the Customer Data Requestor.

Let's suppose that Alice unexpectedly receives a recommendation from Avantara, despite having told them that she doesn't want this. Since Alice has access to the Audit Info where all transactions are recorded, she can verify the permission of the received email and complain to Avantara about the mistreatment of her personal data.

Design Considerations

In this section we outline the six principles, upon which our approach is based. The purpose of this exercise is to demonstrate the feasibility of these principles.

Consent

The guarantee that explicit consent is required before personal data can be stored or utilized for further purposes has turned out to be a challenge. The first premise is to be absolutely sure where our personal data is stored. With the constitution of a data protection law, this requirement can be fulfilled.

Data which belongs to a person can be distinguished between being assignable or not assignable to a person's identity. Assignable data is, for instance, our surname, forename, address, telephone number, email address, etc., and can be directly assigned to a person, i.e. a person can be identified with this information. From now on, the term personal data will be used instead of assignable data. Data which is not assignable includes a person´s age, the items he or she purchased the previous month, the amount of rent he or she pays, etc. This information, viewed separately, can belong to anybody and isn't directly assignable to a certain identity.

Organizations mainly produce business data, as opposed to assignable data, and for the majority of processes, such as data mining, market research activities or individual steps within a whole business process, they do not require personal data. Therefore data which is not assignable can

theoretically be used for these purposes without the person's consent. Anyway, within the described system the usage of these data could be controlled, too. There are a few cases in which it makes sense, as for example data mining applications within medical data.) However as soon as institutions claim to use assignable data, the person's consent must be obtained.

The first step, concerning how data storage can be regulated, is the identification of all existing data islands. Possible institutions and service providers who may retain personal data are: Education, Financial and Legal, Government, Health and Medicine, Home, Media & Telecommunication, Personal Care & Recreation, Shopping, Travel and Transportation. The list is not complete and can, without doubt, be extended. We simply want to illustrate how widespread personal information can be, and how easy or difficult it is to get consent. The astonishing result is that most institutions can theoretically obtain a person's consent for collecting personal data very easily.

Now we will take a look at some cases where it is more difficult to obtain consent, or where organizations are not concerned with obtaining consent.

In cases of criminal investigation, it is particularly difficult to obtain consent. For instance, DNA information and fingerprints of suspicious persons are collected, although the individuals are not asked for their consent. In Great Britain, the DNA database already holds 1,8 million samples [1]. If persons behave in a suspicious way, information is recorded about them without their knowledge. For example, telephone calls can be intercepted, or the caller's position can be located via mobile phone. These privileges are regulated by law, and are only permitted to certain security institutions, such as the police, the civil defense agencies or the military, all of which are legally allowed only for specific purposes. The informational selfdetermination concept does not hinder this kind of investigation, however all transactions, where personal data is involved, are registered and can be used in cases of law abuse.

Another hidden data record is to be found in buildings and areas where high security is needed. Examples of this may be airports or banks, where face scans and observation cameras are installed. In such buildings, any suspicious persons must be identified in order to control their access rights and to observe their behavior. For this form of identification, it is difficult to obtain consent and is often not reasonable.

Furthermore, the protection of data privacy is particularly difficult when institutions hold various personal data. Administrative bodies, for example, hold all sorts of personal information: birth certificates, marriage and/or divorce papers, official documents certifying a person's citizenship and religion, employment contracts or registration cards, information concerning taxes, penal records or monetary records. Especially under E-Government, numerous web applications are integrated, and are used by various national administrative bodies. Any interactions and information flows, which take place for the processing of services between these bodies, must be revealed and consented by the persons involved.

If personal data is obtained illegally, often for use in marketing purposes, it is more difficult to retrace. Institutions sometimes carry out indiscreet market researches or advertising, without the express permission of the person being interviewed. Another problem arises when personal information is handed out to third parties, who carry out instructions on behalf of an institution. Some institutions are even requested to collect personal information and to resell it to other interested institutions. Many are network marketing specialists: they make home visits in order to present their products, and in return, they expect the host to provide them with the addresses of friends and acquaintances. Friends and acquaintances generally know a lot of information about us, and could hand out private information to organizations without realizing that revealing these data may be unwanted. Another simple method of collecting personal data is by organizing lotteries and contests. Afterwards, these entrusted personal data may be further used for unapproved marketing purposes. Some software companies even gather personal data when persons try to get help or get an update from their websites. At the time a person installs these programs, he/she is not explicitly asked for their consent, and in most cases he/she is not even aware that personal data is illegally stored in organizational systems. To prevent illegal data usage, persons must insist on more transparency on the part of institutions. Transparency can be obtained by specifying all business processes and transactions when personal data is used, as described in this paper. Illegal treatment of data is not hindered with the informational self-determination concept, but the detection of such data handling is strongly supported.

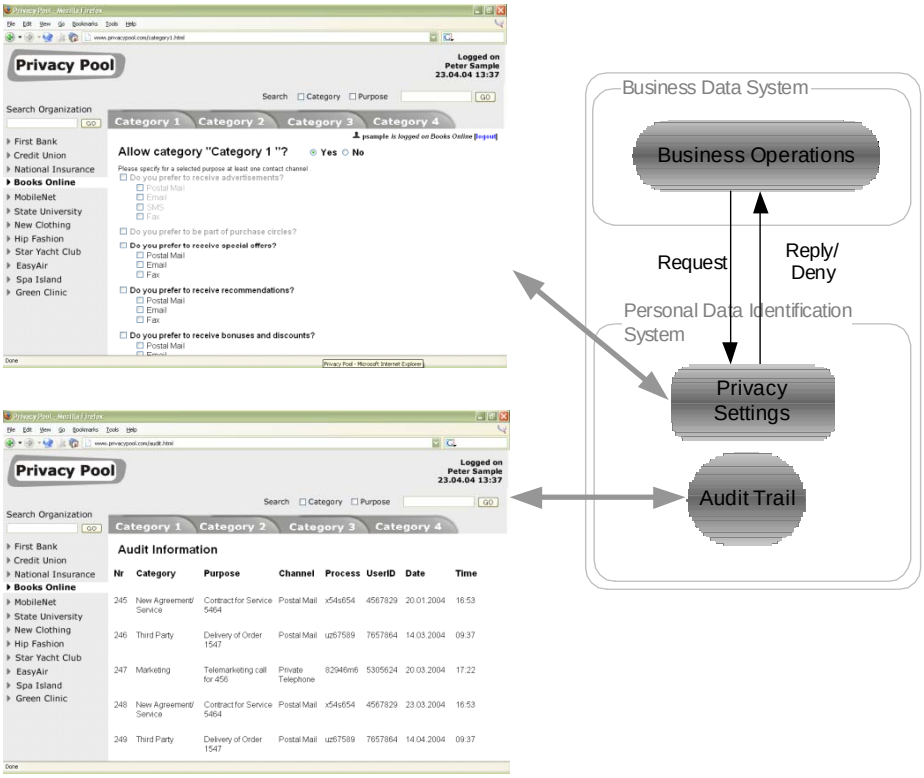

Figure 2 Data access

Purpose

Once institutions have been given a person's explicit consent that personal data can be stored on their institutional databases, the person wants to know what is going to happen with their personal data in the future. Therefore individuals must be able to access their data, in order to view audit information or to specify future purposes for which their data may be used.

As already mentioned in the founding principles, the user relevant information can be provided to persons via telephone, mail, forms or internet. In this part of the paper, we concentrate above all on a person's participation via the Internet. This approach makes it easy to access relevant data on the institutional side, in order to define settings for data usage or to view audit information, as shown in Fig. 2. Due to the fact that no personal data is stored directly on the privacy pools, there is less risk of unauthorized data being viewed.

The access can be realized by means of privacy pools, which are comparable with web portals. Everybody registers themselves at the privacy pool of the institution which holds their data. After having done this, they can log in from their personal computer over a web browser.

Many institutions can be registered at the same privacy pool which means that persons access a portal where different services of different institutions are available. To minimize the risk of unauthorized access, separate access information for each institution can be provided. This solution assumes that only the access information to the privacy pool is the same for all institutions. (see Fig. 3) gives a concrete illustration of three privacy pools with their registered organizations. Different possibilities are outlined, as to how persons can use their personal computer to log in to the privacy system. Due to the fact that access via the Internet is not secure, the identification of and communication between customers and organizations must be secure and reliable.

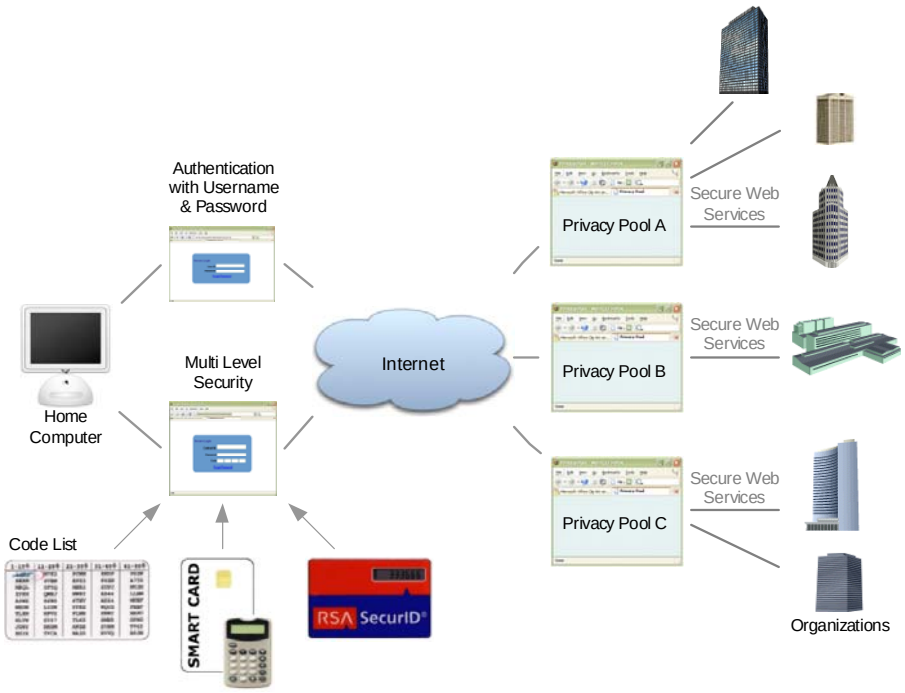

Figure 3 User acces to privacy pool

Separation

The separation between the personal data identification system and the business data system can be either logical or physical. For smaller organizations and companies, the personal data identification system could be developed and operated by another company.

Audit

Audit is connected to the described queries and visualized within the portal.

Participation

Participation is an integral part of ease of use.

Ease of Use

Considering that the list of all possible data islands can reach an unmanageable complexity, a simplification of the data access is a prerequisite. For the requirement "Ease of Use", the access to the different privacy pools can be unified to one single point of entry. As shown in Fig. 3 a person logs directly into the shared access point where a separate username and password is needed. This information is stored at each organization. To access resources on organizational Web servers, separate authorization and authentication is needed, otherwise the loss of security would be unacceptable. From the single point of entry, persons are redirected over the internet to the corresponding privacy pools.

To apply a unified access to the privacy pools, a standard must be established which fulfills the highest security requirements and is accepted by a great many institutions.

Categories

In part 4.1 situations and organizations possibly holding personal data were identified. To enhance the protection of personal data, an institution can only access these data if a reason or purpose for the data usage can be proven and if the owner of the data allows this access. In order that persons can choose if they would like to be contacted regarding a certain purpose, or that they can stipulate for which processes their data may be used, institutions must define all probable purposes and present these to their customers. Moreover, institutions shouldn't be able to introduce and define new purposes, which weren't initially defined, without first obtaining a customer's permission for these additions.

If every institution or organization would define their own purposes, this would quickly become unmanageably complex. Therefore all existing data usage purposes are defined in such a way that

every institution is able to apply them for their own specific services. Purposes which are similar and belong together are grouped into rough categories. The categories must be of central importance in order that a wide range of purposes can be covered.

Order customization: For customization reasons, the preferences of existing customers, such as order history or order status, are recorded and can be requested to facilitate the next order.

Payment: Personal information, such as address, is needed in order to send invoices, or the credit card number must be known, if the payment transaction should be made this way.

Shipment: Addresses of persons are also needed in order to deliver the ordered items.

Abstracts of accounts: To generate and send abstracts of accounts, a link to personal data must be made. Abstracts of accounts can be made by banks, insurances, bonus or shopping card companies, etc.

Personal customer care / services: In order to provide services, customer consultants or front office employees generally need access to personal information, through counter applications for example.

Agreements for new / altered services: The first registration usually only contains basic personal information, such as name, address and telephone number. In order to perform new or different services, additional data input is usually required.

Internet & computer information: Organizational web servers store cookies and information about browsers, operating systems, internet service providers, IP numbers, websites visited, along with the time, date, and duration of the visit.

Marketing: The marketing category contains advertisements, purchase circles, telemarketing, special offers, recommendations, etc.

Data mining & market researches: Data mining and market researches can often be performed without personal data, but in some cases access to personal information is necessary.

Information brochures & newsletters: Contrary to the marketing category, information brochures and newsletters primarily inform customers.

Third parties: Third parties access personal information to perform functions on behalf of an organization. For instance, this could be the delivery of orders, postal mail, etc.

Legal regulations: This category contains all purposes for which data access is permitted without the explicit consent of the data holder, for instance, the data protection law and banking secrecy.

Customizing Categories

Each organization chooses categories which cover their services. The categories and corresponding purposes are then customized in order that services can be provided correctly and all specific business features are taken into account. Customers access these adjusted categories and purposes via a privacy pool, and agree on which service they want to accept or not (compare Fig. 4).

The next question which arises is how to present these categories and purposes to customers. Where convenience is concerned, some customers want to make simple and, at the same time, very general decisions. Other customers want to customize each purpose or category individually. For example, one customer wants to define that he/she never wants to be contacted by telephone. Furthermore, he/she can deny entire categories. If somebody is sure that he/she never wants personal data to be used for any marketing purpose, the complete category can be quickly and simply disabled. However disabling an entire category carries the risk that desired purposes are also denied. By disabling the whole marketing category, for instance, purposes such as bonuses or discounts are also disabled. Therefore, before a category can be disabled, a combined extract of purposes must be shown under the category, or customers should at least be informed about the undesired effects of disabling a complete category.

Some categories cannot be disabled at all. Generally speaking, the importance and adaptability of categories differs considerably. Categories like payment, shipment, abstracts of accounts and personal customer care necessitate personal data so that minimal services can be provided at all. Some information should reach the customers in any case. For instance, abstracts of accounts are necessary in order to ensure transparency. The exclusion of the definition of new/altered agreements makes it impossible for new services to be offered. If payment and shipment details could also be disabled, it would be impossible to sell anything. Legal regulations are the exception in this respect

since they only inform, and must be accepted by the customer in any case. These were examples where categories can't be disabled. The accepted categories can be better adapted to meet personal needs and requirements. Categories such as customization, internet information, marketing, data mining & market researches, brochures & newsletters, and to a certain extent, third parties, can therefore theoretically be disabled.

Could there perhaps be further possibilities than manually defining if a category or purpose can be accepted or disabled?

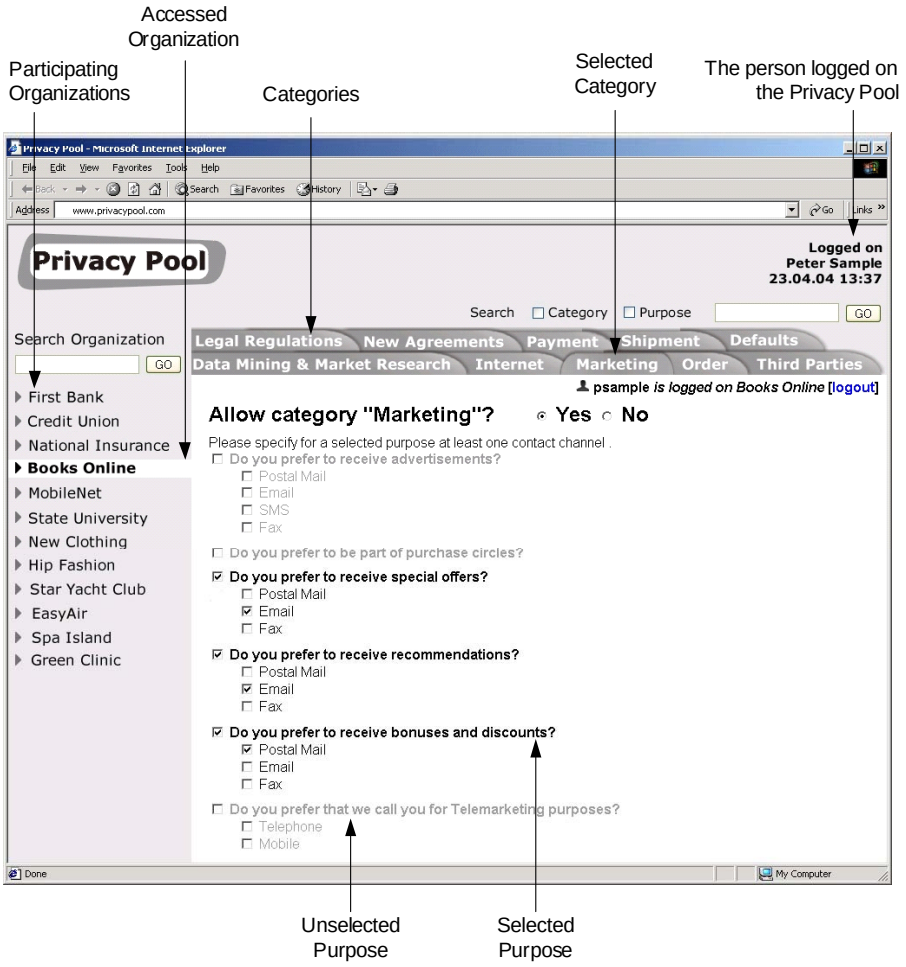

Figure 4 Settings for Books Online

Institutions define templates with privacy levels to simplify the customers` choice between categories and purposes. A template for customers with very high data security requirements denies all unnecessary data usage. A template with standard security precautions assumes that customers want customization, but don't want to receive all information, such as newsletters or advertisements. Furthermore, all competencies can be delegated to an institution. This template contains the most costeffective settings. Customers are preferably contacted via email as opposed to via telephone or postal mail, because this method is cheaper.

Manually defined settings are also stored as a template. The advantage of this is that customers can always return to settings which have once been made: for instance, a customer can remove the category marketing while he/she is on vacation, and can later reset it. The general and manually-defined templates should be available to all institutions which are part of the privacy pool.

Fig. 4 illustrates an example where the person Peter Sample is logged into the privacy pool, and at the same time, into the organization Books Online, which sells books and media over the internet. As shown in Fig. 4, he is customizing the category

'Marketing'. He has allowed the category 'Marketing', and he has selected the preferred purposes and contact channels. The unselected purposes are inactive and are therefore colored gray.

New Challenges

Now we describe some interesting problems which we identified in our principles and design. This list is by no means complete; its purpose is to initiate discussions.

Consent

The cornerstone for informational self-determination database systems would be a new international data protection law, which requests the explicit consent of a person before personal data can be stored. Furthermore, the law stipulates that this person must have access to their data, to specify purposes and to control audit information.

Within this law, several questions are raised. There will be a certain amount of administrative work and it will not always be clear how to set the process up. For instance, the user must first give his / her consent, before his / her personal data is stored, and not the other way around. How can organizations which do not care about this law be identified? Are normal individuals qualified to handle their personal data or instead to instruct a company specialized for this purpose?

However - and this is a crucial point – at least a person knows which databases store information about him / her.

Purpose

At a first glance, purpose specification may appear easy. However selecting what kind of usage from personal data a person allows depends heavily on the way in which this can be achieved and how these usages can be presented and categorized. No one is willing to spend several minutes specifying purposes, therefore a low amount of fixed categorizations have to be defined in which each category includes several purpose specifications. Then people can choose to make settings either only on the category level and / or for each purpose. The categorization must also be independent of the branch or industry. To setup, define and become widely accepted, such a general categorization of purposes is essential and its development may be a tough task.

Separation

Business data and personal data are often already separated in large-sized companies. Different applications use these data. On the other hand, in small and middle-sized companies these data are normally stored together and are only used by one main application. A physical or logical separation is necessary according to the principle of separation. This makes any IT-architecture more complicated. In addition, the architecture has to be extended with a strong identification functionality. To increase trust and confidentiality, the 'Personal Data Identification System' (see Fig. 2) should be certified by a third party.

Audit

Generating audit trails that are in the hands of the people affected could provide a strong and powerful tool for protecting privacy. First of all, these audit trails can be investigated by the organizations themselves in order to detect internal misuse. Secondly, each person can scan these data and convince himself / herself in compliance with the audit trail of his / her personal data, or in the case of misuse, can place a complaint. Last but not least, a person can engage external software agents to monitor his / her audit information and to be automatically informed if a violation is detected. Within this scenario, three main questions arise. How can an individual set up his / her complaint and who will receive this message? What kind of competence or interest could such a 'compliance office' persecutes? What kind of consequences may occur for the principal offender? Furthermore, 'Rule Compliance Validator' agents activated by the customer represent several security and privacy risks, despite being convenient for the customer.

Participation

Participation requests a certain kind of connection to the control equipment of the purpose specification and audit information. This communication and requested identification must be secure. Misuse cannot be tolerated.

Ease of Use

We propose a hybrid solution. Each person can decide how centralized he / she would like to treat

his / her personal data. A centralized system is quicker and easier to handle but encompasses more privacy risks than a decentralized system; however they could both provide a higher level of security. A centralized system is a far more attractive target for illegal transactions, because full data profiles related to specific users are available. The system's structure should at least be digitally secured against possible misuse and should guarantee the respect of a citizen's privacy.

Intercultural Perspective

Basically, an Informational Self-Determination Database System is applicable in every institution and country. But the global differences between countries could theoretically hinder a worldwide diffusion of a self-determination system. Considering the technological, cultural and legal differences between countries the question arises if it is reasonable at all to introduce an Informational Self-Determination Database System in countries for instance with a lower technological development or a totally different cultural and legal background then where I live. Therefore, it will be discussed now which oft these factors are globally responsible if the system is accepted in a country or not?

Data Protection

For a successful implementation existing data protection regulations of the different countries plays an important role. The differences between national legislations complicate transborder data flows or made it impossible. How to solve this problem was the major task of the European Data Protection Directive and as well it is a major global problem. Optimally, in each country a specific data protection directive for the privacy enhancing system is postulated or at least the system is supported by existing data protection regulations. To show the different impact of data protection the current situation in Japan, India and Latin America will be outlined.

The consciousness for sensibility of personal data is in Japan very high. According to a study of the Center for Social and Legal Research [14] Japanese people are in equal measure skeptic when personal data is used by the government and by organizations for commercial purposes. For the study 1000 people were interviewed per telephone. The result points out that the fear of potential abuses is high. 74% of the interviewees are disappointed how the government maintains personal data and 67% believe that consumers have no control how organizations handle personal data. The uncertainty toward the government is reinforced through the occurrences in the year 2001 where the government has passed to the Public Security Investigation Agency (PSIA) personal files without the consent of the data holders. This privacy violation was allowed under the excuse that PSIA examines groups which are suspected to threaten the national security. Although the law stipulates that data requests must be constituted by law this personal information were handled out to PSIA. Furthermore, persons were illegally intercepted or their privacy was otherwise violated. Nevertheless, a lot of people in Japan still believe that telephone and email intercepts are necessary to minimize the growing number of crimes.

To this topic in the late nineties it has been lot of media: „There has been a flurry of news reports on privacy and data security violations. Likewise, government privacy initiatives, including the revised Residence Registration Act, the new Wiretapping Law, the Freedom of Information Act and the proposal for a comprehensive data protection law, have received broad media coverage. The news media has publicly aired comments and reservations to the draft for new comprehensive privacy legislation." [4]. According to that most Japanese were concerned about privacy issues. Finally, in March 2001 new data protection regulations were enacted to form a framework for the commercial usage of personal information. Main content is that personal data can not be passed to third parties without the consent of the person concerned. Every institution is liable to disclose which personal information is stored. The usage of personal data is prohibited for other purposes then claimed at the beginning. Collections of personal information must be transparent. Japan's Personal Information Protection Act which regulates both private and public sector was finally passed in May 2002.

The security needs and consciousness on privacy of Japanese population is clearly present and definitive. Additional data protection which is provided by the Informational Self-Determination Database System people definitely gains more confidence in institutions including the government.

In contrast, in India the importance of outsourcing is crucial for the economic. Out-sourcing is the act of transferring some function, for instance software maintenance or development, operation of a data processing center, or operation of a "call center", from one location or company to another. India is

particularly attractive for outsourcing because the salary structure is much lower then in the United States or in Europe and there is a multitude of highly trained individuals who are comfortable speaking English [3].

A data protection and privacy law such as the EU Data Protection Directive or the Safe Harbor is inexistent. So far, according to the Information Technology Act of 2000 only unauthorized access and data theft from computers and networks are prosecuted with a monetary penalty, but specific provisions relating to privacy of data are not covered.

The absence of data privacy legislation in India has also proved to be a disadvantage for Business Process Outsourcing (BPO) to Indian companies and is a strong reason for stopping the movement of BPO work to the country [11]. The only way to beware India from an outsourcing stop is to enact new data protecting regulations. The Indian Business Process Outsourcing industry has already pressurized the Indian government to enact a data protection law in order to prevent from adverse impact on the economy. The other concern is that the Indian BPO companies and their employees are becoming privy to personal data of the clients and customers of outsourcers [9]. There was even a case reported of an employee in a call centre, who has misused credit card information and other details of a US citizen [11].

"It is becoming extremely important for India to have in place a distinctive legal regime promoting data protection," said Pavan Duggal, a Delhi-based cyber law consultant. "This is necessary to create appropriate confidence among investors and foreign companies to the effect that the data they send to India for back-office operations is indeed safe, and there are appropriate statutory mechanisms in place should a breach of data take place." [12].

The Indian government is on the way to insert new clauses in the Information Technology Act of 2000. The main objective of the new clauses is to conform to the socalled adequacy norms of the European Union's Data Protection Directive and the Safe Harbor privacy principles of the US. "The adequacy norms allow the EU to declare that third-party countries have levels of data protection that conform to European standards and thus allow data on EU citizens to be transmitted outside of the union" [11].

Similarly to Japan, India is anxious to gain more confidence, in order to keeps the supremacy as BPO offering country.

In central and south America various countries including Argentina, Brazil, Chile and Peru have already implemented data protection laws. In Latin America privacy is referred as Habeas Data. The constitutional right shows variations from country to country, but in general, it is designed to protect among other things the privacy and information self-determination of persons. Habeas Data has been described as: "a procedure designed to safeguard individual freedom from abuse in the information age" [5]. An objective of the Habeas is to comply with European Standards in the first instance because the European Directive on Data Protection requires its members to impose strict restrictions against the transfer of data to countries that do not posses data protection regulation as postulated in the Article 26 (4) of the Directive 95/46/EC. Chile has enacted a data protection law that regulates data handling and storage in a very European way. "Brazil and Argentina have also decided to follow the European lead." [10]. Being based on the existing legislation of the EU, it is fair to assume that it will provide more protection than the existing Habeas Data Constitutional provisions and that it will include some of the principles required for obtaining adequacy level from the EU [7].

Since July 2003, the European commission recognizes that Argentina provides an adequate data protection level for personal data [8]. Argentina has become the first country in Latin America which has received the EU Data Protection Working Party's approval for its data protection framework. This means that data flows between Argentina and EU member states are freely and do not violate the European Data Protection Directive.

The effort of Latin American states to adjust their data protection regulations, e.g. the Habeas to the European Data Protection Directive shows an increasing importance of data protection. The Indian government is constrained by the Business Process Out-sourcing industry to enact new data protection regulations. These efforts cohere definitely with the hope to enhance own abilities to compete on the market through closer cooperation with EU member states or the US. Missing data protection regulation can definitely harm trading partnerships or the people's confidence in institution's trust-worthiness. But out of the need for more privacy regulations and protections it can be concluded that an Informational Self-Determination Database System

which supports additional personal data protection will be faster accepted and implemented. Furthermore, it strengthens the established privacy laws. For instance, in countries where data protection regulations have not yet a European protection level an establishment of an Informational Self-Determination Database System can nevertheless enable institutions to act internationally. This can be either to trade with international institutions or to fulfill services for those where personal data is needed or to offer international customers equally high or even higher data protection. These gain of high importance since the explosion of the Internet (where people are able to buy goods world wide).

Cultural Aspect

The question here is how the culture influences people in respect of personal data. As we have seen, in some countries data protection regulations are far behind the EU Data Protection Directives. Because of differences in data protection regulations it is assumable that people in every country seize data protection differently. For instance, the Indian citizens have been more open in divulging their personal information. This can be explained through the lower and less explosive increase of the technical progress. Comparing to the developed countries in the west, in India the penetration of the Internet and technology was much lower. Thus data privacy has not yet become such a concern as it is in the west. Before the enormous growth of BPO industries there was no pressure on the government to enact a data protection law as it is now [9].

This interdependency between people's attitude to privacy and the existing data protection regulation is also observable in other countries with a similar technological development. The technological achievements produce higher flows of information. These higher data flows implicate also higher data usages. Therewith, also the possibility of undesired data insights and abuses escalates. Mailings and telemarketing calls are bothering and telephone interceptions or observations are a deep cut in the privacy. Illegal data misuses can lead to people's exclusion from services, such as insurances, accesses to places can be denied or they can be excluded from schools or jobs. Hens, according to all the possible harms that can appear, people get independently from their culture and origin equally concerned about privacy. An Informational Self-Determination Database System helps to prevent from possible harms. But the people must be informed of the advantages before the system can be accepted everywhere. The importance, efficiency and profit of an enhanced database system must be transparently communicated. Important for people is to know what additional rights are supported through the system. For instance, that the people can gain awareness where personal data is stored and for what purposes personal data can be used for.

As already mentioned data protection has in each country a different significance. In Europe everyone is considered that his / her data is handled very carefully for commercial purposes because people have made already bad experiences. Personal data is collected illegally. People often receive unwanted telephone calls or advertising mails without knowing the initiator. In contrast, in the USA or Japan the people are more sensible when personal information is used by the government. Since people were intercepted without their consent or even any knowledge about it. Consequently, everybody and everywhere appreciate that his /her personal data is protected by additional privacy regulation, independently which institution is using personal data, and this does not depend on the culture.

Technological Aspect

According to the Technology Achievement Index from the year 2001 [13] which reflects the capacity to participate in the technological innovations of the network age most Latin American countries are either "potential leaders" or "dynamic adapters" for creating and diffusing technology. India is as well part of the "dynamic adapters". Mainly between the "dynamic adapters" and some "leading" European member states huge technical deficiencies exist. Only countries with a certain technological development will consider the possibility to implement an Informational Self-Determination Database System. Furthermore, the slower development of technologies has also leaded to a slower development of data protection regulations. This can be explained among other things by the marginal amount and frequency of personal data usages by institutions. In contrary, Japan ranks among "leading" countries and has nevertheless vary late adapted privacy regulations for personal data. Absolutely not all technologically developed countries have sophisticated data protections; there are always additional aspects relevant. In Japan, for instance, the political development and powerful position of the government are also responsible for the slow development.

Position of the Institution

Besides the technological aspect of a country also the position of each single institution is important. A difference between small and medium sized businesses and major enterprises must be made. Generally with the size of an institution also more resources are available. Missing monetary and human resources or the absence of technical skills and know-how can hinder an adoption. However, every interested institution should be able to implement the system. For institutions which have rare resources a ready-made Informational Self-Determination Database System can be purchased where the complexity is reduced to the minimum and the system can be easily installed. Existing lacks of the underlying infrastructure and technical fundamentals can be thereby made up. Looking at the global aspect, then this can arise especially in less developed regions and countries.

Nevertheless, institutions which have a particular size and a subsidiary structure can easier support a privacy enhanced system then smaller organizations. Additionally, if an institution is already a global player, i.e. internationally acting a new system will be rather adopted, mainly to conform to other cooperating global players or countries which already have a higher standard for data protection.

Conclusion

The aspects which could hinder an implementation worldwide were so far outlined, but what happens if one country successfully introduces the system and another which is for instance an important trading partner does not? Are data flows between trading partners entirely exposed and vulnerable to unauthorized misuses?

The privacy enhancing system guaranties that customers are always informed about third parties, which services are carried out by them and for what purposes which data must be handed on. Only information is handed on which is absolutely necessary to provide services. Nevertheless, third parties can misuse and divulge personal data and their partner institutions are not able to hinder them. Consequently institution will rather enter into a partnership where they can guarantee customers' the trustworthiness of their partners. Hens, cooperation where data flows are necessary are definitely securer if all partner support an Informational Self-Determination Database System.

A lot of institutions do not have third parties or transborder data flows. This is applicable in areas such as the education, health care, insurances or home services, etc. Then, in the first instance an institution and their customers can profit from the system. These institutions are independent thereof if another country or institution decides to implement the system or not.

But if the government supports an Informational Self-Determination Database System then it is necessary that also all authorities implement the system; whereas institutions are free to implement the system.

However, there are two problem areas. Institutions which act globally and support privacy enhanced databases have not the same profit from the system in countries where the system is not yet known or achieved. Nevertheless, the institution can still offer the system and try to convince their international partners of the system's efficiency.

An institution which is already acting globally and has not yet the system must adjust their system in order to be able to cooperate with these countries which already have adapted the Informational Self-Determination Database System.

Generally an institution can stand out from other institutions by offering better personal data protection. However, an international adjustment would be the optimal solution.

Proceedings of the symposium "Localizing the Internet. Ethical Issues in Intercultural Perspective" sponsored by Volkswagen*Stiftung*, 4-6 October 2004, Zentrum für Kunst und Medientechnologie (ZKM, Karlsruhe)

---

1 Plans to extend taking of fingerprint and DNA samples. Available from: http://http://www.number-10.gov.uk/output/Page3359.asp. Last Visited: May 2004

2 Agrawal, R., J. Kiernan, et al. Hippocratic Databases. In Proceedings of VLDB 2002. Hong Kong, China, Morgan Kaufmann, San Francisco. pp. 143-154, 2002

3 Brender, D. Data Protection Law in India: A Change of Direction. Available from: http://www.whitecase.com/article_data_protection_law_in_india_a_change_of_direction_1_12-2004.html. Last Visited: July 2004

4 Center for Social and Legal Research. Guide to Privacy and Data Protection Program in Japan and Guide to E-Commerce and Privacy Development in Japan. 2000

5 Falcón, E. Hábeas Data: Concepto y Procedimiento. Buenos Aires, pp. 28, 1996

6 Karjoth, G., M. Schunter, et al. Platform for Enterprise Privacy Practices: Privacy-enabled Management of Customer Data. In R. Dingledine and P. Syverson, Privacy

7 Guadamuz, A. Habeas Data vs the European Data Protection Directive, The Journal of Information, Law and Technology (JILT). Available from: http://elj.warwick.ac.uk/jilt/01-3/guadamuz.html. Last Visited: July 2004

8 IP/03/932. Data protection: Commission recognizes that Argentina provides adequate protection for personal data. Available from: http://europa.eu.int/rapid/pressReleasesAction.do?reference=IP/03/932&format=HTML&aged=0&language=EN&guiLanguage=en. Last Visited: July 2004

9 Kathpalia, V. and V. Parikh. India under pressure to enact a data protection law. Available from: http://economictimes.indiatimes.com/articleshow/610334.cms. Last Visited: July 2004

10 Palazzi, P. Protección de Datos Personales, Privacidad y Habeas Data en América Latina. Available from: http://www.ulpiano.com/Recursos_Privacy_LatinAmerica.htm. Last Visited: July 2004

11 Ribeiro, J. India poised to tighten data protection law. Available from: http://www.computerweekly.com/Article130076.htm. Last Visited: July 2004

12 Ribeiro, J. India works on data protection law. Available from: http://www.computerweekly.com/Article122612.htm. Last Visited: July 2004

13 The Technology Achievement Index. A new Measure of Countries' Ability to participate in the Network Age. Available from: http://hdr.undp.org/reports/global/2001/en/pdf/techindex.pdf. Last Visited: July 2004

14 Westin, A. F. Japan Consumer Privacy. A National Survey of the Japanese Public and Comparisons. U.S. Center for Social and Legal Research, 2000

15 Westin, A. F. Privacy and Freedom. Atheneum, New York, 1967

Appendix

The explosive development of privacy-invasive technology like identifying technologies [12] [40] (biometric technologies, radio frequency identification, bio-implants, DNA sniffers), location based services (cellular systems, wireless local area networks, bluetooth, ultrawide band) and ambient intelligence [13] [41] shifts privacy issues onto a global level.

In this paper- also in the context of a digitalized world - we interpret privacy as the right of individuals to exercise autonomy in controlling their personal data. In order counter privacy-threats resulting from today's information systems, privacy-enhancing technologies such as digital identity managers [Registratiekamer 1995] [16] [26] [23], pseudonymous credentials [10] [8], anonymous communication technology for the internet [19] [35] [6] [9], and the platform for privacy preferences or privacy-protection systems at the enterprise level [25] [2] [42] were developed. All these systems contribute valuable solutions for enhancing privacy.

We are inspired by the tenet of autonomy from Immanuel Kant, and promulgate the informational self-determination database systems. It is time for individuals to regain control over their private data, and that people get control over their virtual shadows, which are spread over a number of information systems in different organizations. We argue that future database systems must provide autonomy with regard to data processing. We will enunciate the key principles for data processing systems that pertain to autonomy in data processing. Our principles are built on current privacy legislations and guidelines, and do not only address technical issues, but also include legal and organizational points. We propose a design based on our principles, identify privacy and security challenges, and suggest some approaches to solving these problems.

Privacy Invasive Technology: an Overview

This chapter examines new and emerging technologies which potentially threaten an individual's informational privacy. These are technologies related to identification, location-based services and ambient intelligence technologies. However this chapter does not claim to cover the topic in full.

Identity-Related Technologies

Identification is the process of establishing the identity of a person [32]. This is achieved by means of a set of characteristics that describe a person. After all, the essential and unique characteristics of an individual are the features which give it an identity. With the ongoing shift towards electronic transactions in both commerce and government, the need for electronic identification of individuals is growing. The term digital identity, however, is as difficult to define concisely as is the concept of human identity. It should be noted that no commonly agreed upon definition can be found in literature. On a very general level it can be said that a digital identity is a machine-readable representation of a human identity which is used in electronic systems for interactions with local or remote machines or people. The purpose of a digital identity is to tie a particular transaction or a set of data in an information system to an identifiable individual, and also to enable access control functionality. With the help of a digital identity, a user can be identified, authenticated and authorised to access a given resource or service. The security of an information system relies to a large extent on the ability to identify and authenticate users [34]. The identity-building process involves unity, permanence, and physical characteristics. Digital identity [12], comprising digitized human characteristics such as identity, behavior, biological features, etc., will in many settings replace today's indicators like, for example, name, telephone number, etc. This new form of identity enables new digital services but at the same time brings new risks. A uniform system to identify users in cyberspace would have dramatic consequences.

To manage and control these many electronic identities that a person may have, identity management systems were developed. A unique access tool manages the many parts of the citizen's online identities. The advantages and disadvantages of identity management systems are discussed within the appendix.

Under biometric technologies we understand the use of physiological or behavioural properties for identification of users [3]. Using biometrics for authentication is itself not new, but that machines are able to process biometrics is a new dimension. This technology, using unique human characteristics such as fingerprints, iris, face, voice and DNA, is the quasi-perfect solution for identification. Some methods, like iris scan and face recognition, are contactless biometrics technologies. Several

biometric measurements can be combined to try to achieve a higher level of protection. In addition, it is very difficult to change biometrics for the user.

From a privacy point of view, biometrics are a threat as they constitute a very strong form of identification. Such a string means of identification may not be necessary in many applications. The biggest problem is that biometric measurements include more data than are needed for an identification. A retina scan may give hints concerning a person's health. DNA samples taken by sniffers would enable a service provider to learn about the user's genetic disposition to illnesses etc. Furthermore, there is also a risk loss of loss: we leave fingerprints almost on everything we touch. There is thus a risk of counterfeiting. Last but not least, it is still debated by scientists and privacy activists which biometrics are really ready for deployment at the current point in time.

Radio frequency identification (RFID) technology is a wireless system for identification.It allows remote non-contact automatic reading of RFID-enabled objects. These objects are built-in 'active' and 'passive' tags. Active tags, powered by an incorporated energy supply, offer a permanent connection and a long distance communication. Passive tags are energized by an antenna emitting radio signals. They just have a short-distance communication, up to about four meters. These tags can be embedded in nearly any object, such as bank notes, clothes or even razor blades, because they are almost invisible.

Future identification technologies are bio-implants and DNA sniffers. A bio-implant is a tiny implanted chip which has communication capability. This could be management of access levels, location data, personalization of the nearby environment, or communication with other chips (e.g. bio-sensors) or with real-time medical systems, for example. Bio-implants can build an 'augmented' human body [40] and can there-fore also be used in creating an identification process. DNA sniffers work on the basis of DNA fingerprints, a far simpler method than DNA sequencing. It can be compared with RFID, because identification also occurs without direct contact. The sniffer correspond to the RFID reader and human cells act as the equivalent of tags. This technology is the leading candidate for future identification systems.

Location Based Services

Location-based services (LBS) is a term that describes services offered to users based on their current location. Providing services based on location implies that a user's position can be determined with a given accuracy. LBS can be deployed in a variety of services ranging from commercial, location-specific content for tourists to services as diverse as health administration or entertainment services.

Wireless communication technologies serve as the basis for providing location-based services. We will briefly describe some wireless technologies in this section, such as cellular systems, wireless local area networks, bluetooth and ultrawide band.

Cellular systems are the most common type. The European standard GSM (Global System for Mobile communications) has become the main mobile system world-wide with about 909 million subscribers across 200 countries (September 2003) [22]. As a 'third generation' standard, UMTS (Universal Mobile Telecommunications System) will succeed GSM and bring broadband services to handsets.

Wireless local area networks (WLAN) is another wireless technology that has a connectivity range of about 100 meters, more commonly known as 'hot spots' (physical locations where WLAN access is provided). It is often used in train stations, airports, city halls, hotels, business centers, university campus, enterprise premises, as well as in private homes.

Bluetooth was developed to replace cables with devices up to a range of about 10 meters, but can be extended to more than 100 meters.

Ultrawide band technology enables the reuse of frequencies already assigned to wireless services and is therefore an alternative to cellular systems.

The geographic coverage is mapped by cells in a cellular system. User equipments run in a specific cell, which can always be determined by the operator. By means of enhanced observed time difference and observed time difference of arrival techniques, measurements from a pair of downlink transmissions, the position of an electronic device can be located with an accuracy of around a hundred meters. Bluetooth and WLAN are able to compute any user location from the position of the fixed access points. WLAN can do this with an accuracy of about 100 meters, and Bluetooth, with the additional possibility of getting their positions from other recently located users, to within about 30 meters. By using the signals transmitted from a satellite constellation, users can compute their

position with the global positioning systems (GPS). Such satellite techniques are accurate up to half a meter. Further advantages of GPS are its global coverage and low impact on existing communication networks; its disadvantage is the signal's weakness indoors. A major impact on the performance of location technologies is achieved via the combination of the different terrestrial and satellite techniques.

Different location computation technologies exist, based on the underlying wireless technology. Some of these services allow users to keep control over their location data. Satellite techniques (Navstar, Glonass, GNSS) are controlled entirely by the user, whereas terrestrial techniques (Cell id, observed time difference, Bluetooth and WLAN) are normally processed by the operators. Privacy problems arise as operators can determine a user's position without the user's consent. Wireless technologies may even enable network operatos to seamlessly track an individual throughout their network. It is clear that such location data is highly privacy-sensitive and also valuable for providers of commercial services.

Ambient Intelligence Space

Ambient intelligence and virtual residence seemed once a vision of the future that is by now – at least in part – becoming reality. Humans will be surrounded by intelligent interfaces supported by computing and networking technology which will be embedded in everything. Smart objects, such as smart paper, smart roads, smart furniture, smart clothes etc. will be ubiquitous and always "on". Smart devices will be able to interact with each other and with the environment. These devices will become increasingly smaller and cheaper and will be able to sense, think and communicate [41]. Several terms reflect this vision: ubiquitous computing, pervasive computing, disappearing computing, proactive computing, sentient computing, affective computing, wearable computing and ambient intelligence. The term 'ubiquitous computing' was coined by [41] "the most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it". This vision carries with it a high risk of losing one's privacy. Consequently, the Information Society Advisory Group has stressed the importance of giving control over ambient intelligence services and interfaces to ordinary people. Ambient intelligence is a vision which would has tremendous social implications. The academia and industry investments in research and technological development within this field are enormous.

Privacy Issues

The abovementioned technologies certainly raise a number of privacy concerns. Some noteworthy points about these technology are discussed in this part of the paper.

Biometric data are sensitive and of a personal nature. Therefore, even if forbidden by law, the risk of disclosure to a third party is given. Biometric data fully identify a person and provide additional and sensitive information. Medical specificity can be found in fingerprints, iris image, and retina scan, for example. A further danger is that some biometric measurements can be taken without physical contact between person and sensor. For instance, face recognition can be performed without consent by the concerned individual. It therefore poses a more serious threat for privacy, as sensors can be hidden in the local environment.

RFID tags can be accessed as well in a contact-less manner. Therefore RFID tags raise specific privacy concerns such as user awareness and empowerment. RFID tags represent a reliable form of identification as soon as the tags can be linked to the owner of an object.

In the near future, cellular system and WLAN technologies will bring mobile broadband services. Wireless communication has the potential to raise privacy concerns, especially regarding location data. Negative consequences may arise for users when databases of network operators are mined. Different parties are involved in the value chain of location-based services, therefore there is an even higher risk for the protection of privacy.

Monitoring and surveillance capabilities, using ambient intelligence, will emerge on a large scale. This kind of technology constantly detects and monitors what people are doing, both offline and online. Some argue that this represents the end of privacy [17]. "The right to be left alone" [39] would not exist any more. Furthermore these technologies create the opportunity to 'cross the border' [29]. Crossing borders usually implies a privacy-invaded feeling. These borders have natural (walls, doors, clothes), social (family, doctors, judges, work colleagues), spatial , temporal, (different parts of a persons life is conveyed to different target groups), ephemeral or transitory (information may get lost and have to be deleted) characteristics. Ambient intelligence makes the crossing of these borders

easier and more likely, even though the borders are always fluid, relative, multidimensional, and dependant on context, culture and personal preferences. This new world of interconnected objects creating smart environments could become an Orwellian nightmare without privacy, data protection laws, organizations and technology [30]. The 'smart home' [20] and 'virtual residence'[5] concepts are just two examples of visions within this field.

Security Issues

Privacy concerns are deeply intertwined with security issues. IT security in general comprises measures both at a technical and an organizational level to achieve the generic security goals of confidentiality, integrity, availability, accountability and authenticity [14]. Security consequently has to be seen as a prerequisite for enforcing data protection. It is a 'conditio sine qua non' for informational privacy. Notwithstanding, issues of data security constitute only a small part of the considerations comprised in the field of informational privacy.

We will discuss selected security issues that are related to the privacy-invasive techniques described above.

Cellular systems from the 2G digital network communication are rather insecure while 3G security features tend to be more efficient. The GSM encryption is fairly easy to break and the lack of strong security in GSM cellular networks allows for a wide range of fraud [33]. WLAN security is even less efficient. 802.1x, 802.11b and 802,1i standards offer strong authentication between access points and wireless LAN cards. Wired equivalent privacy (WEP), dynamic WEP and WiFi protected access (WPA) provide a better layer of armor against hackers.

Location-based services and the accompanying data, including where and who the user is, can improve security in certain situations (e.g. by making it easier to locate accident victims). The main danger of wireless services is however the increasing surveillance in the information society. The collection of location data is made possible and provides interesting information regarding users habits. This situation leads to data mining, discrimination and surveillance, even if the data is only processed by machines. These data might be stolen and could therefore threaten personal security.

The scale of ambient intelligence, its mobility requirements, its heterogeneity, the complexity of its hardware and software, and its distribution of knowledge and resources increase security concerns in matters of trust and dependability. Paradoxically, ambient intelligence best reflects our real world interactions. This paradigm can be described with attributes such as flexibility, mobility, temporality, context dependency, heterogeneity, decentralization, dynamism and change. Interactions will be based on trust and confidence

Conclusion

Though many benefits are gained from identity-related technologies such as location-based services and ambient intelligence space, the potential dangers of monitoring, surveillance, data searches and mining cannot be ignored. At the very least, protection of citizens from various types of intrusion and law enforcement must be ensured when using these technologies.

Balancing security and privacy in the information society [28] [38] will be a tough task. Respecting somebody's private life has to be weighed up against issues of national security, public safety, economic wellbeing, prevention of disorder and crime, protection of health and rights and freedom of others. It is impossible to make a prediction as to which side the future will lie on, but the risk of losing privacy, the "right to be left alone" [39], "the right to select what personal information about me is known to what people" [42] in the information society is rather high.

From our point of view, citizens will lose their entire privacy if nothing is done against current developments. To strengthen privacy and security, actions on legal, organizational and technical issues are required [27]. These three elements are included in our approach to privacy-enhanced database systems. In the following section we summarize what has already been done in the field of privacy-enhancing technology in order to combat the aforementioned risks within this area.

State-of-the-Art Privacy Enhancing Technology

In this section, we consider the concept of privacy-enhancing technologies. We will discuss the PETs that are available today and illustrate their benefits and shortcomings. We will consider identity management, P3P, digital credentials, anonymisers and privacy-enhanced database systems.

The term Privacy-Enhancing Technology (PET) originated in the midnineties from a study that investigated technological measures to curb the use of identifying data in information systems [Registratiekamer 1995]. Nowadays the term PET is widely used, and refers to technologies which aim to eliminate the use of personal data in information systems or to restore the user's control over the revelation of personal data [7]. In a wider sense, one could say that the term PET represents all technologies which pertain to protecting an individual's privacy.

Identity Management

Identity management aims at giving users of electronic services the power to determine for themselves which data concerning their identity should be disclosed to other parties in the course of an electronic transaction. It intends to restore the power of informational self-determination to the user. For that purpose, an electronic identity manager is installed on the user's machine that assists the user in all electronic transactions. Such a software lets the user create several profiles for transactions on the Web that each contain different amounts of personal data. Furthermore, an identity manager supports the user in the creation and management of pseudonymous identities. Such identities may be realized with the help of pseudonymous credentials.

The identity protector as proposed in [Registratiekamer 1995] was the first proposal for an identity manager. A Web-based identity manager was developed by Bell Laboratories [16]. Identity managers were also proposed on the basis of PDAs (Personal Digital Assistants), which the user can carry along with him at all times [26] [23]. Users conduct all electronic transactions with the help of a PDA, on which the identity manager is installed.

The use of identity management solutions alone is not effective enough to prevent the creation of personal data. Nowadays, most higher value transactions require the disclosure of an identity. In such settings, identity management is hardly efficient. Therefore pseudonymous credentials (see below) must be combined with the approach of identity management to allow for anonymous transactions which provide security to service providers (e.g. by guaranteeing that users who engage in unlawful behavior can be traced). Another problem is that users can't control how their data is processed once they have released it. We see the potential of identity management solutions in the context that they may help users to manage pseudonymous identity while at the same time hiding the complexity of credential systems from them.

[10] introduced pseudonymous digital credentials as a building block for an electronic transaction system which lets users conduct anonymous, unlinkable transactions. Users setup a different pseudonym with every organisation they deal with and conduct all transactions under pseudonyms. Since several pseudonyms of the same user can't be linked, transactions can't be traced beyond organizational boundaries. Users can obtain credentials from organizations which are used to prove statements about the holder and thus serve the purpose of establishing trust. Pseudonymous credentials also incorporate a mechanism to hold users accountable for their actions. This may e.g. be a trusted third party who can divulge the identity behind a pseudonym in case of unlawful behavior.

Pseudonymous credentials can be used to achieve anonymous electronic transactions while maintaining security. Anonymous transactions are clearly the most effective way of avoiding the creation of personal data records. Since statements in credentials can be disclosed selectively, they also pertain to the privacy goal of data minimization [42]. Currently, the most advanced implementation of a pseudonymous credential system is the one by [8].

Although credentials afford users the possibility of anonymous transactions, it has to said that these technologies are rather complex and may be difficult for users to understand. Anecdotal evidence suggests that many users even find the handling of X.509 certificates, which have been around for much longer, rather cumbersome. Some of the complexity of these systems can be hidden from the user via measures taken at the level of interaction design. Identity management solutions can make such systems manageable even for the average user. Deploying such systems at the current point in time may be difficult, as there are not yet any official standards regarding algorithms, key and message formats.

Anonymous communications and transactions in the Internet can only be achieved if the underlying network allows for the creation of anonymous communication channels. Several proxy services exist that afford anonymous Internet communication to users and enable users to surf the Web anonymously: examples include onion routing [19], crowds [35] or the Java Anonymity Proxy (JAP) [6]. Onion routing and JAP make use of the mix

approach, a technique proposed by [9] to enable anonymous untraceable email communications.

There are also tools for anonymous email communication. Such tools enable users to send and receive email under pseudonymous addresses. Two types of systems exist: The first type removes identifying information in a message and forwards it. The second type uses mix networks to anonymize messages. A very well-known remailer service was anon.penet.fi, which was closed down by its owner after Finnish police demanded the disclosure of a user's identity.

On a political level, giving users the possibility to use Internet-based services in a fully anonymous manner is often perceived as a danger to society. Anonymity makes it more difficult to pursue offenders who use the Internet to access illegal content. In the current political climate, it is more difficult than ever to argue for fully anonymous communications in the Internet.

Privacy in Ubiquitous Computing

It can be argued that the vision of pervasive computers in combination with powerful, new sensor technology poses a threat to an individual's privacy. This threat creates a need for technology to counteract the negative effects on privacy that ubiquitous computing environments may bring about. As an example, one might cite RFID tags: once clothes are tagged with RFID-based price labels, it is possible to read the information contained in these labels in a number of situations. It then becomes possible, for example, to bar entrance to clubs or restaurants to people wearing clothes that are more than 12 months old.

However sensors such as DNA sniffers, surveillance cameras or RFID tag readers make it difficult to come up with technological solutions that protect an individual. Unless RFID tags are destroyed, they can be read out. Similarly, contactless smart cards pose a risk of operation without a user's consent. Unless such cards are carried in a steel envelope that shields them from contactless card terminals, access is possible at any time.

Currently, the most promising approach to protecting privacy seems to be an approach that relies on an integration of P3P (platform for privacy preferences) into ubiquitous computing environments. People would then declare their privacy preferences. Service providers would have to read these statements (e.g. via wireless communication) and dynamically react to personal privacy settings. If users do not express consent to data processing in a ubiquitous computing environment, they can have services deactivated. Such an approach would again rely on machine-readable privacy policies such as P3P. However users must trust service providers that their privacy preferences will be respected. Thus, such an approach still requires users to put a fair amount of trust in service providers. Furthermore, it may be difficult to react to every user's privacy preferences in settings where many users are active (e.g. in a public place where users are under constant camera observation). Thus many problems remain to be solved in the area of ubiquitous computing, if privacy is to be maintained in scenarios such as these.

Privacy Enabled Data Processing

The Platform for Privacy Preferences (P3P) is a W3C standard which enables users to inform themselves about a Web site's privacy policy and to discover potential discrepancies with their own privacy preferences. Organizations declare their privacy practices in a machine-readable format which, with the help of a P3P-enabled Web-browser, can be compared with the user's own privacy settings. Depending on the browser's comparison, a user can choose not to visit a site, or to 'opt-in' to or 'opt-out' of a specific use of data.

P3P is useful for warning users about sites that engage in privacy-invasive data processing. It also helps users to discover sites which offer them a higher level of privacy. There has also been some criticism of P3P however: first and foremost, users have no way of telling whether service providers really adhere to the principles stated in P3P policies. Unless sites are audited and certified with regard to policy implementation, users do not know whether policies are really implemented. It is also debatable whether P3P really empowers the user. In many cases, a user does not have the option of selecting a site and will just have to accept the data processing practices of a given site. In the opinion of the authors, P3P won't dramatically change the power balance between organizations and consumers.

Privacy-protection at the enterprise level as well as privacy policies which are published on Web sites are essentially promises made by organizations that they will ad-here to certain data processing practices. Users have no way of verifying whether these promises are kept. The Platform for Enterprise Privacy Policies (E-P3P) is an approach to privacy enable the processing of personal data. Privacy

policies are formalized and are then automatically enforced throughout an enterprise [25].

Users are presented with privacy policies at the time of data collection and can consent to a specific use of data. The consent of the user to a given purpose is stored along with all data items which were collected from a user. Whenever personal data is to be processed for a given purpose, the system consults the policy attached to the data and denies an operation if it is not in line with the practices stated in the policy. This leads to a system that effectively prevents the misuse (including unauthorized disclosure) of personal data.

Such a system can guarantee that a user's data can only be processed in accordance with a published policy – provided the system is administered correctly. However, existing systems need to be modified in order to support this approach.

[2] propose a new category of privacy-enhanced database systems called 'Hippocratic Databases'; these include responsibility for the privacy of data as a central design goal. The name is inspired by the Hippocratic Oath, which has guided the professional conduct of physicians for centuries. The founding principles for these databases mostly stem from privacy legislation and guidelines, such as the Fair Informational Practices [42].

When data is collected, users express consent to the processing of specific data items for a specific purpose. A Hippocratic Database keeps privacy metadata which records for every data item: the agreed processing purpose, external recipients (if any), authorized users and the retention period. Based on this metadata, the system checks every query and only executes queries that are compatible with these policies. Further components include a data retention manager that deletes data when no longer needed and a query intrusion detector that flags suspicious queries based on heuristics.

Conclusion and motivation for autonomic databases

Many privacy enhancing technologies aim to allow anonymous transactions and anonymous communication in the Internet. While this is clearly the most effective approach to avoid the creation of personal data, it remains to be seen whether service providers are willing to embrace these technologies. The approach of enterprise-level privacy policies promises to guarantee that enterprises do indeed process data according to their declared policy.

We propose autonomic databases as a technology that complements existing privacy-enhancing technologies. The approach is different from existing technologies. Autonomic databases are intended for settings in which personal data is processed and in which an individual's identity is stored in the database. We perceive that transactions should be conducted anonymous wherever possible and perceive pseudonymous credentials as the most effective technological means to support a migration towards anonymous transactions.

The approach of autonomic databases further develops existing approaches to privacy-enabled data processing. We envision a data processing system that guarantees by technological measures that data is processed in line with polices. The approach of autonomic databases has this characteristic in common with the approach of [Agarwal 2002] and also with the approach of [25] for privacy in data processing at the enterprise level. However, we see further need to tailor data processing to the needs of the individual if privacy is to be maintained.

Our approach comprises new legislative measures to complement the existing legal framework on data protection. On the one hand, we propose a differentiation between personal data and transactional data. Individuals are to be given full access to personal data, but not to transactional data (which is thought to be owned by the company rather than by the individual). Furthermore, we aim to bring more transparency to data processing: through a structure of portal services, an individual can monitor data processing in two ways: First, an individual can view all personal data that is stored about him or her, and second, for every data item, an organization must state how this item of personal data was acquired.

The portal aggregates views on all organization who store data concerning this individual. Through the use of a portal service, individuals do not have to manage accounts with several organizations who store data about that individual. Instead, all data can be accessed through a single point of entry.

The approach of autonomic databases thus aims to give users more control in wettings where identified transactions take place. In the next section, the design goals for such a system are stated.

The Concept of Privacy Revisited

More than 100 years ago, Warren and Brandeis wrote the landmark paper 'The Right to Privacy', published in the Harvard Law Review in 1890 [39]. They defined privacy as 'the right to be let alone' and argued that legislation should give this right to every individual: "Political, social, and economic changes entail the recognition of new rights". In the twentieth century, many legal scholars and philosophers have attempted to define the concept of privacy [21]. However, it is impossible to come up with a universally valid definition of privacy as the concept depends on social aspects, cultural values and the legal framework. The issues of privacy are "fundamentally matters of values, interests and power" [18].

An implication of privacy as an interest, is that it has to be balanced against other competing interests. People´s interest in their own privacy may conflict with the interests of other people or organizations [15]. The concept of privacy does not apply to mere information only. Privacy rights have a long tradition and are implemented in many fields [36]:

- territorial privacy: protects the physical surroundings of a person, i.e. in a domestic or other environment
- bodily privacy: protects the physical integrity of a person against undue interference (e.g. physical searches, DNA testing)
- communication privacy: protects the personal communication of a person against monitoring by other persons or organizations
- informational privacy: the right of a person to control what data about his resp. her person can be gathered, processed and disseminated

In the context of information systems, the consideration of privacy leads naturally to the notion of informational privacy. This restriction makes sense as an information system usually does not affect territorial or bodily privacy (with the exception of robotics applications or some ubiquitous computing devices, which are outside the scope of this paper).

A very common and well-accepted definition of informational privacy is the one given by Alan Westin in his classical work on privacy. Westin defines informational privacy as

'...the claim of individuals, groups and institutions to determine for themselves when, how and to what extent information about them is communicated to others' [42].

At the heart of the notion of informational privacy lies the understanding that certain information about a person is not public but rather private, however it is not possible to give a precise definition as to which data falls into which category. Such a notion depends on cultural understanding and personal views. Informational privacy is, just like other forms of privacy, the interest of an individual and that may compete with the interests of other parties.

With the wide-spread use of information systems, the focus on privacy shifts towards an understanding of privacy as the right to informational self-determination. An individual should have the right to control the release and dissemination of personal data as well as the context the data is going to be used in, to the greatest possible extent. In addition to Alan Westin's definition of informational privacy, we state that in general informational privacy and the measures to protect it should address

- the release and dissemination of personal data
- the right to remain unidentified (anonymous) when we choose to
- the protection of highly sensitive data in electronic systems (see 5.1.3)
- the latent danger of tracking and logging of users and their activities
- the right to be let alone when we choose to be let alone
- the right to live without the threat of constant surveillance by electronic means

We claim that the advent of new technologies poses a threat to the citizen's privacy. The fact that computers are becoming ubiquitous - and that information technology is becoming more and more a part of our daily lives – leads to an erosion of informational privacy. An awareness for privacy problems must therefore be created urgently. We maintain that any technology that can enhance privacy is thus worth discussing. We see our paper as a contribution to the discussion on privacy issues and aim to point out new directions in which technology and legal frameworks may be developed in order to work towards offsetting the negative effects that information technology has on privacy.

The next section explores the tenets of participation and transparency. Transparency and participation are considered in the context of the data protection tradition. They will be discussed in the context of private public sector data processing. We consider how these two principles are implemented by our architecture and explain why the architecture leads to both more transparency and better participation as compared to most of today's data processing systems.

Motivation for Transparency and Participation

The interest in informational privacy increased in the 1960's and 1970's due to the wide-spread use of information technology. Legislative bodies began addressing the problem in the 1970's. The first modern data protection act was adopted by the German State of Hesse, the first national law by Sweden in 1974. A very influential piece of data protection legislation is the US Privacy Act. The act was passed by the Congress in 1974, thereby acknowledging that the rapid development of information systems posed a threat to personal privacy. Although the Data Protection Act was not very successful in the US, it found much attention abroad. This resulted in the fact that many elements of this policy can be found in data protection laws of other countries.

The US privacy act was crafted after the work of an advisory committee, which established the notion of 'fair informational practices'. This concept turned out to be very influential in shaping data protection legislation around the world. These practices are based on work by Alan Westin. Westin stated 8 important principles for fair information processing [42]. These principles are also incorporated into the OECD guideline on data protection of 1980 [31] and the EU directive 95/46/EC on the protection of individuals with regard to the processing of personal data (EC95 1995).

One of those principles is the principle of openness and transparency. It states that there should be a general policy of openness about collections of personal data. Especially, there should be no secret data collections. Means of establishing the existence and nature of collections, the main purposes of their use as well as the identity of the data controller should generally be known. Another important principle is the principle of individual participation: an individual should have the right to request information from a controller as to whether a collection contains data about them. Requests should be answered within a reasonable period of time and at a reasonable price. Furthermore, individuals should have the right to have records rectified, completed or erased where appropriate (i.e. in the case of incorrect or illegally stored data).

Transparency in e-Commerce Data Processing

Various surveys have shown that privacy is a substantial concern on the Internet, particularly in e-commerce transactions [1]. Users are obliged to divulge personal data in almost every transaction, and in so doing, leave traces each time such a transaction is carried out. In most business relationships, users have neither insight into what data the other party collects nor do they have access to these data.

For e-commerce purposes, P3P is slowly gaining in popularity. This standard, how-ever, only addresses privacy declarations. The use of P3P does not lead to any form of participation or to a much enhanced transparency. There are very few companies who allow users to see ther personal data and to control how this data is to be used. An approach such as EPAL is therefore a step in the right direction: such technologies ensure that data is processed in accordance with specified polices. However EPAL does not lead to a heightened user participation in data processing.

We thus conclude that in the domain of e-commerce, participation and transparency in data collections is the exception rather than the rule. An approach such as the one presented in this paper can help to make data collections more transparent and to give users more participative power in data collections. We propose that portals should be operated that give individuals access to the audit data that is stored about them and thus increase both transparency and participation.

Transparency in e-Government Data Processing

Informational privacy is an especially important issue in e-government. The data that are processed in e-government environments are often of a much more sensitive nature than the data processed in the domain of electronic business [24]. People are increasingly concerned about privacy issues related to e-government, and tend to feel the same way about citizen cards [11], [4]. Although information and communication technology provide tremendous opportunities for reshaping the relationship between government and stakeholders and creating more efficiency in bureaucratic systems, it also creates significant security and privacy challenges.

Data in governmental databases contain highly sensitive data such as social security numbers, information related to individual taxation, data concerning religious beliefs, criminal records, demographic information and medical records. Furthermore, governmental bodies process high volumes of data. They are empowered by public law to collect data on citizens and can enforce their right to do so. Governments thus have the potential to accumulate large data collections, which may create potential conflicts with the citizen's interest in informational privacy [37]. Given these facts, it is even more desirable that citizens know what data administration keeps about them.

Administrative cultures and procedures in Europe vary, and so do the views on the sensitivity of data. The religious affiliation is considered a very sensitive issue in the Netherlands and in Greece, while inhabitants of Finland are very sensitive about data that relates to the gender of a person. Many other examples can be found illustrating the differences that exist with regard to the sensitivity of data.

We feel that there is still a general lack of transparency and participation in governmental data collections. In most European countries, citizens do not have the right to access their own data in governmental data collections. An exception is the country of Sweden: all data that is collected by the state is deemed public. As a consequence, any citizen has the right to see e.g. their neighbour's tax declaration. Another fairly advanced country (with regard to participation) is the Netherlands: here it is currently being discussed if citizens should have access to their own data in all governmental data collections. In most European countries though, citizens do not automatically get access to their own public records.

Conclusion

With the wide-spread use of information systems, the focus on privacy shifts towards an understanding of privacy as the right to informational self-determination. An individual should have the right to control the release and dissemination of personal data as well as the context the data is going to be used in, to the greatest possible extent.

1 Ackermann, M. S., L. F. Cranor, et al. Privacy in e-Commerce: Examining User Scenarios and Privacy Preferences.In S. Feldman and M. Wellmann, Proceedings of the 1st ACM conference on Electronic commerce. Denver, Colorado, ACM Press. pp., 1999

2 Agrawal, R., J. Kiernan, et al. Hippocratic Databases.In Proceedings of VLDB 2002. Hong Kong, China, Morgan Kaufmann, San Francisco. pp. 143-154, 2002

3 [Ashbourn 200]

4 BBC News 2003. Public oppose ID card scheme (July 2003). Available from: http://news.bbc.co.uk/2/hi/technology/3004376.stm. Last Visited: November. 2003

5 Benson, D. Distributed Identities: Managing privacy in pervasive computing. Explored View-points. SRI Consulting Business Intelligence.2003

6 Berthold, O., H. Federrath, et al. Web MIXes: A System for Anonymous and Unobservable Internet Access.In H. Federrath, Designing Privacy Enhancing Technologies. International Workshop on Design Issues in Anonymity and Unobservability. Springer, 2009. pp. 113-129, 2001

7 Burkert, H. Privacy-Enhancing Technologies: Typology, Critique, Vision.In P. Agre and M. Rotenberg, Technology and Privacy - The new Landscape. MIT Press. pp., 1997

8 Camenisch, J. and A. Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation.In B. Pfitzmann, Advances in Cryptology - EUROCRYPT 2001, International Conference on the Theory and Application of Cryptographic Techniques, Innsbruck, Proceedings. Berlin, Springer, 2045. pp. 93-118, 2001

9 Chaum, D. L. "Untraceable Electronic Mail, Return Addresses and Digital Pseudonyms." Communications of the ACM 24(2): 84-88.1981

10 Chaum, D. "Security without identification: Transaction systems to make Big Brother obsolete." Communications of the ACM 28(10): 1030-1044.1985

11 Campbell, R. D. Collins English Dictionary. Harper Collins, Glasgow, 1998

12 Covell, P., S. Gordon, et al. Digital Identity in Cyberspace.In Proceedings of Conference on Legal/Technical Architectures of Cyberspace. Harvard Law School, Cambridge, Massachusetts, USA. pp., 1998

13 Ducatel, K., M. Bogdanowicz, et al. Scenarios for Ambient Intelligence in 2010 - Final Report. IPTS Publications.2001

14 Eckert, C. IT-Sicherheit: Konezpte, Verfahren, Protokolle. Oldenbourg, München, 2001

15 [Etzioni 1999]

16 Gabber, E., P. B. Gibbons, et al. How to Make Personalized Web Browising Simple, Secure, and Anonymous.In R. Hirschfel, Proceedings of First International Conference on Financial Cryptography. Springer, 1318. pp. 17-32, 1997

17 Garfinkel, S. Database Nation, The Death of Privacy in the 21st Century. O'Reilly, 2001

18 Gellman, R. Does Privacy Work?In P. Agre and M. Rothenberg, Technology and Privacy: The New Landscape. Cambdridge, Massachussetts, MIT Press. pp. 193-218, 1998

19 Goldschlag, D., M. Reed, et al. "Onion Routing." Communications of the ACM 42(2).1999

20 Gooley, C. and T. Saponas. Privacy issues of the Aware Home. Paper on the Georgia Tech Aware Home project.2003

21 Gormley, K. "One Hundred Years of Privacy." Wisconsin Law Review (1335).1992

22 GSMWorld. Available from: http://www.gsmworld.com/news/press_2003/press_25.shtml. Last Visited: April 2004

23 Jendricke, U. Sichere Kommunikation zum Schutz der Privatsphäre durch Identitätsmanagement. Berlin, 2003

24 Joshi, J., A. Ghafoor, et al. Security and Privacy Challenges of Digital Government.In W. J. McIver and A. K. Elmagarmid, Advances in Digital Government. Dordrecht, Kluwer Academic Publishers. pp. 121-136, 2002

25 Karjoth, G., M. Schunter, et al. Platform for Enterprise Privacy Practices: Privacy-enabled Management of Customer Data.In R. Dingledine and P. Syverson, Privacy Enhancing Technologies. Springer. pp., 2003

26 Köhntopp, M. Generisches Identitätsmanagement im Endgerät.In R. Grimm and A. Röhm, GI Workshop Sicherheit und Electronic Commerce - WSSEC 2000'. Bonn, Köllen. pp., 2000

27 Lessig, L. Code and Other Laws of Cyberspace. Basic Books, 1999

28 Maghiros, I., C. Centeno, et al. Security and Privacy for the Citizen in the Post-September 11 Digital Age: A Prospective Overview. IPTS Publications.2003

29 Marx, G. T. "Murky conceptual waters: the Public and the Private." Ethics and Information Technology 3(3): 157-169.2001

30 Mattern, F. Ubiquitous Computing: Scenarios for an informatized world. ETH Zurich.2003

31 OECD Organisation for Economic Cooperation and Development. Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.1980

32 Oxford OED, U. P. The Oxford English Dictionary, 2nd Edition. Oxford, University Press, Oxford, 1989

33 Partnership, T. G. Technical Specification Group Services and System Aspects 3G Security; Security Principles and Objectives (3G TS 33.120 version 3.0.0).1999

34 Pfleeger, C. Security in Computing, 2nd Edition. Prentice Hall, Upper Saddle River, 1996

[Registratiekamer 1995] Registratiekamer, T. N. Privacy-Enhancing Technologies: The Path to Anonymity.In., 2002. pp., 1995

35 Reiter, M. K. and A. D. Rubin. "Crowds: Anonymity for Web Transactions." ACM Transactions on Information and System Security 1(1): 66-92.1998

36 Rosenberg, R. The Social Impact of Computers. Academic Press, San Diego, 1992

37 Schweizer, R. and H. Burkert. Verwaltungsinformationsrecht.In H. Koller, G. Müller, R. Rhinow and U. Zimmerli. Basel, Helbling & Lichtenhahn. pp., 1996

38 Walters, G. J. "Privacy and security: an ethical analysis." ACM SIGCAS Computers and Society 31(2).2001

39 Warren, S. D. and L. D. Brandeis. "The Right to Privacy." Harvard Law Review 193(4).1890

40 Warwick, K. Identity and Privacy Issues raised by Biomedical Implants. IPTS Publications.2002

41 Weiser, M. "The computer of the 21st century." Scientific American: 94-101.1991

42 Westin, A. F. Privacy and Freedom. Atheneum, New York, 1967