Top 10 Mistakes in System Design from a Privacy Perspective and Privacy Protection Goals

Marit Hansen

▶ **To cite this version:**

Marit Hansen. Top 10 Mistakes in System Design from a Privacy Perspective and Privacy Protection Goals. 7th PrimeLife International Summer School (PRIMELIFE), Sep 2011, Trento, Italy. pp.14-31, 10.1007/978-3-642-31668-5_2 . hal-01517612

HAL Id: hal-01517612
https://inria.hal.science/hal-01517612

Submitted on 3 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Top 10 Mistakes in System Design from a Privacy Perspective and Privacy Protection Goals

Marit Hansen,

Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein,
Holstenstr. 98, 24103 Kiel, Germany
marit.hansen@datenschutzzentrum.de

Abstract. Privacy requirements are often not well considered in system design. The objective of this paper is to help interested system designers in three ways: First, it is discussed how "privacy" should be understood when designing systems that take into account the protection of individuals' rights and their private spheres. Here specifically the concept of linkage control as an essence of privacy is introduced. Second, the paper presents a list of ten issues in system design collected during the daily work of a Data Protection Authority. Some of the mistakes are based on today's design of data processing systems; some belong to typical attitudes or mindsets of various disciplines dealing with system design (technology, law, economics and others). Third, it is explained how working with protection goals can improve system design: In addition to the well-known information security protection goals, namely confidentiality, integrity and availability, three complementing privacy protection goals – unlinkability, transparency and intervenability – are proposed.

Keywords: Privacy, Privacy Mistakes, System Design, Privacy Protection Goal, Unlinkability, Transparency, Intervenability.

1 Introduction

IT security consultants have been publishing information on typical security mistakes for a long time. From these mistakes, organizations and individuals can learn, and thereby they may avoid repeating the same mistakes all over again. Several of the mistakes might reside in human nature or in the professional socialization, for some mistakes poor design of data processing systems may be accounted. The same is true for "privacy mistakes", or to narrow it down: mistakes in system design from a privacy perspective.

The findings of this paper are derived from the experiences of the author after having worked for more than 15 years in a Data Protection Authority. Being a computer scientist herself, the author has collaborated with people from various disciplines and thereby identified some typical attitudes or mindsets of system designers that may explain the vulnerability for various mistakes and other wrong-doings, be it intentional or not. The collection of Top 10 mistakes have been presented first at the IFIP Summer School 2011 on privacy and identity management, thereafter the list has been

further discussed and developed. Cases illustrating the mistakes have been used for education and sensibilization purposes. Moreover, they support the work with privacy-specific protection goals that complement the well-known IT security protection goals confidentiality, integrity and availability.

The text is organized as follows: Section 2 describes the notion of privacy that is used in this text and gives an overview of general principles. The Top 10 mistakes in system design are presented in Section 3. Section 4 focuses on protection goals, how to use them for system design and why it makes sense to extend the widely employed set by the privacy-specific protection goals unlinkability, transparency and intervenability. Finally, Section 5 summarizes the results and gives an outlook.

2 Privacy Perspective

The terms “privacy” and “data protection” are very often used with varying meanings [1]. This paper does not try to give the one-and-only definition. It is motivated by the author’s experiences working for a German Data Protection Authority whose role is defined by data protection law. However, this does not cover all possible influences to the private sphere or the personality rights of a human being. When discussing system design, the view should be broadened – this is meant by “privacy perspective” in this text.

This section firstly introduces basic definitions (2.1), lists relevant data protection principles (2.2) and then widens the scope to achieve an understanding of an extended notion of privacy (2.3).

2.1 Classical Definitions of “Privacy” and “Data Protection”

Although privacy properties have been playing a role since the ancient times, famous definitions are much younger. Warren and Brandeis stated an individual’s “right to privacy” in the meaning of a “right to be let alone” in 1890 [2]. It is often overlooked that their work must not be reduced to this single statement; in fact, they already considered balancing between private and public interests and touched upon the relevance of context. Westin provided the following definition of a right to privacy in 1967: “Individuals, groups, or institutions have the right to control, edit, manage, and delete information about them and decide when, how, and to what extent that information is communicated to others.”[3] Note that this definition is not restricted to an individual, but also includes groups and institutions. Again with a focus on individuals, a similar definition – not of privacy, but of a so-called right to informational self-determination – stems from the ruling of the German Federal Constitutional Court on the 1983 census and demands that each person can ascertain at any time who knows what about him or her [4].

The census decision has become an important cornerstone for data protection in Germany and beyond. A great number of legal norms have been created to specifically regulate processing of personal data. “Data protection” was chosen as the term that should not only express the IT security notion of ensuring the data’s confidential-

ity or integrity, but its objective should be “to protect the individual against infringement of his/her personality right as the result of the handling of his/her personal data” (§ 1(1) of the German Federal Data Protection Act). The pivotal elements for data protection are “personal data” of an individual, the “data subject”, which are handled by a “data controller” (determining the purposes and means of the processing) and processed by a “data processor” (on the controller’s behalf).

2.2 Data Protection Principles

The following seven principles (further developed from [5]) show the main characteristics of data protection. Note that the first principle includes already the possibility of exceptions to the other principles. In this case the exceptions have to be laid down in statutory provisions.

1. **Lawfulness:** Processing of personal data is lawful only if a statutory provision permits it or if the data subject has consented.
2. **Consent:** Consent means a freely given specific, informed and explicit indication of the data subject’s wish.
3. **Purpose Binding:** Personal data obtained for one purpose must not be processed for other purposes.
4. **Necessity and Data Minimization:** Only personal data necessary for the respective purpose may be processed. Personal data must be erased as soon as they are not needed anymore.
5. **Transparency and Data Subject Rights:** Collection and use of personal data has to be transparent for data subjects. Data subjects have rights to access and rectification as well as (constrained) to blocking and erasure of their personal data.
6. **Data Security:** Unauthorized access to personal data must be prevented by technical and organizational safeguards.
7. **Audit and Control:** Internal and external auditing and controlling of the data processing is a necessity.

2.3 Extended Notion of Privacy

On the one hand, the legally specified obligations concerning data protection provide a good instrument to work with – on the other hand, they still offer a lot of room for interpretation (e.g., what data are really necessary?), and many aspects of a protection against infringement of the private sphere are not tackled (e.g., related to profiling and derived decisions that affect individuals or groups). While Westin’s definition already contained the idea of a group’s right to privacy [3], this has not been widely discussed. Also newer approaches that address already privacy-enhancing system design and thereby widen the idea of privacy, have hardly gained any practical effect [6].

For clarifying the scope of system design, several typical phases in enriching information on a person or a group of persons have been identified and discussed [7], as illustrated in Fig. 1.

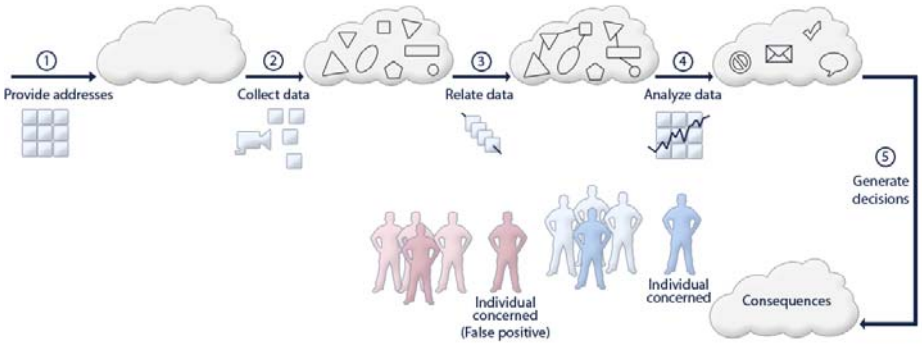

Fig. 1. Enriching information with effects on individuals.

Fig. 1 shows the following: data – in the beginning not necessarily personal data – may be observed and collected by various parties. They can be related with other data, this information can be analyzed, and on this basis decisions can be generated which lead to consequences for single individuals or for groups of people. It is not guaranteed that the data and the aggregation methods are accurate, and even on the basis of correct information the decisions may be false or affect the wrong people. In any case it may be difficult for all persons whose private sphere is concerned to find out what exactly went wrong, who is to be held responsible and how to achieve remedy.

Having this setting in mind, it has been proposed to establish the paradigm of "linkage control" as the key element of privacy [8]. Linkage control would rely on three components: unlinkability when possible and desired, transparency on possible and actual linkages, and the feasibility for data subjects to exercise control or at least intervene in the processing of data.

This leads to an extended notion of what "privacy perspective" means in this text: the protection of individuals (single and in groups) against infringement of their private spheres and their personality, in particular as the result of handling of their data.

3 Top 10 Mistakes in System Design From a Privacy Perspective

Almost each year computer magazines and blogs publish lists and reports like "Top 10 Security Mistakes". Sometimes these lists have a quite general content ("Trusting people", top 1 mistake in [9]), sometimes they go into detail ("The not-so-subtle Post-it Note", top 1 mistake in [10]). Some of the publications are mainly related to technical issues, again with a more general flavor ("Connecting systems to the Internet before hardening them", top 1 mistake in [11]) or being more specific ("Sending sensitive data in unencrypted email", top 1 mistake in [12]).

Of course, mistakes in IT security are relevant for privacy issues, too. But there are many mistakes that should be pointed out from a specific privacy perspective. This

chapter lists ten mistakes in system design that have been put together having in mind years of experience working for the Data Protection Authority of Schleswig-Holstein, Germany (Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein, ULD). These mistakes cover all kinds of problematic properties or wrongdoings – whether intentional or unintentional. Each of the listed mistakes has appeared several times in various circumstances, but neither is there proof that the list is complete nor should the order of mistakes be overrated. The author is not aware of a truly scientific evaluation of mistakes or wrongdoings in system design from a privacy perspective.

Note that term "system design" mainly addresses the development of information technology systems, but in several cases the observations are also valid for the design of political or societal systems.

3.1 Mistake 1: Storage as Default

Storing data is a precondition for all kinds of data processing. At least the data have to be stored for a short time, e.g., in the internal memory of the data processing system. Volatile memory typically requires power to maintain the stored information, e.g., most parts of the random access memory (RAM), whereas for long-term persistent storage non-volatile memory is needed, e.g., hard disks. Sometimes, by so-called swapping, the data held in the RAM is temporarily put in a dedicated section on a hard disk. Similarly temporary files are often stored in various locations, among others in caches. Erasing data from non-volatile memory is frequently not implemented in a way that all parts are gone; in many cases data can be reconstructed after having been "removed". In addition, standard data transfer protocols make use of various (often dynamically chosen) hops as intermediary stations. Again, this requires storage for some time.

All these facts show that storage to some extent is a technical necessity which has to be considered when assessing privacy risks. Further, development of IT systems involves storing data for functionality tests or for debugging, and even later it is convenient to track errors if there are informative logfiles. Finally, the comparably low costs for memory while increasing the storage capacity have led to an attitude of "You never know when you're going to need it. So better keep it."

From the privacy perspective, it is difficult to evaluate IT systems because there are numerous possibilities for storing or moving the data. The data may often reside on a plenitude of IT systems with individual providers. All the same there is no guarantee that the data will be effectively erased as soon as they are not necessary any more. In particular, temporary files and logfiles are regularly neglected when assessing privacy risks.

On a higher level, this has been discussed as the "virtue of forgetting" [13], and even the proposal for a European General Data Protection Regulation [14] foresees a right to be forgotten and to erasure (Art. 17) which extends the scope of erasure according to the current Data Protection Directive 95/46/EC.

3.2 Mistake 2: Linkability as Default

For data processing, it is easier to address objects by specific identifiers, and generally this means to assign unique identifiers that enable the linkage "identifier – object" and "identifier – same identifier". When designing relational databases, the so-called database normalization of fields and tables aims at minimizing redundancy. If, e.g., the postal address of a person changes, this should be entered into the database only once and it should be immediately valid for all instances of the data representing the person. This makes sense for keeping information up-to-date which can be good from a privacy perspective.

However, pushing the idea further would lead to a central world-wide database of all subjects and objects where different parties would get different access rights. From a privacy perspective this would be a nightmare because of the mass of linkable data. In particular the combination of data would make separation of powers difficult, and also the principle of purpose binding could hardly be realized. Therefore data minimization is based on unlinkability as far as possible [15], and linkage control for data subjects are key for their privacy [7][8].

Several examples have shown that database entries that have been pseudonymized (in order to remove the relation to the data subject) often can be linked to the right persons, e.g., in the cases of the published logfiles of the AOL search engine with pseudonymized IP addresses [16] and of the Netflix Prize Dataset with pseudonymized movie rental information that could be linked to public background information from the Internet Movie Database [17].

3.3 Mistake 3: Real Name as Default

Developing further the thoughts on linkability, a special case is the real name policy of many services, among others Facebook and Google in their social networks. These providers as well as various politicians who discuss the topic consider it suspicious if users prefer to act under one or more pseudonyms [18]. The role of the real name is treated differently across cultures [7]. But even outside the online world – in the "real world – the use of nicknames that may differ with different peer groups is socially acceptable. Very often it is not necessary to state one's name or to prove one's identity by showing an official ID document.

In the online world it is so difficult to be really anonymous, and linkable data trails are hard to prevent. So from a privacy perspective it should be the tradition for the online world to use not only a few, but a great number of pseudonyms and be anonymous whenever possible. With technologies such as private credentials, anonymity and accountability can be achieved at the same time, so the often debated lack of accountability with pseudonyms is not a valid argument.

However, not many system designers consider pseudonyms, and even if the state in their privacy policy that pseudonyms are accepted, this is not always reflected in their forms and database schemas that contain a mandatory first name and last name.

3.4 Mistake 4: Function Creep as Feature

"Function creep" means a widening of the data processing beyond the original purpose or context. This violates the principle of purpose binding and can pose risks to privacy that have to be considered when assessing the system [19]. However, computer science aims at re-using code developed once. The art of programming usually offers many degrees of freedom in adapting IT systems for new contexts. As a related matter, interoperability is highly appreciated. Typically the scope of data processing is not limited to one or very few purposes only. Economists are even more trained to exploit available data for multi-purpose usage. Here often context-spanning identifiers are being assigned so that new usage possibilities of the data, thereby linkable across contexts, can be created later. Here function creep is not regarded as a bug, but as a feature.

Function creep is very much related to de-contextualization, i.e., data are taken out of the original context which can lead to wrong conclusions when interpreting the data. Instead, the principle of purpose binding and the objective of contextual integrity [20][21] should be taken serious.

3.5 Mistake 5: Fuzzy or Incomplete Information as Default

From a privacy perspective, accurate and complete information on the planned and performed data processing is a necessity: Data controllers and data processors have to know how their IT systems and organizational procedures work, and this information is required when asking data subjects for consent or being asked by supervisory authorities.

However, most people and organizations do not want to commit themselves more than really necessary. Privacy policies are a good example: Usually they are not drafted by the technology department, but by lawyers or marketing people who do not know the exact details of the data processing. In general, IT people do not tend to invest much time in documenting or explaining their work because developing new things is more interesting. Further, it makes sense to be a bit fuzzy because then there is no need to create a new version of the privacy policy in case of small changes in the IT systems or organizational procedures. Being exact and complete also may lead to long texts that are not very attractive to read.

Some keywords easily show fuzziness in privacy policies, e.g., "including, but not limited to" when discussing data types, business partners to transfer the data or purposes of data processing, similar "such as" or "for <xy> purposes or otherwise". Much harder to detect is when an organization has omitted specific issues in its privacy policy. Very often the statements are unclear and can be misleading [22].

Sloppy system descriptions and unclear responsibilities bear further risks to privacy. Sometimes the data processor does not provide the exact documentation by default, unless extra charges are being paid. Even when there is a contract between the data controller and the data processor, this does not guarantee that the necessary information is provided without extra costs; this has happened for rule sets of Internet firewalls that were hosted on behalf of the data controller.

3.6 Mistake 6: "Location Does Not Matter"

In principle, technology such as dynamic routing on the Internet or the dynamic assignment of resources in cloud computing offers a dissociation from the location where the data processing takes place. In short: "location does not matter."

However, location does matter in law. Very often the jurisdiction is determined by the location of an action or the place of business when a service is being provided. Since there is no common world-wide valid and accepted law, the location of data processing is definitely relevant. This is not only true for data protection law, but for all kinds of access to the data that is stored or transmitted in a country.

At least for the last ten years U.S. American intelligence officials have warned about possible intelligence and military consequences because Internet data more and more bypass the U.S. [23]. In 2011 it became widely known that U.S. companies and their subsidiaries have to comply with U.S. government requests concerning data under the control by the companies – even if the data are stored in Europe. Legal grounds are, among others, the Patriot Act and the Foreign Intelligence Surveillance Act (FISA) [24][25]. Note that in addition to the United States there are several countries that have similar legally based access rights, most outside Europe. But also within Europe there are examples, in particular the Swedish FRA law ("Försvarets radioanstalt lagen") that entitles the government agency FRA to intercept all Internet communication that crosses Swedish borders. Summarizing, for all kinds of risk assessment, location does matter from a privacy perspective.

3.7 Mistake 7: No Lifecycle Assessment

Many problems occur because the system design did not consider the full lifecycle of the data, the organization or the system itself [26]. For instance, data are created without equally planning how and under which conditions to remove them later on. Further, often there are no plans for emergency management, e.g., in case an incident happens. One reason for that is that it seems to be more important to provide a quick and dirty solution, use the momentum, be early on the market and create precedents than to plan ahead and develop a proper solution from the beginning. Even if the developers have intended to clean their system, "quick & dirty" often survives.

Related are lock-in effects where data portability is not offered: If a user has not foreseen an exit strategy, it may be hard to change a provider because there are already established dependencies. Or if the provider can be changed, there is no guarantee that the data are erased on the provider's side.

However, long-term thinking and planning is difficult for human beings, and there are few incentives to think ahead for more than, say, five years: This is similar in the political sector with a session of parliament and in the economic sector when calculating the return on investment [27].

3.8 Mistake 8: Changing Assumptions or Surplus Functionality

Related to the problem of considering the full lifecycle are changing assumptions: In the beginning of designing a system, the developers focus on the functionality. Meanwhile several system designers think of implementing some privacy functions. But this effort may be completely in vain if later the assumptions change or surplus functionality is being implemented. All privacy "guarantees" may be gone, and this has to be communicated.

For instance, in one case a (legally based) cancer registry processed pseudonymized data with some introduced fuzziness, i.e., with some probability, entries belonging to the same person got different pseudonyms, and entries belonging to different persons got the same pseudonym. This was fine for statistics purposes. But after some years, the requirements changed: The cancer registry should establish a feedback system to the persons whose data were processed, but since there was (on purpose) no bijective assignment, this would have created the risk of informing the wrong persons on a very sensitive issue.

Even if we assume that a privacy-compliant service with exemplary data minimization and transparency has been developed, surplus functionality may water down or even contradict the intended privacy guarantees. In particular, a surplus payment method, a business model basing on profiling and advertising, or obligations from the police or homeland security could render all privacy efforts useless.

3.9 Mistake 9: No Intervenability Foreseen

Many system developers – no matter whether from the technological, legal or economic discipline – try to build systems that work well to solve a problem. Sometimes they forget an important property: It has to be possible to change the system, and it has to be possible to shut it off. The possibility to intervene is relevant for the entities processing the data, for the supervisory authorities that may inspect the data processing system, and – at least partially – for data subjects whose data are being processed, simply because of their data subject rights.

For instance think of an ambient assisted living scenario where a person living in her household is being monitored by video cameras so that a guard can immediately call for help in case of an accident. Still there should be the possibility to deactivate the surveillance for the person concerned if she does not want to be monitored for some time. Of course this would change the obligations of the guard who cannot react without the signals [28].

3.10 Mistake 10: Consent Not Providing a Valid Legal Ground

All processing of personal data is only lawful if a statutory provision permits it or if the data subject has consented. In many cases this means that the data controller has to ask the data subjects for their consent. This is easier said than done: Consent that provides a valid legal ground has to meet various requirements (this has been made

explicit in the draft of the European General Data Protection Regulation [14], but has been discussed and implemented in the law of various Member States):

- "The data subject's consent" means any freely given specific, informed and explicit indication of his or her wishes by which the data subject, either by a statement or by a clear affirmative action, signifies agreement to personal data relating to them being processed.
- The data subject is aware that and to what extent consent is given.
- The consent has to be freely given; this is not the case if the data subject has no genuine and free choice and is subsequently not able to refuse or withdraw consent without detriment.
- Consent should not provide a valid legal ground for the processing of personal data, where there is a clear imbalance between the data subject and the controller.

For instance if the given information is not accurate and comprehensive, the consent does not provide a valid legal ground. The consent must not be hidden in the privacy policy or terms and conditions; statements such as "By using <xy>, you agree to the <xy> Terms of Service" do not form a valid consent. The consent of one person does not cover the consent of others, e.g., the "consent" in a social network that an application may "access my friends' information".

And it is doubtable that anybody would give consent to the following phrase from the (former) "Terms of Use" of World of Warcraft/Blizzard: "Blizzard may monitor, record, review, modify and/or disclose your chat sessions, whether voice or text, without notice to you, and you hereby consent to such monitoring, recording, review, modification and/or disclosure." Or who would consent to all possible modifications of one's chat session without being notified?

4 Privacy Protection Goals

The ten mistakes listed in the previous section illustrate main areas to specifically look at when designing or assessing systems. However, a more general method can be employed that is widely used in IT security system design: working with protection goals (4.1). IT security protection goals such as confidentiality, integrity and availability are well known. Here we present a complementary set of privacy protection goals (4.2) and give advice how to employ them (4.3). Further, the relation of these protection goals with the data protection principles elaborated in Section 2.2 and with the "privacy by design" method that is being propagated by Cavoukian [29] is discussed (4.4).

4.1 Working With Protection Goals

For decades, skilled system designers and engineers have been working with the traditional security protection goals confidentiality, integrity and availability, also called the "CIA triad". These protection goals are driving factors for assessing the risks and investigating potential damages if the desired level of protection cannot be achieved.

Thus, they are part of the work on the information security management system (ISMS) and its core element, the security concept [30].

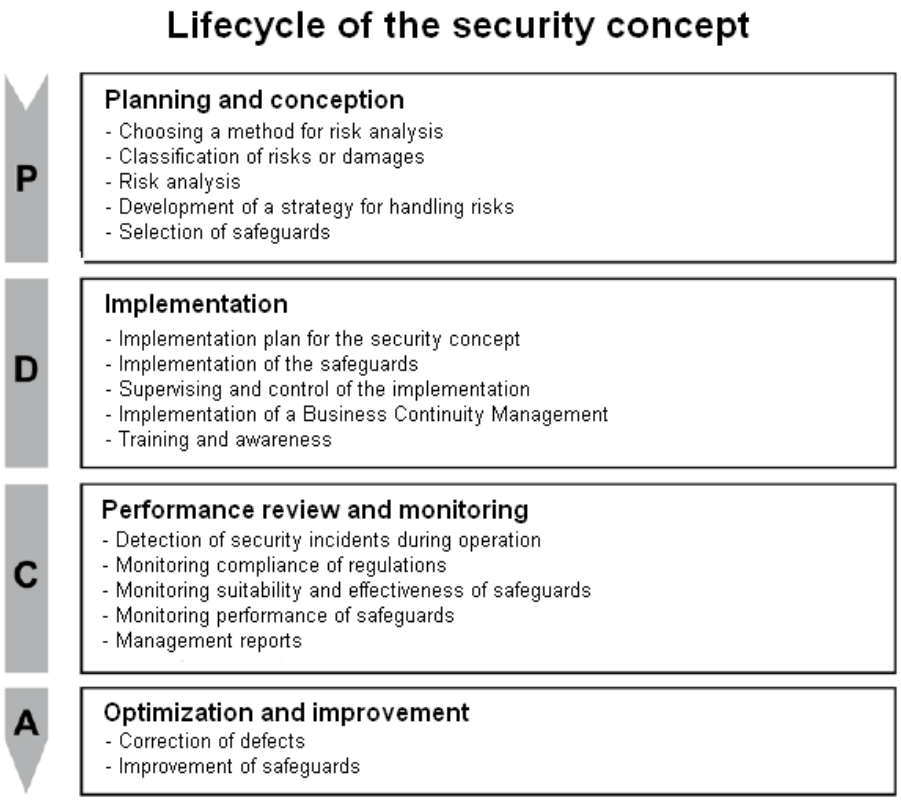

Fig. 2. Iterative procedure to conceptualize, implement, review and improve the security concept and appropriate safeguards (revised version from [30]). It follows the Deming cycle of Plan (P), Do (D), Check (C) and Act (A) that is to be repeated for further adaption.

In the realm of IT security, the established procedures for setting up and maintaining ISMS are described in detail in a series of documents (beginning with [30]). They are part of the audit on the basis of IT Baseline Protection standardized by the German Federal Office for Security in Information Technology which is also part of the ISO 27001 certification. Note that the security concept and the appropriate safeguards are not grafted on top of the designed systems, but have to be interwoven. Thus, the work on and with ISMS is highly relevant for system design from a security perspective throughout the full lifecycle of development.

4.2 Privacy Protection Goals

For privacy-related issues, manuals and catalogues comparable to the elaborated security protection goals are still missing. In particular it would not work to stick to the CIA triad because it reflects the information security perspective only. Therefore, privacy-specific protection goals have been proposed that represent the most important requirements from a privacy perspective: unlinkability, transparency and intervenability [31][32]. These privacy protection goals are described in the following.

Unlinkability aims at separating data and processes: This means that processes must be operated in such a way that the privacy-relevant data are unlinkable to any other set of privacy-relevant data outside of the domain. If full unlinkability cannot be achieved, it should be realized to the extent that linking would require disproportionate efforts for the entity establishing such linkage. The objective of this protection goal is the minimization of the risk to privacy by misusing privacy-relevant data.

Since unlinkability covers separation from personal data and the related data subjects, it is the key element for data minimization. Further, the separation of data sets belonging to different purposes supports the principle of purpose binding. Clearly, unlinkability and separation of powers are related.

Examples for achieving and enhancing unlinkability comprise data avoidance, separation of contexts by different identifiers, anonymization and pseudonymization mechanisms, and early erasure.

Transparency aims at an adequate level of clarity of the processes in privacy-relevant data processing so that the collection, processing and use of the information can be understood and reconstructed at any time. Further, it is important that all parties involved can comprehend the legal, technical, and organizational conditions setting the scope for this processing. This information has to be available before, during and after the processing takes place. Thus, transparency has to cover not only the actual processing, but also the planned processing (ex-ante transparency) and the time after the processing has taken place to know what exactly happened (ex-post transparency).

For data controllers, comprehensive transparency is needed, e.g., they have to know exactly how their data processors handle the data. On request, this level of transparency may be demanded by the supervisory authority. For data subjects, it is important that they can obtain full information on their own personal data and the most important information how the data are processed. This comprises for what purposes the data are processed, for how long, which recipients receive the data, the logic of the data that are undergoing the processing and the intended and possible consequences of such processing, e.g., in case of profiling [14]. All parties should know the risks to privacy and have sufficient information on countermeasures, how to employ them and what limitations they have.

Examples for achieving or enhancing transparency comprise reporting mechanisms, an understandable documentation covering technology, organization and responsibilities, the source code, privacy policies, information of and communication with the data subject.

Intervenability aims at the possibility for parties involved in any privacy-relevant data processing to interfere with the ongoing or planned data processing. The objective of intervenability is the application of corrective measures and counterbalances where necessary.

For data controllers, it is necessary to be able to effectively control the data processor and the used IT systems to influence or stop the data processing at any time. For data subjects, the rights to rectification and erasure of data as well as the right to withdraw consent are part of the intervenability. Moreover, intervenability addresses the data subject's right to lodge a claim or to raise a dispute to achieve remedy. Supervisory authorities may intervene by requesting or enforcing the blocking, erasure or destruction of data or even shutting off the system.

Examples for achieving or enhancing intervenability are established processes for influencing or stopping the data processing fully or partially, manually overturning an automated decision, data portability precautions to prevent lock-in at a data processor, breaking glass policies, single points of contact for data subjects' intervention requests, switches for users to change a setting, e.g., changing to a non-personalized, empty-profile version of a search engine or recommendation system, or deactivating an auto pilot or a monitoring system for some time (see 3.9 for the ambient assisted living scenario).

4.3 Working With Privacy Protection Goals

Working with privacy protection goals means to consider the CIA triad as well as unlinkability, transparency and intervenability. In addition to the ISMS, a complementing "Privacy Protection Management System" could be set up. Note that the protection goals have dependencies with each other: For instance, high confidentiality with complex access control mechanisms and encrypted files could complicate the access for authorized persons, too, so that the level of availability would be decreased. Perfect technical integrity of a data collection could hinder necessary contentwise corrections, demanded by intervenability. Also, integrity, availability and transparency of data traces would work against unlinkability and confidentiality. This shows that balancing of the requirements derived from the six protection goals is needed. However, it does not mean that there is necessarily a zero-sum balance between privacy and security [29], but depending on the choice of instruments improvements in several or all areas may be possible.

All of the protection goals can in principle be applied on the data themselves as well as on technical and organizational processes. The perspectives of all parties involved, such as data controllers, data processors, data subjects or third parties, have to be considered when assessing the value of assets and the consequences in case of damage or loss. Privacy protection goals can help to structure the risk analysis as well as the choice of safeguarding instruments when designing the system. For some application scenarios such as ambient assisted living and smart meters the use of the six protection goals has been tested [28]: A three-dimensional matrix has been developed to map protection goals, data types and processes, applying the perspectives of all parties involved. Even without ready-to-use catalogues enlisting safeguarding instru-

ments and their characteristics, a walk-through structured according to the developed matrix has revealed where decisions for system design would have to be made and how they could be argued.

The privacy protection goals are proposed for standardization of the ISO Privacy Reference Architecture [33] and have become part of the revised Data Protection Act Schleswig-Holstein [34].

4.4 Positioning the Protection Goals with Regard to Related Approaches

The most distinct characteristic of the privacy protection goals is their structural similarity to the CIA triad: Thereby, for employing these goals the well established procedures from [30] can be inherited, and system designers and engineers will quickly learn how work with them. This can be further promoted by elaborating catalogues with safeguarding instruments assigned to the respective protection goals. However, the application of privacy protection goals does not guarantee lawfulness, see Table 1.

Table 1. Relation of the protection goals and the seven data protection principles.

	Unlinkability	**Transparency**	**Intervenability**	**Other**
Lawfulness				
Consent		X	X	
Purpose binding	X			
Necessity and data minimization	X			
Transparency and data subject rights		X	X	
Data security				CIA
Audit and control		X	X	Integrity

Unsurprisingly, legal compliance that is based on the existence of statutory provisions with their own complexity cannot be evaluated in an abstract way – without knowing the actual law. But the legal provisions are important factors for the appropriate balancing of the protection goals and for the choice of safeguarding instruments.

Similarly Cavoukian's pleading for privacy by design (PbD) [29] stresses the pro-privacy balancing when employing the protection goals. Table 2 shows which PbD principle belongs to which characteristics when employing protection goals.

Table 2. Relation of the protection goals and the seven Privacy by Design [29] principles.

	Part of the design process	**Balancing criteria**	**Addressing specific protection goal**
1. Proactive not reactive – preventative not remedial	X (prior risk assessment)		Risk avoidance: see entry for 5.
2. Privacy as the default setting		X	
3. Privacy embedded into design	X		
4. Full functionality – positive-sum, not zero-sum	X (choice of safeguards)	X	
5. End-to-end security – full lifecycle protection	X (full life-cycle)	X	CIA, possibly unlinkability
6. Visibility and transparency – keep it open		X	Transparency
7. Respect for user privacy – keep it individual and user-centric		X	Intervenability (for users)

Some of the seven PbD principles (obviously principle 3, but also 1, 4 and 5) focus on the "by design" part, i.e., they tackle specifically the embedding in the design process, partially with a concentration on specific properties (risk assessment, choice of safeguards, throughout the full lifecycle).

Some principles play a role on balancing the interplay between the various protection goals and of course the main functionality of the system: principles 2 and 4 only emphasize a bias on privacy and encourage to find solutions that maximize both privacy and security; principles 5 to 7 address specific protection goals (5: the CIA triad, 6: transparency, 7: intervenability). Note that mainly the user perspective is taken, probably because this perspective is usually least considered by system designers. However, transparency and intervenability are important for the other parties involved, too.

How much unlinkability, or data minimization, play a role in the seven PbD principles, is not fully clear. Unlinkability could be seen as an instrument to "prevent privacy breaches from occurring" [29] and it also fits when discussing the lifecycle of personal information "from the point of collection through to its secure and timely destruction" [29] although this issue concentrates on "strong security controls".

All in all, privacy by design principles can help working with protection goals by a pro-privacy attitude and a focus on users, while the three-dimensional usage matrix [28] and the standardized procedure laid down in [30] have the advantage of a more comprehensive approach, supported by an ISMS and possibly a privacy protection management system.

5 Conclusion and Outlook

This paper has discussed how privacy may be defined and what to keep in mind when designing systems that process data. Here the concept of linkage control is introduced as an essence for privacy protection.

The list of ten important and typical mistakes reveals challenges for system designers. Talking about these mistakes in this paper will not be sufficient to prevent that these mistakes will be repeated, because they are caused or emphasized by the intrinsic logic of current data processing systems and business procedures as well as the education of people designing systems with a technical, legal or economic background. However, it will generate awareness for such problems.

A more general approach, in fact an extension of the idea of linkage control as key to privacy, is the proposal for introducing three protection goals complementing the CIA triad: unlinkability, transparency and intervenability. Since many system designers are aware of the function of protection goals and know procedures for risk assessment, generating security concepts or maintaining information security management systems, this well established set of instruments can be extended by the notion of privacy protection goals. The discussion of related approaches illustrates that a legal analysis of requirements would be still necessary because there might be very specific statutory provisions. A combination with the privacy by design approach could easily be done. The protection goals even could extend it a bit by strengthening the role of unlinkability and considering further parties involved.

Further effort has to be invested to elaborate catalogues that list appropriate instruments for the various privacy protection goals and point out possible dependencies between the goals. The first test cases have been done in the area of upcoming technologies (ambient assisted living, smart meter, cyber-physical systems) rather than fully specified procedures. Working with the privacy protection goals will become daily business for the Data Protection Authority Schleswig-Holstein, Germany, because since January 2012 they have to be considered when designing automatic systems for the public sector in the region.

References

1. Solove, D.J.: Understanding Privacy, Harvard University Press (2008)
2. Warren, S.D., Brandeis, L.D.: The Right to Privacy. Harvard Law Review, Vol. 4, No. 5, 193–220 (1890)
3. Westin, A.F.: Privacy and Freedom. Atheneum, New York (1967)
4. Benda, E., Simon, H., Hesse, K., Katzenstein, D., Niemeyer, G., Heußner, H., Henschel, J.F.: BVerfGE 65, 1. In: Mitglieder des Bundesverfassungsgerichts (eds.) Entscheidungen des Bundesverfassungsgerichts. 65, pp. 1–71. Mohr, Tübingen (1983)
5. Bizer, J.: Sieben Goldene Regeln des Datenschutzes. Datenschutz und Datensicherheit (DuD), Vol. 31, No. 5, 350–356 (2007)
6. Phillips, D.J.: Privacy Policy and PETs – The Influence of Policy Regimes on the Development and Social Implications of Privacy Enhancing Technologies. New Media & Society, London, Thousand Oaks, CA and New Delhi: SAGE Publications, Vol. 6, No. 6, 691–706 (2004)

7. Hansen, M.: Linkage Control – Integrating the Essence of Privacy Protection into Identity Management Systems. In: Cunningham, P., Cunningham, M. (eds.) Collaboration and the Knowledge Economy: Issues, Applications, Case Studies. Proceedings of eChallenges 2008, pp. 1585–1592. IOS Press, Amsterdam (2008)
8. Hansen, M.; Meissner, S. (eds.): Verkettung digitaler Identitäten, Untersuchung im Auftrag des Bundesministeriums für Bildung und Forschung, Kiel, https://www.datenschutzzentrum.de/projekte/verkettung/2007-uld-tudverkettung-digitaler-identitaeten-bmbf.pdf (2007)
9. McCormick, J.: Top 10 security mistakes to avoid. TechRepublic, 30.04.2007, http://www.techrepublic.com/blog/security/top-10-security-mistakes-to-avoid/221 (2007)
10. Horowitz, A.S.: Top 10 Security Mistakes. Computerworld, 09.07.2001, https://www.computerworld.com/s/article/61986/Top_10_Security_Mistakes (2001)
11. SANS (SysAdmin, Audit, Network, Security) Institute: The Ten Worst Security Mistakes Information Technology People Make. 10.09.2005, ttps://www.sans.org/security-resources/mistakes.php (2005)
12. Ashish: 10 common security mistakes that should never be made. Mind Tree, 22.08.2008, http://www.hurricanesoftwares.com/10-common-security-mistakes-that-should-never-be-made/ (2008)
13. Mayer-Schönberger, V.: Useful Void: The Art of Forgetting in the Age of Ubiquitous Computing. Faculty Research Working Papers Series No. RWP07-022. John F. Kennedy School of Government – Harvard University, http://www.vmsweb.net/attachments/pdf/Useful_Void.pdf (2007)
14. European Commission: Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). COM(2012) 11 final, Brussels, 25.01.2012, http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf (2012)
15. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management, v0.34, 10.08.2010, http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf (2010)
16. Electronic Frontier Foundation: AOL's Massive Data Leak, http://w2.eff. org/Privacy/AOL/ (2006)
17. Narayanan, A., Shmatikov, V.: Robust De-anonymization of Large Sparse Datasets. IEEE Symposium on Security and Privacy 2008, pp. 111–125 (2008)
18. boyd, d.: "Real Names" Policies Are an Abuse of Power. Blogpost from 04.08.2011, http://www.zephoria.org/thoughts/archives/2011/08/04/real-names.html (2011)
19. Information Commissioner's Office (ICO): Privacy Impact Assessment Handbook, Version 2.0, http://www.ico.gov.uk/upload/documents/pia_handbook_html_v2/ (2009)
20. Nissenbaum, H.: Protecting Privacy in an Information Age: The Problem of Privacy in Public. Law and Philosophy, Vol. 17, No. 5, 559–596 (1998)
21. Borcea-Pfitzmann, K., Pfitzmann, A., Berg, M.: Privacy 3.0 := data minimization + user control + contextual integrity. it – Information Technology, Vol. 53, No. 1, 34–40 (2011)
22. Gomez, J., Pinnick, T., Soltani, A.: KnowPrivacy. 01.06.2009, http://www.knowprivacy.org/report/KnowPrivacy_Final_Report.pdf (2009)
23. Markoff, J.: Internet Traffic Begins to Bypass the U.S. The New York Times, 30.08.2008, https://www.nytimes.com/2008/08/30/business/30pipes.html (2008)
24. Bowden, C.: Privacy and surveillance on the Internet – What happened, and what to expect next… Presentation slides from 20.09.2011, http://wolnyinternet.panoptykon.org/sites/default/files/internet_surveillance_caspar_bowden.pdf (2011)

25. Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD): Inanspruchnahme des Patriot Acts und anderer US-rechtlicher Regelungen zur Beschaffung von personenbezogenen Daten aus dem Raum der Europäischen Union durch US-Behörden. Position paper, Kiel, 15.11.2011, https://www.datenschutzzentrum.de/internationales/20111115-patriot-act.html (2011)
26. Storf, K., Hansen, M., Raguse, M. (eds.): Requirements and Concepts for Identity Management throughout Life. Deliverable H1.3.5 of the EU FP7 Project PrimeLife, Zürich/Kiel 2009, http://www.primelife.eu/images/stories/deliverables/h1.3.5-requirements_and_concepts_for_idm_throughout_life-public.pdf (2009)
27. Hansen, M.: Towards future-proof privacy-respecting identity management systems. In: Pohlmann, N., Reimer, H., Schneider, W. (eds.) ISSE 2010 – Securing Electronic Business Processes, Highlights of the Information Security Solutions Europe 2010 Conference, pp. 182–190. Vieweg + Teubner Verlag, Wiesbaden (2010)
28. Rost, M.: Datenschutz in 3D – Daten, Prozesse und Schutzziele in einem Modell. DuD, Vol. 35, No. 5, 351–355 (2011)
29. Cavoukian, A.: A Foundational Framework for a Privacy by Design – Privacy Impact Assessment. http://privacybydesign.ca/content/uploads/2011/11/PbD-PIA-Foundational-Framework.pdf (2011)
30. Bundesamt für Sicherheit in der Informationstechnik (BSI): BSI-Standard 100-1: Information Security Management Systems (ISMS), Version 1.5, Bonn, https://www.bsi.bund.de/cae/servlet/contentblob/471428/publicationFile/28221/standard_100-1_e_pdf.pdf (2008)
31. Rost, M., Pfitzmann, A.: Datenschutz-Schutzziele – revisited. DuD, Vol. 33, No. 12, 353–358 (2009)
32. Rost, M., Bock, K.: Privacy By Design und die Neuen Schutzziele – Grundsätze, Ziele und Anforderungen. DuD, Vol. 35, No. 1, 30–35 (2011)
33. Hedbom, H., Schallaböck, J., Wenning, R., Hansen, M.: Contributions to Standardisation. In: Camenisch, J., Fischer-Hübner, S., Rannenberg, K. (eds.) Privacy and Identity Management for Life, pp. 479–492. Springer, Berlin (2011)
34. Schleswig-Holsteinisches Gesetz zum Schutz personenbezogener Informationen (Landesdatenschutzgesetz - LDSG -). Version after the last change that has been published in: Gesetz- und Verordnungsblatt für Schleswig-Holstein (GVOBl. SH 2012, No. 2, pp. 78–82), https://www.datenschutzzentrum.de/gesetze/ldsg.html (2012)