Summary  Over the last two decades, privacy has been fading away. Some people have even stated: You have zero privacy – get over it! As privacy researchers, we are not willing to accept this statement. Therefore, we analyze the causes for this fading away of privacy, and develop a set of approaches to preserve or even regain privacy. We argue that Privacy 3.0 should be a combination of (1) Data minimization, (2) User control of personal information disclosure, and (3) Contextual integrity. Data minimization is one of the main motivations for the development of privacy-enhancing technologies, which aim to limit collection and processing of personal data by data controllers. User control of personal information disclosure supports users in deciding which personal information is released to whom and in which situation. Contextual integrity provides a new quality of privacy by making the original context in which particular personal data have been generated easily accessible to all entities that are aware of that particular personal data.

Keywords  Privacy, data minimization, user control, privacy-related context, contextual integrity

1 Introduction
Recent debates about privacy, including the one on the upcoming census in 2011 in Germany, have shown that many people claim to have a right to privacy. However, they do not perceive that their privacy is granted given the ever more pervasive deployment and use of information technology. So, privacy still provides quite a large field for exploration regarding use cases, definitions, legal regulations, etc.

To prepare our problem analysis, we will briefly survey the evolution of privacy as a concept in the following Sect. 2. In Sect. 3, we will discuss two important privacy issues related to human beings interacting with and via computers. Following this, Sect. 4 provides a three-component approach to privacy arguing that the traditional approach of data minimization (it may be referred to as Privacy 1.0) based on legal regulations is not always feasible and certainly not in every situation. In addition, disclosure of personal data has to be user-controlled, which may be called Privacy 2.0. Since even this is not realistic for ubiquitous computing, we put up for discussion the concept of contextual integrity as a third component of privacy (this combination of the three components is referred to as Privacy 3.0). Conse-
Privacy 3.0 := Data Minimization + User Control + Contextual Integrity

2 History of Privacy

The general concept of privacy has been known from time immemorial – in the meaning of hiding private things from unauthorized others (i.e., establishing a private sphere) which coined the term confidentiality. Particular concerns about privacy were developed over a hundred years ago when modern technologies to store and transmit information gained importance in everyday life. In the 1970s, when information technology was establishing itself in companies, governmental institutions, and, last but not least, people’s personal lives, the idea of privacy reached a new dimension: the protection of people against unnecessary storage and processing of their personal data. The limitation of the storage and processing personal data is referred to as data minimization. In the 1990s and 2000s, the requirement of data minimization became part of according European legal regulations, cf. the data minimization principle (Directive 2002/58/EC) and the purpose binding principle (Art. 6 (1b), Directive 1995/46/EC).

Since then, information technology has experienced a remarkable evolution. It has strongly influenced human beings’ perception and demands with regard to their privacy. One major milestone relating to privacy was the attempt to conduct a large-scale population census in Germany in 1983. Many people considered some of the questions proposed for the census to be too intrusive. Further, it was planned to update data in the central population register using the data gathered during the census. This, however, would have contradicted the principle of purpose binding. As a result, the German Federal Constitutional Court stopped the census by a fundamental policy decision which elaborated the right to informational self-determination.

Currently, the discussion on privacy is experiencing a second major wave through the developments occurring in the field of social computing. People use social networking platforms to connect to each other, to communicate, and to collaborate. However, the “average Joe and Jane” lose track of the implications of their use of networking platforms. They are usually not aware that the audience of their contributions has the possibility to access the personal data of the user that is stored with the platforms. The discussions related to the census and networking platforms differ in one major point: The increased awareness of people about their right of privacy during the census activities in 1983 was mainly triggered by the governmental plans to force people to indicate accurate and complete personal information, thus demanding that German citizens give up control over their data. In comparison, users of social networking platforms feel that they are in control of their data and are not (really) forced by some institution to reveal things they would rather not indicate. The privacy discussions raised in the context of “networked” users are driven mainly by people who work in the field of privacy (such as privacy researchers, and journalists).

Looking at the historical evolution of the concept of privacy from a more general point of view, we have to admit that the approach towards how privacy is to be understood has changed a lot. It started from a more or less social understanding of private life and reached in the 1980s a kind of formalization by a first approach to legally capture the actual concept (i.e., the right to privacy, the right to self-determination). The aim was to protect individuals when they interact with organizations. Since, at those times, individuals did not have the possibility to acquire and use appropriate security tools and mechanisms, they could only rely on legal regulations as their “privacy guard”. Thus, privacy was understood as being controllable within a certain legal jurisdiction.

Today, the situation has changed dramatically. Firstly, state borders do not play a significant role in delimiting legal jurisdictions on the Internet from each other, i.e., people access applications from servers which may be located in any country of the world and, to make it legally even more challenging, the service providers reside in yet another country. Privacy, therefore, is no longer controllable within a particular legal jurisdiction. Secondly, individuals no longer communicate exclusively with organizations, but largely with other individuals. This requires adjustments of privacy-related legal regulations to also consider civil interactions. Thirdly, privacy as we understand it today also implies the development of security technology – users may use tools to protect their data.

This immense change in information technology, on the one hand, and of the requirements for privacy on the other hand, requires a closer look at the problems that originate when human beings interact with technical devices or their interactions are mediated by technical devices.

3 Identifying and Analyzing Major Privacy Problems

Two major problems with respect to informational privacy (i.e., protection of personal data being subject of data processing) can be identified:

(1) Available information technology is inherently insecure. This problem got even worse with information technology having undergone an enormous growth of functionality whereby this as well as feature interactions within the applications (e.g., mashup applications combining data and functionality from several sources) contribute to the increase of insecurity.
(2) Human beings' perception of their interactions with information technology are similar to a person's perception towards another human being [2]. This means that human beings unconsciously apply rules to their interaction with computers that are very similar to those of social interaction. The usual rules of social interaction are:

• **nondisclosure**, i.e., human beings tend to presume that other human beings do not disclose information that is very personal.

• **natural process of oblivion**, i.e., normally human beings gradually forget or lose track of information that they did not use for some time.

• **conscience**, i.e., normally human beings are able to distinguish right from wrong and factor this into their interactions with other human beings.

These effects even get proliferated these days with developers of human-computer interfaces putting a considerable effort into concealing the technology behind the interfaces and ensuring that the human-computer interface approximates human-like behavior. This is a concession to ordinary users who use computers as a medium only to communicate with others, i.e., they do not interact with, but via computers. Consequently, such an approach inherently favors the issue of human-like perception of technology by human beings: the computer as a technical device fades away, but it still does not follow the rule of nondisclosure (unless they are programmed that way and contain tamper-resistant hardware), as it does not forget nor does it possess conscience.

Looking into the future, we have to admit that none of the indicated problems will vanish, i.e., it is neither to be expected that information technology will be 100% secure nor that human beings will change their manner of interacting with and via computers. Consequently, an approach inherently favors the issue of human-like perception of technology by human beings: the computer as a technical device fades away, but it still does not follow the rule of nondisclosure (unless they are programmed that way and contain tamper-resistant hardware), as it does not forget nor does it possess conscience.

A first approach to meeting the problems of a false perception of technology by human beings would be to avoid making technology human-like. One possibility to resolve this issue is the integration of awareness features into the indicated issues would be to accept that:

**Data minimization as a privacy property is neither technically realizable in every respect nor is it socially acceptable in every situation.**

Regarding this, privacy may be understood as a three-component concept as described in the following paragraphs. We will see from this discussion that the differentiation into the three components is mainly driven by the constraints of the historical evolution of information technology and, above all, the kind of interaction with and via computers of the users:

(1) **Data minimization:** In the early 1980s, computers were primarily used for work. Communication via networks served as a fast means of information flow to support cooperation between co-workers. At that era, data minimization was technically controllable as the applications' functions were rather simple. Consequently, data minimization is to be understood as the traditional driving concept of the field of PET (privacy-enhancing technology) research (e.g., anonymous communication and pseudonymous e-commerce). The research concepts based their considerations on the fact that interactions took place mainly between individuals and organizations. However, with the evolution of Web 2.0 technologies and thus with the support for active user participation, the concept of data minimization as the prime property of privacy has reached its limits. Human beings require more and more functionality, and this desire dominates their demand for privacy in regards to data minimization.

(2) **User control of personal data disclosure:** In the 1990s, the human beings' perception and type of use of computers in the sense of computers becoming "just" a medium for transporting information gained more and more importance. Thus, social interaction has become technically supported. Consequently, Westin's approach of defining privacy as a "claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [11] became a particular issue for a world with computers, i.e., software developers need to consider providing possibilities for user control with regard to the kind of sharing of the users' personal data to other users.

As social interaction is very common and varied, the functions of the interaction-supporting technology have achieved a level of sophistication and mutual interdependence that is not legally and technically controllable with regard to privacy any more. Additionally, it was becoming impossible to serve the competing human demands for comfort and some kind of bargain-hunter mentality2, on the one hand, and keep personal data confidential or minimize disclosed data, on the other hand. To overcome this dilemma, the disclosure of personal data had to become user-controlled. Thus, the users themselves had to accept more responsibility for their data disclosure.

---

2 Human beings typically tend to accept offers in which their personal data as well as their behavioral data are collected if they get better and/or more service or discounts for it.
Currently, different approaches are subject to research and applications aiming at technical support for the users in that task. These comprise privacy-enhancing identity management and education of the general public about the consequences of disclosure of personal data. Further work is being done in developing semi-automated decision making processes in regards to which personal data should be disclosed to whom in which situation by defining privacy contexts [3; 4] (see also Sect. 5). Another idea is comprised of mechanisms that make the user aware of potential consequences when personal data is disclosed [7].

3 Contextual integrity: A third privacy component is just becoming important. These days, devices and software are being developed to satisfy the desire for availability of functionality everywhere and at any time. This, together with the massively increasing possibilities for storing and processing of information available in various forms (i.e., multimedia) leads to the problem that neither the realization of data minimization nor that of user control is always possible. So, if none of the two indicated privacy characteristics can be fulfilled, we will have to deal with a third one, which aims at protecting individuals from embarrassment by falsifying the context in which information has been communicated. Thus, the objective of the third privacy component is to assure contextual integrity of communicated information and, that way, to protect people from the decontextualization of communicated information. The term of contextual integrity is actually not new to privacy research. It was first used by Helen Nissenbaum in 1998 [6]. For her, contextual integrity means that personal data must not be transferred from one social context (defined by norms, which govern the gathering and dissemination of personal data) into another. This understanding differs from the one to which we are referring in so far as our concept allows for the transfer of personal data between different contexts. However, our concept further demands that data describing the context from where the personal data originates is transferred together with the personal data.

On the topic of protecting users from embarrassment, danah boyd coined the term “Personally Embarrassing Information” in comparison to “Personally Identifiable Information” [8]. This term, however, relates only to the use case of disclosing information about oneself that could potentially be embarrassing. It does not consider turning any innocuous information into mortifying information by putting it into a wrong context.

5 The Concept of Context in Privacy Considerations

Privacy advocates often associate the new possibilities of ubiquitous technology with the loss of privacy, as more and more devices become able to sense environmental properties (attribute values of certain characteristics of the environment in which a user acts), including sensing human beings. This facilitates the creation of even more detailed user profiles.

Given that it is not possible to fully prevent such situations, the question could be asked whether the conceptual ideas in the field of ubiquity, in particular the concept of context, could be used to maintain privacy in the sense of the second (user-controlled disclosure of personal data) and/or the third (contextual integrity) components. This would mean that if it is not possible to achieve data minimization (first component of Privacy 3.0), environmental properties should be used to determine the context in which a to-be-protected activity occurs.

Context is a well-established concept of research regarding ubiquitous computing. It was introduced by Schilit et al. as “the location of use, the collection of nearby people, hosts, and accessible devices as well as changes to these aspects over time” [9]. The definition has been subject to several discussions since then. However, the main point that all approaches share is the relation of context to some physical environment.

We argue that context can also be used for application environments in social computing where the instances of the software as well as the data are distributed to a huge number of computing devices and the states of that software and of the data are not fully determinable any more. In these cases, context can be used to define the properties of an environment that are important for selected issues.

We see two different aspects where context plays a role in privacy considerations, whereby the first one corresponds to approaches of supporting users in controlling the disclosure of their personal data (the second component of the three-component approach to privacy), which we will call privacy-related contexts of disclosure or in short – disclosure contexts. The second aspect follows the idea of contextual integrity (the third component). We will call these contexts privacy-related contexts of integrity, which is comprised of integrity contexts.

1) Supporting users in selecting personal data for disclosure: Users distinguish between situations (classified in contexts) in which they disclose selected information relevant to the context and also distinguish between different levels of detail of information. For example, they indicate the date of their birthdays (i.e., day, month, and year of birth) in of-

3 To give just two examples: Video surveillance in public transportation or on streets cannot be escaped. Also, it can hardly be prevented that, when people post descriptions of events to the Internet, they include remarks about another person.

4 Environmental properties are not limited to physical location-related characteristics, but may also comprise software spaces, e.g., namespaces in Wikis or workspaces in groupware.

5 Other terms for ubiquitous computing are: pervasive computing, and ambient intelligence.
ficial forms of governmental institutions while they reveal just their ages in informal situations where they are not yet confident in their communication partners. Thus, the concept of context is used to classify situations with similar environmental properties to determine which personal data should be disclosed to whom and in which situation. Related to this, users tend to maintain contexts and to keep the disclosed information the same as long as the context remains the same. To give an example, a random user A meets another user B in a chat. As long as their relationship does not change qualitatively, user A will not disclose (much) more information to user B than A already disclosed. In the event that user A has the feeling that s/he would benefit from intensifying the relationship with user B, user A could disclose some more information in order to motivate user B to deepen the relationship. This would mean that user A adjusts the context and all situations that will now fall into that context, i.e., the users A and B meet in environments with similar characteristics and are associated with all disclosed information in this context.

(2) Supporting users in maintaining contextual integrity of data: Users want to be protected against falsification of the semantics of their personal data by putting information into a different context. Thus, their interest is to see their personal information continuing to be associated with the actual context in which it was disclosed. This needs to be ensured along the whole chain of information transmissions. To give an example: let user A indicate in the mood section of an instant messaging service “I am blue today”. Without knowing the actual context, other people could interpret this message very differently: one person may argue that user A painted with his/her child and s/he got blotted with blue color by the child. Another user might interpret the message meaning that user A feels sad. If a third user is not familiar with the English language, s/he might think that user A is drunk, as a German equivalent (“blau”) of the English word “blue” means to be inebriated. This misunderstanding could get even worse if the information were to be transferred to a number of different people who could use the (misinterpreted) information to insult user A. Attaching information about the context why user A is “blue today” to the actual statement would prevent others from misinterpreting the message.

6 Implementability of Privacy-related Contexts

6.1 Definitions

Disclosure contexts are connected to the well-known privacy sphere model, the idea of which was published in the form of the privacy sphere theory in [5]. The sphere model in its generalized form defines a set of concentric circles, i.e., the spheres are associated with different levels of protection. The most secret sphere is the inner circle and the public sphere is the outer circle, and there are some graduations in between. The drawbacks of the sphere model are that the spheres are hardly able to be isolated from each other and that each individual defines his/her spheres differently. Also, it is not always reasonable to make the differentiation of which information to disclose and in which detail depending on the spheres only. For example, even if we would define one of the spheres as “trusted family”, human beings typically still reveal different information to their parents than they would towards their siblings.

This is why we define a disclosure context more generally as follows:

Definition: Privacy-related context of disclosure (disclosure context) is a user-defined classification of situations, in which the user discloses personal data that consists of similar environmental properties. Disclosure contexts are used in decision-making processes about which personal data to disclose.

Regarding this, the phrase “similar properties” refers to a selected set of properties that is the same for all situations grouped by the context class. Thus, a system that supports disclosure contexts should encompass the definition of contexts as well as the assignment of selected information to the context, the recognition of a situation being classified as a particular context and, of course, the management of personal information.

Integrity contexts differ from disclosure contexts in at least three aspects:

(1) Encapsulation: Integrity contexts are not used to encapsulate or to classify, respectively, different situations. Instead, integrity contexts are only associated with the disclosed information.

(2) Objective: The objective of integrity contexts is not to support making the decision which information to disclose to whom, but integrity contexts are used to link to a disclosed information in order to prevent mis-/dis-interpretation6 of the actual information.

(3) Visibility: While disclosure contexts are managed by and visible to the disclosing user only (unless the user discloses her/his contexts to others), integrity contexts have to be visible to all recipients of the disclosed information and the actual environmental properties may also be negotiated between the parties involved.

Nevertheless, disclosure contexts and integrity contexts have one characteristic in common: The user defines a particular set of environmental properties that s/he perceives as essential for the respective task. Consequently, we define integrity contexts as follows:

6 Here, we equate interpretation with information. We refer to mis-information when talking about unconscious falsification of the semantics of an information and we refer to dis-information when a person has intended falsification in mind, e.g. in bad faith.
Definition: *Privacy-related context of integrity* (integrity context) is a set of relevant environmental properties describing the situation in which some information is generated.

In this regard, “relevant environmental properties” comprise such properties that are reasonably required to genuinely understand information.

### 6.2 Initialization

The question of how to initialize privacy-related contexts is an important one for disclosure contexts, in the first place. As disclosure contexts are used to find a reasonable trade-off between service consumption and data minimization, it is suggested that contexts are initially defined in such a way that no information or only absolutely necessary information is disclosed.

To illustrate the issue by an example that is still rather rough: An initial context might be *<all communication>* associated with *<no disclosing of information>*. After working a while, the contexts become more differentiated by associating *<username IM>* and *<password IM>* to context *<IM>*. *<username e-Mail>* and *<password e-Mail>* to context *<e-Mail>* and so on. In the same manner, several contexts within one application could also be differentiated, e.g., activities like browsing an e-shop, creating a wish-list in the e-shop, buying products in the e-shop, and creating reviews of products bought in the e-shop. In such a case, browsing the e-shop could be carried out fully anonymously, i.e., the context *<browsing the e-shop>* is associated with *<no information to disclose>*. The creation of the *<wish-list>* has to be associated with a *<wish-list name>* of the user. If others should be able to access this wish-list, the name should be chosen in such a way that those others are able to find it when searching. The context *<buying in e-shop>* requires indicating the information necessary for the creation of the contract of purchase. These are, e.g., *<name>, <first name>, <address>,* and payment information.

Defining integrity contexts coincides with the occurrence of information disclosure. The parties involved have to agree on a specific set of environmental properties (types as well as values) that they consider as comprehensively describing the context of the information.

### 6.3 Dynamics and Statics of Contexts

Integrity contexts must not allow for contexts to be changed as the contrary case would carry the motivation for contextual integrity of an information ad absurdum. So once a context description has been created and mutually agreed on by all parties involved, it has to stay the same for ever. In order to assure that unnoticed changes of integrity contexts are not possible, it is recommended that all the parties involved digitally sign the context description of the disclosed information. For long-term security, signed integrity contexts might be stored in trusted archives that time-stamp each information that they store.

In comparison to this, disclosure contexts are not required to remain unaltered. As our lives can change in an instant and our relationships to others gain different qualities, the underlying patterns (i.e., the disclosure contexts) for decision making will have to be adjusted as well. For example, let us consider a project team needs to be created. According to a model of group development, there are five stages (forming, storming, norming, performing, adjourning) in which the kind of relationships and interactions between the involved parties changes [10]. In such a case, the actual context would remain the same. However, its attributes change, i.e., during the forming stage, members of the team try to disclose as little information about themselves as possible. The storming and norming stages are those phases where the team members negotiate with each other about what to disclose to the others whereby rules are also defined indicating which additional information should be disclosed during the performing phase to the other members of the team as well as to “outsiders”. The adjourning stage can be understood as a negotiation phase in which information exchanged during the project can be re-used in further projects (and contexts, respectively).

There are different types of changes of disclosure contexts:

- **Refine a context description:** A context description is adjusted to the new circumstances of interaction.
- **Create a sub-context:** A context can be refined by adding a sub-context in cases where it is necessary for the original (more general) context description to still play a role, but additional conditions might appear leading the new sub-context. E.g., a user discusses particular issues of recent classes with a group of students. This situation may find itself in the user’s context *<class chat>*. But, if the tutor of the class joins in, a new sub-context, e.g., *<class chat + tutor>* is entered.

- **Create a new context:** Some changes in circumstances may lead to the creation of a new context. In this case, the original context remains and a new context is created and described.

As seen, the approaches of using contexts for managing an individual’s privacy vary in different characteristics. Further, as the ideas described are of rather of theoretical nature, they open up lots of new research questions to be answered by different disciplines. Examples of such research questions are: What are the long-term implications of linking integrity contexts to disclosed personal data? Do the cultural backgrounds of the human beings involved play a particular role when they use integrity and disclosure contexts?
7 Conclusions and Outlook

Considerations of the historical evolution of privacy indicate some rather sceptical results. Some people even gave up caring for their privacy all together. Others, especially young people confronted with the benefits of being members of social networks, did not have the opportunity to protect their privacy or have not yet learned what it means to expose personal data over the Internet. However, the elaborations on privacy supported by a three-component approach have shown that it is possible to maintain the interests of the individuals in maintaining their privacy, on the one hand, and utilizing services, on the other.

Having described the theoretical basis for our ideas, this opens up further research questions concerning consequences for the users with regard to acceptance and costs in terms of the way applications need to be designed, e.g. to support users in formulating, negotiating, and signing contexts. Additionally, side effects of such an approach have to be examined and considered when developing applications as well as human-computer interfaces. Further, interdependencies that may occur when applying individual components of the Privacy 3.0 concept need to be studied as well as the findings considered during the design of applications.

Acknowledgements

We would like to thank Marit Hansen and Diane Whithouse for their helpful suggestions to improve this article.

The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 216483 for the PrimeLife project. The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any particular purpose. The PrimeLife consortium members shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

References


Received: September 1, 2010