Lifelong Privacy: Privacy and Identity Management for Life

Andreas Pfitzmann, Katrin Borcea-Pfitzmann

Lifelong Privacy: Privacy and Identity Management for Life

Andreas Pfitzmann and Katrin Borcea-Pfitzmann

Technische Universität Dresden, Faculty of Computer Science
D-01062 Dresden, Germany
{andreas.pfitzmann,katrin.borcea}@tu-dresden.de
http://dud.inf.tu-dresden.de

Abstract. The design of identity management preserving an individual's privacy must not stop at supporting the user in managing her/his present identities. Instead, since any kind of privacy intrusion may have implications on the individual's future life, it is necessary that we identify and understand the issues related to longterm aspects of privacy-enhancing identity management. Only that way, according solutions can be developed, which enable users to control the disclosure of their personal data throughout their whole lives, comprising past, present, and future.
This paper will give a general overview about concepts supporting privacy-enhancing identity management. Further, it introduces the reader to the problem field of privacy management by means of privacy-enhancing identity management during various stages of life as well as in various areas of life. Statements about required mechanisms will be given as well as directions regarding the three most important aspects to consider when managing one's identities: communication infrastructure as well as selection of communication partners and tools.

Key words: Privacy, Identity Management, Lifelong Aspects, Stages of Life, Areas of Life

1 Introduction

When starting to talk about lifelong privacy[1], first we have to state that we're talking about a timeframe of nearly 100 years. Inclusion of genetics and children inheriting DNA codes from their parents into the considerations may even extend this timeframe essentially. To give a point of reference, the military would be quite happy if they could keep their secrets for about 30 years. So, what the researchers in the field of lifelong privacy are talking about is an extremely long time span.

[1] "Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." [Wes67]

During that timeframe, an individual's world changes a lot, i.e., information and communication technology develops (remember the changes in this area during the last 40 years, which were very impressive; an even more perceivable evolution is to be expected during the next decades), and each individual's appreciation of privacy will change several times in her or his life, too.

What is really hard and, to the authors opinion, not possible to achieve is to make data fade away. Each time a user is using the Internet, possibly, s/he creates lots of traces. What s/he cannot do is reliably cause data to be destroyed on other persons' or organizations' machines – be they smart phones, laptops, desktops, or servers. One particular copy of data can be deleted if the other person or organization cooperates. But nobody knows whether there are other copies somewhere on the Internet. The approach of this problem field is two-fold:

1. *Minimization of personal data* means giving *hiding priority over disclosing data* since "if data is given out, it is out". No-one can ever call it back. At this point, we have to admit that some applications would not work well with users not willing to share personal data. *Identity management* may prove to be the most important approach to cope with this dilemma. It provides a mindset, a means to support people in managing their personal data and in sharing data that they really want to share with those people they really want to share it with. Identity Management will be dealt with in the following section. That section will explain what identity as well as management of identities shall mean. Further, it will introduce means to make identity management privacy-enhancing.
2. *Long-term security* is the second means of enabling lifelong privacy. Thereby, information- theoretically secure cryptography should be used instead of comp-lexity-theoretically secure cryptography wherever possible. Information-theoretically secure cryptography, which is sometimes called unconditionally secure cryptography, provides secure crypto independent of the attacker's computing power and algorithmic knowledge, which may essentially develop further in, e.g., 50 years. Nevertheless, migration to platforms providing stronger security should be done when they become available.

A more detailed discussion about long-term security will not be given in this paper as it is explored to a large extent already (cf., e.g., [CGHN97]). In contrary to this, identity management with regard to long-term aspects can be considered as a rather new research area. Especially, preserving or – to soften that strong term a bit – managing privacy during such a very long period of time by means of identity management takes an interesting perspective on the topic. That is why the following sections will give a general overview on the concepts of *identity*, *identity management*, and how it can be used to support users in managing their privacy throughout their *whole lives* by considering different stages as well as different areas of their lives. We conclude this paper by summarizing those issues important to consider when managing one's lifelong privacy based on privacy-enhancing identity management.

2 Identities and Identity Management

When talking about the concepts of *identity* and *identity management*, the question "the identity of which data subject" [2] needs to be answered. Even if almost each person has in mind natural persons when referring to identities, this could also refer to the identity of legal persons or the identity of computers. The latter is true when a person (let's call him Bob) takes a computer (e.g., being a mobile phone) with him all the time. In this case, if Bob would allow others to have a location tracking service of his computer they could track where he moves. This little example very well shows the need for some identity management for computers acting in place of their owners as well.

The development of the entities being in the position to have identity characteristics during the next 50 years we assume as follows: while the number of natural persons will not change very much (at least in comparison to the other two kinds of entities) – it can be expected that the number of human beings will not exceed the limit of 10^{10} – the number of legal persons will essentially increase (about 10^{11}). The numbers of computers, however, will explode. We expect roughly 10^{14} computing devices in the year 2059.

2.1 Identity – What is it?

Identity is a concept that is less clear than most people would expect. So, it is more than just talking about *names*, which are easy to remember for human beings. Identity is also more than *identifiers*, which usually are unique in a certain context. And, identity is even more than being a means for secure authentication. (If looking into the longer timeframe, i.e., a person's lifetime, identifiers and means of authentication experience much more change than names.) So, identity as we understand it is:

> *Identity* primarily is a set of attribute values related to one and the same data subject.

Some of the attribute values of an identity may change over time. But, if we add a timestamp to each attribute value for which that attribute value is valid[3], then attribute values never change. And, following this train of thoughts, we can further state:

> An *identity* as a set of attribute values valid at a particular time can stay the same or grow, but never shrink.

[2] By data subjects we refer to entities being able to interact via communication infrastructures with other entities, i.e., natural and legal persons as well devices used to represent them in interactions. Sometimes, even sets of persons are called data subjects.

[3] A *valid* attribute value means that it is used to represent its holder in a given setting.

This is true both for a global observer as well as for each party (or set of parties pooling their information) interacting with the entity represented by the identity. Therefore, if an attacker has no access to the change history of each particular attribute, the fact whether a particular subset of attribute values of an entity is an identity, which sufficiently identifies its holder within a set of data subjects, or not may change over time. If the attacker has access to the change history of each particular attribute, any subset of attribute values forming an identity, which sufficiently identifies its holder within a set of data subjects, will form such an identity from his perspective irrespective how attribute values change.

Any reasonable attacker will not just try to figure out attribute values per se, but the points in time (or even the timeframes) they are valid (in). This is because such change histories help a lot in linking data and, thus, in inferring further attribute values. Therefore, it may clarify one's mind to define each *attribute* in such a way that its value(s) cannot get invalid. So, instead of the attribute *location* of a particular individual person, take the set of attributes *location at time* x. Depending on the inferences one is interested in, refining that set as a list ordered concerning *location* or *time* may be helpful.

Partial Identities. Having in mind that identities usually grow over time and, thus, the probability of identification of the entity within the given subset of entities usually grows as well, a solution is needed to get a way out of that privacy-related dilemma. The idea is to subset the identity of an individual, the result of which should be a possibly very large set of so called *partial identities*. Thereby, each partial identity may have its own name, own identifier, and own means of authentication. In a certain sense, each partial identity might be seen as a full-fledged identity of someone or something.

The question that has to be answered now is how the attribute values have to be subset in order to establish reasonable partial identities. Obviously, if subsetting is done badly it won't help out of the privacy-related dilemma and it only makes the life of the related person more complicated. So, the right tools have to be used and subsetting of one's identity has to be done in the right way. Then this does not only help the person whose identity is under consideration, but also the people communicating with her or him since partial identities should consist of only those attribute values, which are really needed within that particular relationship or context.

Figure 1 shows a snapshot of a person's possible partial identities in different contexts. The dark-grey areas represent different partial identities of a person being parts of the full identity of that person represented by the light-grey area. While one may assume that this identity as well as its partial identities are related to activities of the individual in either the online world or the physical world, activities may also spread to the respective other world. The authors even assume that it is really hard to say if there will be any differentiation between those two "worlds" in the next 50 or 100 years. Ambient intelligence and ubiquitous/pervasive computing might make the boundaries blur or even disappear. This means that differentiating between identity-related data of the online and of the physical worlds might not make sense anymore. To conclude,

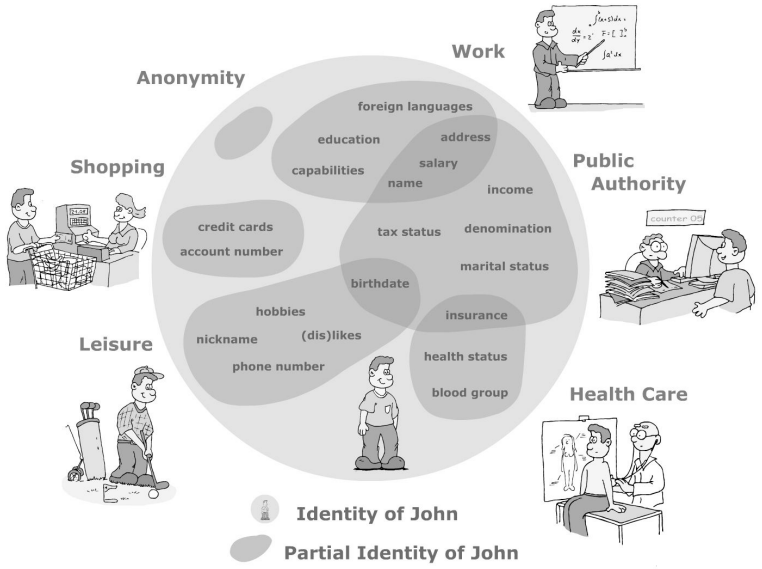

Fig. 1. Partial identities of an individual [HBPP05].

when looking into the future, subsetting the identity/ies is important whenever one strives for privacy.

Requirements for Using Partial Identities. Using partial identities requires a *basic understanding* by the data subject concerned. Of course, government and businesses have to understand it as well since managing one's (partial) identities makes sense only if the surrounding is willing to accept it.

Further, the authors assume that every person has at least one *personal computer* (or some device able to execute the according computations) administrating personal data and executing cryptographic protocols. Thereby, this personal computer is fully controlled by the user (otherwise there is no way to validate privacy properties).[4] The authors are fully aware of this today very daring assumption that all people have a computer being fully under their control. However, every time when people are talking about secure e-commerce they assume the same. So, since there are "major commercial forces" striving for that direction, it could be expected that the assumption the authors have made will become a more realistic one during the next 20 years.

By having a large set of (partial) identities, each of these (partial) identities needs its own means of authentication. Therefore, *digital pseudonyms* are needed

[4] In contrast to the requirement indicated here, whenever somebody talks about digital rights management (DRM) then usually having the user fully in control is not what they have in mind.

to fulfill the requirement for secure authentication (otherwise there is no way to achieve accountability). With digital pseudonyms we refer to bit strings, which represent unique identifiers of the respective (partial) identity and which are used to authenticate items originated by the holder in a way that recipients can check it (based on [PH09]).

Last but not least, *anonymous credentials*[5] are needed to transfer certified attribute values from one partial identity to another partial identity of the same identity. So, anonymous credentials are important because they are the basis for sharing authenticated attributes between partial identities of the same entity. Without anonymous credentials, the applicability of partial identities would be severely reduced.

Important Kinds of Attributes. When looking at attributes of (partial) identities, we can observe several kinds of attributes, each of them implying a particular degree of protection. Besides the already mentioned attribute types name, identifier, and means of authentication, we distinguish biometrics, addresses (used for communication), bank accounts, credit card numbers etc. used for – to a large degree – uniquely identifying entities. Biometrics as one of these represents a well-known concept of the physical world used for identifying persons for hundreds of years. However, biometrics being stored and evaluated by computers is relatively new. Biometrics can be helpful to bind computing devices to a natural person. But, it can also by critical if it is used in contradiction to privacy attitudes of people. When considering long-time aspects, the authors expect a lot of change of identifiers, of means of authentication, in the field of biometrics, and also of addresses.

With respect to classification of identity-related attributes, there are different possibilities:

- One of the main distinctions that can be made with respect to attributes is if they are *authenticated* at all. If so, then there are two possibilities regarding who did authenticate the attribute: First option is that they are authenticated by the first party – the data subject. In this case, it would be a claim the data subject makes about her/himself and the claim would be as trustworthy as the data subject is trustworthy. The second option refers to authentication by a third party. The authors explicitly did not refer to a

[5] The concept of anonymous credentials has been introduced by David Chaum in [Cha85]. According to him, a credential provides evidence of a statement about a particular property (attribute) of a data subject. This evidence is provided by an entity, i.e., the credential issuer, about another entity, i.e., the data subject, adding authentication by the credential issuer. If that credential is transferable between different digital pseudonyms of one and the same holder and using it with these pseudonyms does not prove the sameness of their holder, then it is called an *anonymous credential*. Anonymous credentials can be brought in different representations and used towards different parties. If anonymous credentials are issued to several users, they provide a good level of privacy among those users sharing the same attribute in a certified way.

trusted third party. So, it should be quite natural to ask: The third party is trusted *by whom* and with respect *to what*?
- Another approach of classification refers to *who knows* the attribute value, i.e., is the attribute value known only to the first party (the data subject) or also to second parties (the data subject's communication partner)?
- Attributes can be classified according to the *degree of changeability*. Could attributes values be changed easily or is this hard to do?
- *Variability* of attributes over time is also a possible classification whereby this could range from non-varying to fully varying. In this context: Can changes of attribute values be predicted?
- Attributes can be distinguished according to *who defines* the attribute values, i.e., are the attribute values given to the data subject by an external source or did the data subject her/himself choose the attribute values.[6]
- Another classification of attributes could be the actual *information* it contains. So, are we talking about *pure* attributes whereby the attribute values contain only information about themselves, or do the attribute values also contain significant side information?[7]
- Further, attributes can be classified according to the *relationships* the data subject is in. So, one could ask if an attribute value characterizes a single entity per se or an entity only in its relationship to other entities, e.g., entity `A` likes/loves/hates entity `B`.
- *Sensitivity* of attribute values in particular contexts can be seen as an additional means to classify attributes, though this might be a very subjective approach. However again, if considering long-term aspects, then attributes judged to be non-sensitive today, may become quite sensitive in future times (just think of a possible change of the social order).

From those approaches of classification, the question can be drawn regarding how much protection attributes or attribute values, respectively, need. Supposedly, some attribute values need much more privacy protection than others, e.g., those which

- are not easy to change,[8]
- do not vary over time or can be predicted,
- are given attribute values,
- might contain significant side information,[9] or

[6] To give an example: if we refer to the attribute *color of hair* then its value can be a given (natural hair color) or a chosen (after chemical dyeing) attribute.

[7] Let's assume we use biometrics, i.e., an image of someone's face available in a high resolution. From this, some doctors possibly may conclude some diseases.

[8] To give an example for the necessity to protect those attributes, think of some biometrics gets to be known widely. Then, it might become necessary, but be very hard to change that biometrics (which could mean, e.g., handing out new fingerprints to everybody). In comparison to that, cryptographic keys can easily be revoked and new ones generated.

[9] Nobody knows which algorithms for analysis of side information will become available during the next years.

– are sensitive or might get sensitive, respectively, in at least one context.

These attribute values are part of the *core identity*. Of course, it would be nice to protect everything. But, to be realistic, this is almost not possible. So, whenever starting to manage identity attributes, one has to think what defines her or his core identity, i.e., what attributes really belong to that core identity and need, therefore, according protection. Advancements and use of technology may shift some attributes from core identity to non-core identity. E.g., the address of your house or flat is core, the current address of your laptop maybe not.

Biometrics – the extraordinary identity attribute. Biometrics has already been mentioned in this paper several times. But, since it is an eternal core-identity attribute, it represents the most important example for an attribute requiring outstanding protection. Pfitzmann discussed the issue on "How to (not) use biometrics" in quite a detail in [Pfi08]. The main statements that have been made in that article relate to the following: Biometrics represents a really good concept if it is applied between a personal computing device of the person owning the biometric attribute value(s) and that person only. But it implies serious problems with regard to privacy if it is applied between, e.g., some kind of border control computer, which the person has no control over, and that person. The use of biometrics is, therefore, advised under the following conditions only:

– Biometrics is applied between a person and her/his devices only;
– Authentication is realized by possession and/or knowledge *and* biometrics;
– Classic forensic techniques are not to be devaluated (e.g., by foreign devices reading fingerprints, digital copies will make it into databases of foreign secret services and organized crime, enabling them to leave dedicated false fingerprints at the scenes of crime);
– Privacy problems by side information must be prevented when using biometrics (e.g., biometric-related measurements may also contain medical or psychological side information).

Since the safety problem remains unchanged by using biometrics between a person and her/his devices only, a possibility needs to be provided to switch off biometrics once and for all after successful biometric authentication.

2.2 Identity Management - How it Works

Identity management typically is not only between a person and its personal computer, which would imply some kind of authentication. But usually, identity management is applied within interactions between several persons and/or organizations.

Figure 2 demonstrates an example scenario where a person wants to do business with an organization. The typical data flow is as follows: The person uses a laptop. For authentication with her/his laptop, the person can use possession of the laptop, passwords, physical tokens, or biometrics. The laptop communicates to some infrastructure using addresses and encryption. This forwards the communication content to an end device within the organization.

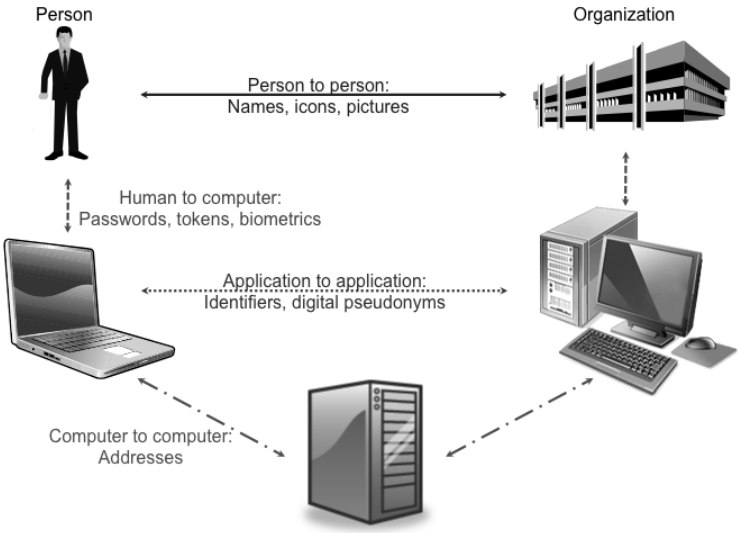

Fig. 2. Data Flows of Identity Management.

However, such data flow is not what the person is really interested in. S/he wants to do a person-to-person communication by using names, icons, or pictures. Since interaction is mediated by a computer-based infrastructure, application-to-application communication is required. At that level, cryptographic authentication is applied using identifiers, digital pseudonyms etc. So, whenever someone talks about digital pseudonyms, s/he is talking about computer-to-computer communication; it does not imply scenarios where human beings talk to each other addressing each other directly using digital pseudonyms.

An architecture of identity management looks like shown in Figure 3. Accordingly, a *user* communicates with a *service provider*. They use a *secure channel* for their communication. On each side, a component providing identity management (including authentication) functionalities is executed. For certain reasons, they may need services provided by so-called *Trusted Third Parties* (cf. our statements with respect to "trusted" third parties on page 7), e.g., identity brokers, PKI service, certification etc.

2.3 Presentation of Identities – Pseudonyms

Considering the use of partial identities in particular, one has to be aware that, first, partial identities have to be consciously created and established; and, second, the usage patterns of the partial identities[10] drive the kind of linkability of the attribute values and, thus, the conclusions that could be inferred. This

[10] When referring to *usage patterns of partial identities*, we address different aspects, e.g., how frequently a partial identity is communicated; how fine-grained is the con-

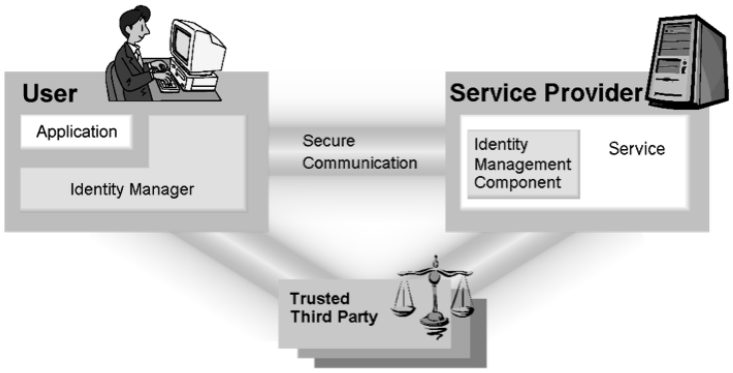

Fig. 3. Architecture of Identity Management.

means that users should do some partitioning of online activities according to contexts – so called context management [BDF+05].

Identities or partial identities of an entity are represented using (digital) pseudonyms. Those are used as identifiers of the (partial) identities, on the one hand, and as addresses of the (partial) identities, on the other hand. In order to indicate holdership of a (partial) identity, an explicit link between the pseudonym and the holder of the attributes of that (partial) identity needs to be created. Thereby, different kinds of initial linking between a pseudonym and its holder can be distinguished:

- *Public pseudonym*: The linking between a pseudonym and its holder may be publicly known from the very beginning, e.g., the phone number with its holder listed in public directories.
- *Initially non-public pseudonym*: The linking between pseudonym and its holder may be known by certain parties (trustees for identity), but is not public at least initially, e.g., a bank account with the bank as trustee for identity.
- *Initially unlinked pseudonym*: The linking between pseudonym and its holder is – at least initially – not known to anybody (except the holder), e.g., biometric characteristics such as DNA (as long as not in some register).

As already mentioned, according to the usage patterns of using partial identities and, connected to them, their pseudonyms, various types of pseudonyms can be distinguished. That differentiation of pseudonyms is closely related to different levels of anonymity that are achievable by the usage patterns.

Figure 4 illustrates that interrelation. According to this, *person pseudonyms*, i.e., names or identifiers directly identifying a real person, imply the lowest degree

text defined in which the partial identity is used; what rules are applied when selecting a particular partial identity for an interaction.

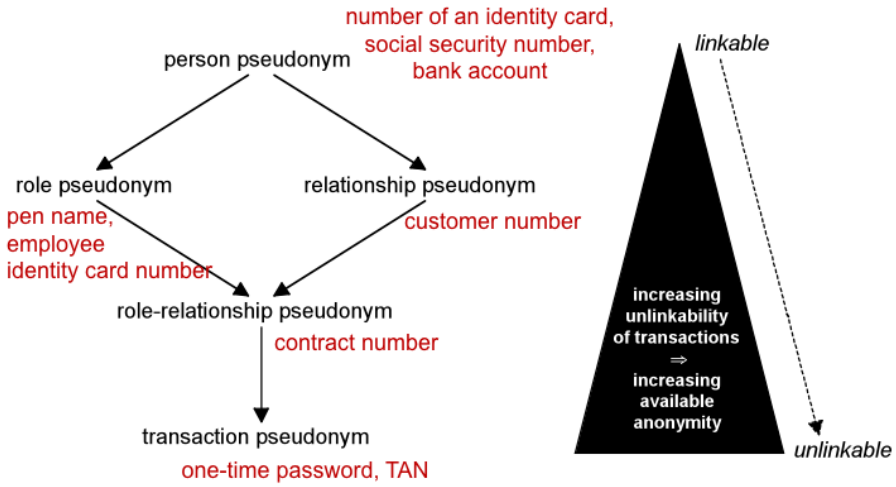

Fig. 4. Pseudonyms – Use in different contexts leading to partial order [PH09].

of anonymity. Examples for such kinds of pseudonyms are numbers of identity cards or the well-known social security number, which are used with very diverse communication partners and in very manifold contexts. Further, they typically are associated with their holders over their whole lifetime. This means, each time a user communicates by indicating her/his person pseudonym, all of the person's activities could potentially be linked together. As a result, a quite detailed profile describing that person could be created.

In comparison, *role pseudonyms* and *relationship pseudonyms* are pseudonyms used within particular contexts only. Thereby, a role pseudonym is used by its holder when acting in a certain role. An example for role pseudonyms are pen names. Similar to role pseudonyms are relationship pseudonyms. Those refer to entities within particular relationships, e.g., a pseudonym denoting someone in his or her relationship to a sports club. In this case, it does not matter if the person represents him- or herself as a trainer or as an athlete. So, the two pseudonym types are distinguished according to the following rules: Whenever a pseudonym specifies a person communicating *with* specified other entities, then we speak of a relationship pseudonym. Instead of this, if users specify *as what/whom* they communicate, then they are using role pseudonyms. Linkability is, therefore, restricted to the activities performed within the given relationship or when acting in a particular role and using the according pseudonym.

Even more privacy in terms of anonymity can be reached with help of role-relationship pseudonyms. The increase of conditions, i.e., used in a particular relationship while appearing in a special role, narrows the variety of a scenario where one and the same pseudonym is used essentially down. So, more role-

relationships (and, connected with them, partial identities) have to be created for more specific contexts.

If the goal is to get utmost anonymity when communicating via a computer network, one should make use of transaction pseudonyms. So, individuals benefit from the one-time use of those transaction pseudonyms. Linkability of different actions of the pseudonym holder via the pseudonyms only is not possible any longer since the user would create a new pseudonym for each interaction that is visible outside the user's personal computer.

The classification as given above is a rather rough means to contribute to tool development supporting the user in decision making with respect to the selection of pseudonyms or partial identities, respectively.

3 Identity Management throughout Life

This section, we would like to start by summing up what the previous sections comprise:

> *An identity management system has to be the communicational gateway of its user to her/his outside world.*

So, the previous sections provide the basis and give necessary information to build on towards identity management throughout life. And passim, the authors already pointed to aspects important for considerations of long-time aspects.

3.1 Identity Management Spanning Areas of Life and Stages of Life

Identity management has to be supported by an identity management system, which needs hardware and software interfaces to, of course, legacy systems, but also to emerging systems. Thereby, the users need to be aware that their identity management systems as a hardware/software implementation will change throughout their lives several times. Further, people's attitudes regarding privacy will change, too, as all individuals run through various phases of life and are related to different areas of life.

Figure 5 is a try to depict disclosures of personal data during an individual's lifetime, which has been sketched in [CHP+09,HPS08]. Usually, even before a human being is born, a lot of personal data about the unborn child is gathered. Such gathering is continuing all the time during a human being's life. The data is stored with various data controllers involved. And, if this is being done well, thereby it is partitioned into various partial identities: Each data controller should know only one partial identity of the human being. It would be even better if, even with respect to the same data controller, one has several distinct partial identities for distinct purposes.

Looking at a particular partial identity, there is a starting point where the partial identity is being established (in Figure 5, marked by "Establishment"). It evolves by either the person concerned adding data or by others appending data to that partial identity (in Figure 5, designated by "Evolvement"). And, finally

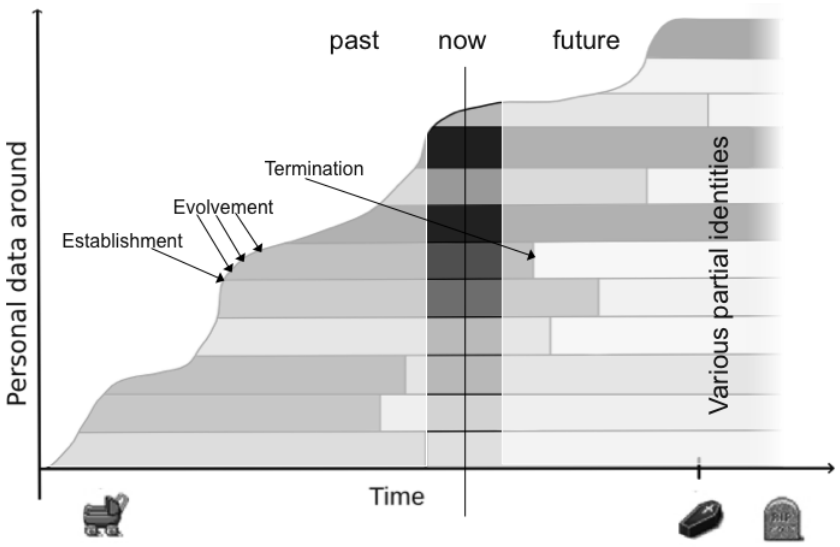

Fig. 5. Example of how partial identities develop throughout life and in various areas of life (based on [CHP+09]).

it is terminated (in Figure 5, this is labeled by "Termination"). But, termination does not mean that the data disappears. In many cases the data will be stored further, e.g., in backups. So, data will stay for quite a long time. Assuming the person died in the moment where Figure 5 shows a coffin, still some data will be stored even after the funeral for quite a long time.

If looking at identity management today, usually it covers a short timeframe only. It takes some history of the past into account and, depending on the attitudes of the user and the possible settings of his or her identity management system, it also looks a little bit into the future. What is actually required for comprehensively managing (partial) identities is a perspective taking into account the whole past and as much of the future as possible. Making a long story short, privacy throughout life means:

- covering the *full lifespan* by considering short-term as well as long-term effects;
- covering *all areas of life* by addressing context-specific as well as context-spanning aspects;
- covering *different stages of life* by respecting constant as well as changing abilities or behavior, respectively, of individuals.

When talking about *areas of life*, formal and informal areas are addressed. In formal areas, i.e., government, education, work, and health care, people have to participate whereas in informal areas, i.e., family, friends, shopping, and church,

one may choose whether to participate or even others decide for the person, respectively.

A *stage of life* of an individual with respect to managing her/his privacy is a period of life in which his rights and abilities to do so remain between defined boundaries characterizing this stage of life. A concrete stage might be defined in different areas of life differently, e.g., in Christian churches, a young man becomes adult after his confirmation (typically at age between 12 or 14) whereas for the point of view of a national government, a young man becomes adult when reaching a certain age (usually 18). Typical formal stages of life are nonage, adulthood, and retirement (cf. Figure 6).

3.2 Delegation as a Means to Overcome Issues Related to (Dis-)Abilities

One important point has to be made when talking about managing one's privacy during one's whole lifetime: The ability of an individual to manage her/his privacy during that time is not constant, cf. Figure 6. Starting with the stage of life which is called nonage, the right of the person to be heard usually grows because the ability[11] to manage her/his own privacy increases. When the person arrives at adulthood, s/he gains full responsibility over her/his life and, thus, over her/his privacy management.

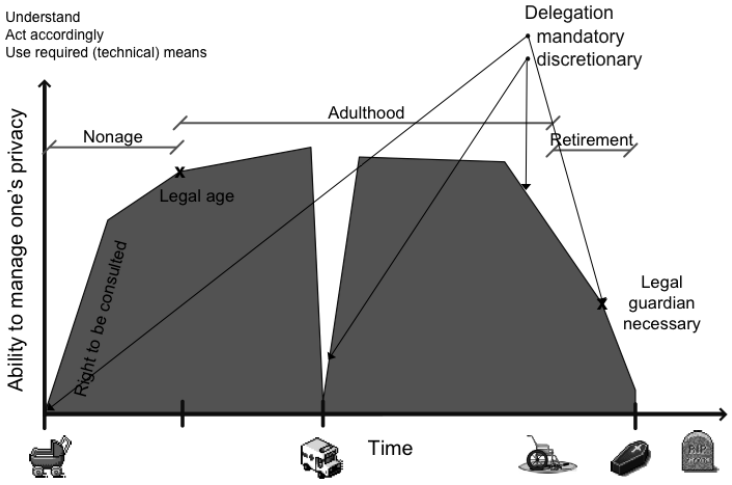

Fig. 6. Ability to manage one's private sphere during an individual's lifetime (based on [CHP+09]).

[11] The semantics of "ability" is: understanding a situation, to act accordingly, and to use the required (technical) means.

The case of accidents, temporary hospitalizations etc. may severely reduce the capability of handling one's data and, related to this, the capability of managing one's privacy. If s/he does not pass away before, each human being will get old (stage of retirement), which again may start reducing her/his abilities including the ones to taking responsibility of her/his privacy. In an extreme case, a legal guardian is needed. So, when looking at the curve in Figure 6, some people will start experiencing a loss of ability with respect to managing their privacy. At the end of their lives, this may lead to an ability similar to the one of children or even of babies.

In result of the considerations regarding the different levels with respect to the ability to manage one's own privacy, a very important concept has to be taken into account: *delegation*. Rights of the child are delegated by law – usually to their parents. This is very similar to situations where people need legal guardians and their rights are delegated to their legal guardians, which, of course, has to include rights and duties concerning privacy. So, the parents should take care of the privacy of their kids as legal guardians have to do so.

Delegation may happen in a mandatory fashion imposed by law. And, it can be realized in a discretionary manner. To give an example for the latter: A person is waking up in a hospital after experiencing an accident and having to undergo an operation. That person may decide for her/himself that s/he will need another person (her/his surrogate) temporarily taking over authorities and duties of the patient. Also, in case of old age, it could be a good decision that the person concerned defines a surrogate who will take over responsibilities by the time when s/he is losing the according abilities.

3.3 Mechanisms

The prime technological concepts and mechanisms required to realize reasonable lifelong privacy based on identity management have been introduced in Section 2. The following list summarizes what concepts and mechanisms are already available:

- Much theoretical work exists describing how to *handle partial identities*.
- How to *minimize personal data* is also known in principle.
- *Enforceable rules for data processing* and how to handle them are known as well.
- Further, researchers elaborated and are still working on several kinds of *transparency functionality* including how to check and inspect computers.

Even though many of the indicated concepts are well-elaborated and studied in theory, it is still an open issue putting them into practice. In addition, there are open issues for research and discussion, which need development effort and/or adaptations:

- As many *areas of life as possible and sensible* have to be covered.
- The *variances of stages of life* have to be regarded.
- Elaborations have to address the *full lifespan* when developing support for the management of an individual's privacy.

4 Conclusion

The authors do not claim having provided *the* solution solving all the problems connected with lifelong privacy. Instead, the problem field has been framed and known solutions have been described.

However, this work has shown that managing one's lifelong privacy affects many aspects people are not yet aware of. So, one of the most important tasks of researchers, educators, government etc. is to tell people to be attentive to managing their identity implying the management of partial identities. Otherwise, others will manage them – in a way that might not foster their privacy.

While managing their privacy, people should find a compromise, which they consider being right, between their desire or need to interact and their privacy. For this, the following three issues have to be considered:

1. Finding a good compromise based on human (subjective) decision is and will be an issue as long as the according *tools* helping to achieve the required compromise are not available. Several project groups are currently working on such tools, e.g., PRIME[12] as well as its succeeding project PrimeLife[13].
2. Besides the mentioned tools, an according privacy-preserving *communication infrastructure* is needed. Such an infrastructure has to prevent attaching permanent identifiers to the communication partners (e.g., network addresses). I.e., if all communication activities of an individual use the same network address then the network address is a globally unique identifier allowing to link all these activities together and, thus, also to link all partial identities. That kind of communication infrastructure would make privacy-enhancing identity management at the application layer void.
3. Finally, the right *communication partners* have to be chosen, i.e., avoiding those which are unnecessarily privacy-invasive. Communication partners labeling themselves as trustworthy should cause quite some distrust with the privacy-aware individual.

Acknowledgments. We like to thank the attendees of the keynote talk related to this article for their very helpful comments. 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.

References

[BDF+05] Katrin Borcea, Hilko Donker, Elke Franz, Katja Liesebach, Hilko Donker, and Hagen Wahrig. Intra-Application Partitioning of Personal Data. In *Proceeding of the Workshop on Privacy-Enhanced Personalization (PEP'05)*, pages 67–74. UC Irvine Institute for Software Research (ISR), July 2005.

[12] https://www.prime-project.eu/
[13] http://www.primelife.eu/

[CGHN97] Ran Canetti, Rosario Gennaro, Amir Herzberg, and Dalit Naor. Proactive security: Long-term protection against break-ins. *RSA Laboratories' CryptoBytes*, 3(1):1–8, 1997.

[Cha85] David Chaum. Security without Identification: Transaction Systems to Make Big Brother Obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

[CHP+09] Sebastian Clauß, Marit Hansen, Andreas Pfitzmann, Maren Raguse, and Sandra Steinbrecher. Tackling the challenge of lifelong privacy. In *eChallenges*, October 2009.

[HBPP05] Marit Hansen, Katrin Borcea-Pfitzmann, and Andreas Pfitzmann. PRIME - Ein europäisches Projekt für nutzerbestimmtes Identitätsmanagement. *it - Information Technology, Oldenbourg*, 6(47):352–359, 2005.

[HPS08] Marit Hansen, Andreas Pfitzmann, and Sandra Steinbrecher. Identity management throughout one's whole life. *Information Security Technical Report*, 13(2):83–94, 2008.

[Pfi08] Andreas Pfitzmann. Biometrics – how to put to use and how not at all? In *TrustBus '08: Proceedings of the 5th international conference on Trust, Privacy and Security in Digital Business*, pages 1–7, Berlin, Heidelberg, 2008. Springer-Verlag.

[PH09] Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management. https://dud.inf.tu-dresden.de/Anon_Terminology.shtml, December 2009. v0.32.

[Wes67] Alan F. Westin. *Privacy and Freedom*. New York Atheneum, 1967.