Identity and privacy issues throughout life

Dobias, J.; Hansen, M.; Köpsell, S.; Raguse, M.; C Roosendaal, A.P.; Pfitzmann, A.; Steinbrecher, S.; Storf, K.; Zwingelberg, H.

Published in:
Privacy and identity management for life

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Dobias, J., Hansen, M., Köpsell, S., Raguse, M., C Roosendaal, A. P., Pfitzmann, A., Steinbrecher, S., Storf, K., & Zwingelberg, H. (2011). Identity and privacy issues throughout life. In J. Camenisch, S. Fischer-Hübner, & K. Rannenberg (Eds.), *Privacy and identity management for life* (pp. 87-114). Springer.

Chapter 4
Identity and Privacy Issues Throughout Life

Jaromir Dobias, Marit Hansen, Stefan Köpsell, Maren Raguse, Arnold Roosendaal, Andreas Pfitzmann, Sandra Steinbrecher, Katalin Storf, and Harald Zwingelberg

4.1 Challenges and Requirements

Much research and development has been done during the past couple of years to assist users in managing their partial identities in the digital world by several types of identity management [BMH05]. A comprehensive privacy-enhancing identity management system would include the following components [CK01]:

- an Identity Manager (IdM) on the user's side;
- IdM support in applications (e.g., at content providers, web shops, etc.);
- various third-party services (e.g., certification authorities, identity providers).

However, current concepts for identity management systems implicitly focus on the present (including the near future and recent past) only. The sensitivity of many identity attributes and the need to protect them throughout a human being's entire lifespan is currently not dealt with. The digital lifespan is the range of time from the emergence of the first information that is related to the human being until the point in time when no more personal data is generated: from the moment of birth until death. Hence, lifespan refers to the temporary aspects of privacy and identity management and, in particular, to the challenges involved in realising (privacy-related) protection goals over very long periods of time. The area of challenges regarding privacy is vast – even when not considering an entire lifespan (see, e.g., [Eni08]). In the following, we describe which additional problems occur concerning lifelong protection of individuals concerning their privacy in a technology-based society.

4.1.1 Dealing with Dynamics

Our society, as well as the individuals that form them, underly dynamics. We distinguish between *dynamics in the surroundings of the individual* and *dynamics in* *the individual's ability or willingness of managing her private sphere on her own* as outlined in the following subsections [CHP$^+$09].

4.1.1.1 Dynamics in the surroundings of the individual

The dynamics of the effects from the outside world – possibly affecting the individual's private sphere – comprise, among others, technological developments, replacement of administration, changes in law and policies, and – last but not least – the evolvement of society.

The least dynamics we have to deal with is the increasing processing of personal data during one's lifetime. This involves the disclosure of personal data to many Data Controllers, partially because the disclosure and processing of data is officially required (e.g., because of school attendance, tax liability), partially because the data are needed to fulfil tasks in the areas of e-commerce, leisure, communication etc. Fig. 4.1 shows a simplified model of increasing data disclosure to different Data Controllers, depicted by coloured stripes. The lighter colours on the right-hand side express that the personal data are not needed anymore for the task to be fulfilled, but the data may still live on in official archives, at Internet services, or in the storage of communication partners [MS09]. The data might not be deleted after the time of death nor after the funeral.

The coloured stripes in Fig. 4.1 might also correspond to several *partial identities* [HPS08], for example (but not exclusively) individuals in different areas of their life. Areas of life are sufficiently distinct domains of social interactions that fulfil a particular purpose (for the Data Subject) or function (for society). Formal areas of life include education, work, and health care. Informal areas of life cover mainly a user's social network including family, friends, leisure, religion etc. Some of these informal areas might become formal by institutionalisation, e.g., for religion in the form of membership in a church.

Another dynamic results from the development in technology including possible risks. The technological progress of the last decades triggers the transformation of our society towards a computerised social community highly dependent on information. The more structures of our society depend on information, the more important is the role that data plays in our everyday lives. During decades of technological evolution, several methods were invented for storing data in various forms and on various types of media. However, the processes and events in the nature, society, and in the life of the Data Subject cause failures, which might lead to loss of the data during the lifetime of the Data Subject. Even if unwanted data loss might not be a common phenomenon encountered within the life of every Data Subject, it may become an evident and serious problem, which emerges in the lifelong extent of time.

Privacy becomes an increasing problem, e.g., unauthorised access to personal data which enables attackers to read them, link them with other data, or modify them. Personal data, which are assumed to be secure at a specific point in time, may be at risk after some time if no additional precautions have been taken.

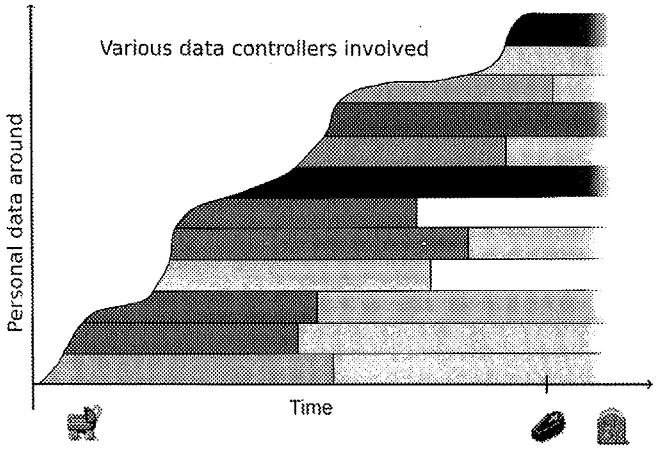

Fig. 4.1: Accumulation of personal data and formation of partial identities

Also, in the political and societal area, dynamics are to be expected during a period of several decades. In particular, it is not predictable how people in a leading role will interpret today's personal data or which policies will apply.

Lifelong privacy mechanisms need to cover not only the near past and future of an individual, but also need to consider the future prospects for a human's lifetime and also beyond, which means about 90 years. The problem is that we only have experience with computer-based technology for about half of this time, and with the growing exchange of computer data over the Internet, even less than a quarter of this time. This means that all privacy mechanisms (including public-key cryptography invented in 1976 based on cryptographic assumptions) could not be tested in practice for a whole lifetime yet. For the selection of privacy technology (hardware and software), attention should be paid to the following aspects:

- The duration of cryptographic security (based on cryptographic assumptions) should be at least an individual's lifetime. If unconditional security (not relying on cryptographic assumptions, but just on pure probability theory) is possible and feasible, it should be provided.
- Migration of data between different hardware and software needs to be assured. When considering the long-term risks in an unpredictable setting, the sensitivity of personal data is of utmost importance. For the choice and protection level of personal data processed, a categorisation regarding their sensitivity with respect to privacy has to be made according to [HPS08, CHP^{+}09].

4.1.1.2 Dynamics in the individual's ability or willingness of managing her private sphere on her own.

During their lifetime, individuals pass through different stages. A stage of life is defined as follows:

> *A stage of life of an individual with respect to handling her privacy is a period in her life in which her ability to manage her private sphere remains between defined boundaries characterising this stage of life.*

The management of one's private sphere comprises the ability to understand privacy- and security-relevant aspects concerning one's private sphere, the ability to (re-)act accordingly, and the ability to use appropriate (often ICT-based) means for one's (re-)actions. Obviously toddlers cannot manage their private sphere on their own, nor can people suffering from pronounced dementia or those in a coma. Even for those who are mentally able to manage their private sphere, it may not be feasible if it requires using technical devices.

Three large stages of life that individuals typically run through are childhood, adulthood and old age, which are depicted in the example shown in Fig. 4.2. It is quite clear that a baby is physically less able than a 10-year-old to interact with technical devices. So the ability of a child to manage her private sphere and her right to be consulted usually increase with her age. However, at least small children are not able to decide on their own how their data are created and processed and how their private sphere can be controlled. Also, adults may have temporary or permanent needs where others support them or even act on their behalf concerning decisions concerning their private sphere. This is true especially in old age. For small children, as well as for very old people, and in the case of emergency, delegation of the right to manage one's private sphere is needed. For children, these delegates are automatically their parents; in case of emergency or for old people it might be a close relative.

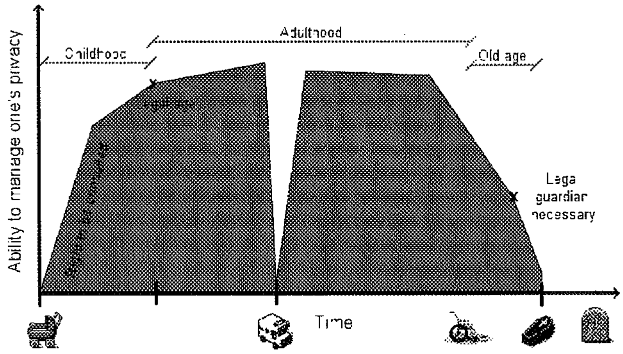

Fig. 4.2: Stages of life: variation in the ability to manage one's privacy

Ability in the legal sense would not be partially linear functions as in Fig. 4.2 but more step-like and then remain constant until legal guardianship is needed again.

Sometimes, individuals who in principle are able to manage their privacy on their own want to involve other parties or delegate their privacy control, e.g., for convenience reasons. These individuals may have the ability, but not the willingness to manage their private spheres on their own. Furthermore, both the ability and the willingness might change depending on the circumstances an individual is in and especially depending on what possible Data Controllers offer to her.

A privacy-enhancing identity management system should support the delegation of duties and authorities. There are three possible situations that might occur regarding delegation from the legal perspective: Firstly, delegation of rights might be made by law automatically for a certain time frame (e.g., for children to their parents). Secondly, delegation might be made willingly by an individual to others for a certain time frame (e.g., delivering mail to others during holidays). Thirdly, delegation of rights of an individual might be initiated by other individuals to achieve delegation of her rights to them or others (e.g., in the case of incapacitating a person), which presumably requires thorough juridical investigation before divesting the person of a right. Delegation can be implemented by different means. Usually, the delegate does not take over the identity of the individual concerned, but receives authorisations to act – often within defined ranges – on behalf or as inheritor, respectively. Technical processes for delegation and digital estate have to be defined in accordance with legal procedures. We come back to this issue in Section 4.1.3.

4.1.2 Digital Footprint

The difficulties regarding partial identities and identity management in digital environments have another important complicating factor. Next to the partial identities consciously created by individuals to perform actions on the Web, huge amounts of data are collected just because of web activities or other electronic transactions. Every interaction with a communication device, consciously or unconsciously, leads to a data log; a digital trace. By accumulating these data and making connections between the data, extensive digital footprints of individuals can be created.

Digital footprints are data that accumulate in information systems and can be indicated as belonging to one individual. This means that data in this sense is not restricted to personal data only. We outline in the following which data can be sensitive and how it can be linked.

4.1.2.1 Sensitive Attributes

Some attributes and attribute values usually need more privacy protection than others. According to [HPS08, CHP^{+}09], we distinguish the following properties of identity attributes, which, alone or in combination, pose specific risks to privacy when being disclosed:

- *Data may be static, or changes are quite accurately predictable:* Data which are static over time and are disclosed in different situations enable linkage of related data. Examples for static data are date and place of birth. Similar to static data are those which are quite accurately predictable or guessable because they follow some rules. Examples are data following mathematical rules like the number of children that will only remain or increase. If static identity information is being used for purposes such as authentication, this bears a risk because these data cannot easily be revoked and substituted: For example, the use of fingerprints with biometric access systems.
- *Data may be (initially) determined by others:* Data that the individual concerned cannot determine herself (e.g., the first name) may persist or it may take a significant amount of time or great effort to change them. A special case is the inheritance of properties from others, e.g., the DNA being inherited from the natural parents.
- *Change of data by oneself may be impossible or hard to achieve:* If data are static (see above) or if data are not under the individual's control, wilful changes may not be possible. Examples are data processed in an organisation.
- *Inclusion of non-detachable information:* There are data that cannot be disclosed without simultaneously also disclosing some side information tied to the data. Examples are simple sequence numbers for identity cards, which often reveal gender, birth data and at least a rough timeframe of when the identity card was issued.
- *Singularising:* If data enable the recognition of an individual within a larger group of individuals, the individual may be tracked or located, even if other personal data of the individual are kept private.
- *Prone to discrimination or social sorting:* There are no data that are definitely resistant against possible discrimination forever. This does not need the individual to be identified or singularised. If some people disclose a property and others resist to do so, this already allows for social sorting or positive discrimination.

Note that this list of sensitive properties extends the enumeration of special categories from Art. 8 Data Protection Directive ("personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life"). Because of the sensitivity of the listed personal data, everybody should be careful with related data processing.

4.1.2.2 Linking Data

When taking into account that each and every transaction in an information system leaves a digital trace behind, it becomes clear that the accumulation of data can be enormous and that it thus becomes easier to link data to a single individual. The more data available, the more unique the combinations of these data are. The more data disclosed to a party over time, the lower the fingerprinting threshold becomes, which means that after a certain amount of time individuals can be uniquely identified and recognised when they appear in a new transaction, even when this new transaction is done from another computer or another location than previous ones [Con09].

The possibilities of linking data that belong to one individual also have their drawbacks on the dynamics in the surroundings of the individual. At first glance, it might seem that changing surroundings can have a positive influence on identity management, because contexts can be separated relatively easily. Nevertheless, linkability appears possible when disclosing information or revealing certain patterns over time, meaning that the different surroundings can be connected as belonging to the same individual as well, therewith creating more challenges with regard to identity management. These challenges are even more supported by the fact that the Internet does not forget. There is no digital oblivion. Once data are revealed somewhere, they persist over time and remain stored in databases, cache memories and so on and so forth.

Obviously, to make the connection between different data sets, some kind of connector is needed. This connector can be created with the help of sophisticated technologies, analysing clicking behaviour, click trails, and key strokes. Linkability between data (sets) can be based on an analysis of the data and revealing patterns. However, this becomes easier when a common identifier exists within different data sets (such as names, registration numbers and IP addresses) and is easiest when this identifier is a so-called unique identifier. Unique identifiers identify a single individual. Unique identification can be supported by governments who issue unique identification numbers to their citizens. These numbers are sometimes even used throughout different domains or areas of life, thereby linking these areas, and the data therein, as belonging to one single individual. In the PrimeLife project, a number of European countries have been investigated on their use of unique identifiers in four formal areas (government, health care, education, employment) and their application range: In the Netherlands a unique identifier, the BSN, is commonly used in several settings which, in principle, allows for the construction of a compound identity, instead of the citizen having distinct identities in different areas. In Germany this is not the case. There is even a separation between the different federal states, which all have their own regime in certain contexts. However, all German citizens of age 16 and above are obliged to have a personal ID card, which is used for identification by public authorities. Up to now, this card only played a role in the physical world. Starting with the 1st of November 2010, the new German eID card will enable trusted services to read out selected attributes over the Internet. Consequently, the holder of the eID card maintains control over which provider is allowed to read out which attributes. Additionally, the new eID card would enable the user to establish a secure pseudonymous communication with a given service provider. Therefore a unique pseudonym will be generated for each relation between an eID card holder and a service provider. France and Austria also have ID cards, but no unique identifier that is used throughout public services. Belgium, Sweden, Ireland, and Poland do use unique identification numbers, often combined with ID cards.

In the four formal areas, some differences do occur. Some countries have national student cards, others do not. The same goes for electronic health cards. The Netherlands has no electronic health card, but is working on a central system, called EPD, which aims to improve health care by providing access to health records electronically. The system progresses towards a general comprehensive medical identity. In the area of employment, all investigated countries show centralised systems to register people that are unemployed.

4.1.3 Concepts for Delegation

In the context of identity management throughout life, one focus lies on investigating the necessity of delegation for people who are not able to manage their needs of privacy for a limited time within a stage of life or forever.

Delegation is a process whereby a delegate (also called "proxy", "mandatory" or "agent") is authorised to act on behalf of a person concerned via a mandate of authority (or for short: mandate) [HRSZ10].

The mandate of authority usually defines in particular

- the scope of authority for the actions of a delegate on behalf of the person concerned and
- when and under which conditions the delegate gets the power of authority to act on behalf of the person concerned.

The delegate shall only act on behalf of the person concerned if the delegate has the actual power of authority and if his action lies within the scope of authority. The simple acting of the delegate with the existence of a mandate while not having the power of authority would not be sufficient. The difference between mandate and power of authority becomes clear in the following example: In working life, the schedule of responsibilities may determine that person A should take over the work of colleague B if the latter is absent. The issuance of the mandate of authority to A is expressed by the schedule of responsibilities, but A's actual power of authority only comes into existence if B is absent. Otherwise A must not act on behalf of B.

The mandate of authority is issued by the delegator (also called "mandator"). This may be the person concerned herself, but there are also cases where other entities explicitly decide on the delegation (e.g., in the case of incapacitation of a person, the guardianship court rules on delegation) or where the delegation is foreseen in law (e.g., when parents are the default delegates of their young children). The mandate of authority is usually assigned for a specific period of time. Similar to the process of issuing a mandate, changing or revoking the mandate can be done by the delegator, i.e., by the person concerned herself or by other entities. The conditions and processes to issue, change, or revoke a mandate can be defined by the underlying contract or law.

Note that the delegate is not always aware of the mandate of authority or of the fact that he actually has the power of authority. So the delegator should implement an appropriate way of informing the delegate (and the person concerned if she is not the delegator herself) about the mandate and the power of authority.

For supervising purposes of the delegation and related actions by the parties involved, one or more impartial delegation supervisors may be appointed by one or more of the actors. In particular, the person concerned may have the need to check whether the delegate really acts as agreed upon.

The current civil legal framework encompasses several instruments regulating legal representation or agency, which have an effect with regards to the exercise of fundamental rights: For minors, the instrument of parental care is known in civil law. Most of the EC Member States also have legal regulations regarding the representation of children. The Article 29 Data Protection Working Party defined in its Opinion 2/2009 [DPW09] principles regarding children's privacy, which we generalise in the following to the relation of persons concerned and delegates regarding privacy-relevant actions:

- The delegate should act in the best interest of the person concerned. This may comprise protection and care, which are necessary for the well-being of the person concerned.
- Guidelines for delegation should be defined beforehand.
- The person concerned and her delegates may have competing interests. If conflicts cannot be avoided, it should be clarified how to sort them out, possibly with the help of external parties. Note that a delegate does not necessarily stand in for all partial identities of the person concerned, which may lead to additional conflicts of interest of parties involved.
- The degree of delegation should be geared to the capabilities of the person concerned regarding privacy and self-determination. This means that the degree of accountability of the person concerned has to be adapted over time, and regarding privacy-relevant decisions taken by the delegate, the person concerned has a right to be consulted.

Each stage of life has significant question on how to handle identity management and in particular personal data and therefore has different requirements. It is quite clear that a baby is physically less able than a 10-year-old to interact with technical devices. It appears that the privacy protection rights of an individual are exercised by different people during the lifetime. This asks for a delegation system where it is clear for all parties involved who can perform which rights at which moment and in which context. The consequences of the delegate's actions may both influence the privacy of the person concerned and the delegate herself to a certain extent. The following subsections explore various stages of life with respect to delegation.

4.1.3.1 Fruit of the womb

Privacy throughout life comprises a very early stage of life, the prenatal phase of an individual. Even in this stage of life, there might be the need to protect personal data, for example, considering the privacy implications of prenatal DNA tests. In many EU Member States, there are discussions about the issue of genetic analysis and the threat that using genetic data poses for individual's right of informational self-determination as well as potential discrimination. More detailed regulations regarding requirements for genetic analysis and the use of genetic data could be a solution.

4.1.3.2 Children and teenagers

Growing autonomy is an important issue in the protection of children's rights, in any area of law. The complexity of situations involving minors is based on the fact that children, despite having full rights, need a representative (delegate) to exercise these rights – including their privacy rights.

Data protection for children starts within the first days after birth and the processing and storage of birth data or medical data within the hospital. The protection of the personal data of children resides more or less in the responsibility of parents or legal guardians as delegates by issued by law. But when a child grows up, other responsible persons for data processing in different areas of life may become involved, such as teachers, doctors or supervisors [HPS08].

The rights of the child, and the exercise of those rights – including that of data protection – should be expressed in a way that recognises both of these aspects of the situation [DPW09]. Until a certain age, children have no way to monitor data processing, simply because they are too young to be involved in certain activities. If their delegates (parents or other representatives) decide, for example, to put the child's pictures on their profile in a social network, it is the delegate who makes the decision about the processing of the children's data and gives the consent to do so on behalf of the child. Normally, putting pictures of another person in a social network profile requires consent of that person, the Data Subject. In the situation described here, the delegate (e.g. parents) is entitled to express the consent in the name of the child. Such situations may put the delegate in the double role – of Data Controllers, while publishing their child's personal information open on the Web, and, at the same time, of consent issuers as the child's representatives. This double role may easily lead to conflicts. Parents must take great care not to cross the line of the child's best interest when processing the child's data.

It is necessary for the delegate (e.g. parents or other representatives) to listen carefully to the interests of the child at least beginning from a certain age and consider those interests when making a privacy-relevant decision, as that decision is binding for the child [DPW09]. When the child reaches legal age, it may want to change recent decisions of the delegate. Therefore the child needs to know what decisions about processing of personal data were made by the representatives. Afterwards, the child needs to give her explicit consent for the processing of personal data. This may be implemented in certain operations in a way that the operator is reminded that the person is over 18 and now the explicit consent is needed. This is relevant in many circumstances, for example, medical matters, recreational activities of the child, school matters, or agreements made by the delegate before the child's majority.

As children and teenagers are in the process of developing physically and mentally, the rights of the child and the exercise of those rights – including the rights of data protection – should be accomplished in a way that recognises these aspects of the situation. Especially the adaptation of the degree of maturity of children and teenagers is a central aspect that has to be taken into account by their delegates. Children gradually become capable of contributing to decisions made about them. It is natural that the level of comprehension is not the same in the case of a 7-year-old child and a 15-year-old teenager.[1] This, in particular, has to be recognised by the children's representatives. Therefore the children should be consulted more regularly by adults, teachers, care-takers or other delegates about the exercise of their rights, including those related to data protection.

The children's delegate should also think about a way to document privacy-relevant decisions so that the children or young adults can later easily understand what personal data have been disclosed to whom and under which conditions. They may also then choose to actively approach certain Data Controllers to give or revoke consent concerning data processing or to request access, rectification or erasure of their personal data.

4.1.3.3 Adults lacking privacy management capabilities

For adults that may have temporary or permanent needs to get support or requiring that others act on their behalf concerning decisions concerning their private sphere, the distinction has to be made between delegation for legally relevant actions and non-legally relevant actions. All legally relevant actions regarding the processing of personal data are based on national legal regulations such as delegation or legal guardianship.

In case of non-legally relevant actions, such as help with a social network or the Internet, in general the person concerned can freely decide what to do. The delegator or person concerned could choose a delegate (for example, a care-taker) to act in the name of the person on the basis of a contract to manage the private sphere. Then the person concerned should clearly define her expectations and needs regarding the representation and the power of disposal.

[1] The level of comprehension is defined in different ways. For instance the US-American Children's Online Privacy Protection Act (COPPA, Title XII – Children's online privacy protection, SEC. 1302) defines a child as an individual under the age of 13.

4.1.3.4 Deceased people

In situations where a person has deceased, the instrument of law of succession applies. The European Data Protection Directive 95/46/EC assigns the right of privacy and data protection to "natural persons" (Article 1). Deceased persons are no longer regarded as Data Subjects. Protection against unregulated processing of data concerning deceased individuals in some European legal frameworks[2] is provided by means of a "post-mortal personality right". In some situations, the instrument offered by the law of succession might not be sufficient – further regulations are needed.

For instance, some users of social networks want their profile to exist even after death or at least would like to be informed as to how the provider handles the personal data and the profile after death. Here, the action of providers of social networks is required to find mechanisms and concepts for handling profiles after the death of the user. Various mechanisms are conceivable, for example, the user could determine how her profile should be handled after death within the registration process (deletion, blocking, proxy to contact, etc.). Therefore, SNS providers need to define clear measures and concepts to determine the handling of profiles after one's death. In some situations, even the autonomous action of the SNS provider might be essential for the protection of users. For example, if a SNS user dies and the press accesses the SNS site to copy pictures, contacts, etc. of the dead user, the provider has to balance the protection of the users rights and her competence to, for example, block the profile without the consent of the legal assignee (because this has to happen very quickly).

Meanwhile, new services appear on the market that offer to send out secure messages to friends after the death of the user. Their goal is to give people a safe way to share account passwords, wills and other information. When users book the service against payment of a fee, they are given options for when to send messages or to delete some messages permanently after their death. It is problematic if authentication credentials of the user have to be transferred to the service, which opens the way for misuse because it is not distinguishable for others whether the user or the service acts.

4.1.3.5 Delegation based on explicit decisions of the Data Subject

The Civil law recognises the instrument of legal representation for cases where the concerned individual is fully in possession of her mental capabilities and decides on her own to transfer the exertion of rights to another person (for example, Articles 172 et seq. German Civil Code[3]). Various reasons exist why a Data Subject may wish to transfer the full or partial legal authority of representation to another indi-

[2] Such as Germany: so-called "Mephisto decision" of the German Constitutional Court; BVerfGE 30, 173.

[3] English translation of Bürgerliches Gesetzbuch: (German Civil Code): `http://bundesrecht.juris.de/englisch_bgb/englisch_bgb.html`.

vidual (the mandate of authority). For example, a person may simply be unavailable for a longer period of time with no access to information technology, which would allow transmitting and enforcing remote decisions (for example, during a scientific or recreational journey to a secluded region). Or a Data Subject may feel that certain services which are handled online are better understood by friends or even a professional data custodian. Actions of and decisions by the delegate (authorised representative) may have consequences for the fundamental rights of the delegator. The delegator may have, at first glance, authorised the delegate to act on behalf via a mandate of authority, which, for example, only granted authority to the delegate to close one contract on the delegator's behalf. Delivering the contractual duties, however, will possibly also require the processing of personal data. The legal authority to represent a person concerned in closing a contract does include the implied authority to initiate the data processing steps necessary to fulfill the primary goal. The instrument of legal representation based on the Data Subject's declared intention may also have an effect after the Data Subject's death. The Data Subject may during her lifetime lay down a last will which binds the heirs. This last will may also comprise decisions regarding how to treat documents or electronic files containing personal data.

The Art. 29 Working Party defined in its Option 2/2009 [DPW09] principles regarding exercising the right of children. These principles may also be helpful for determining principles on delegation in general, because delegators (persons concerned) may have the problem that delegation in privacy-relevant situations might be interpreted in different ways. This means that one may have different needs on good practice of handling privacy.

4.2 Demonstrator

A scenario that is relevant to all areas and stages of life and deals with dynamics and possible delegation - this means the challenges we described in the previous Sections 4.1.1 and 4.1.3 - is the area of backup and synchronisation tools and applications.

Many backup systems and backup strategies, which have been available for many years, are already dealing with the problem of unwanted data loss. However, they are mostly protecting the raw data only and do not involve the Data Subject, his specific characteristics, social relations and interactions as a part of their scope. Existing backup systems and backup strategies also do not reflect the process of evolution of the Data Subject during his lifetime with respect to the possible different states he might pass through during his lifetime and which might have an immense influence on his ability to manage his data on his own behalf (e.g., illness, hospitalisation, or death). Additionally, existing systems and strategies dealing with the problem of unwanted data loss do not cope with boundaries among distinct areas of the Data Subject's social interactions. However, these aspects are nowadays becoming more and more sensible on the level of the data, hand in hand with the massive expansion of the technology.

Therefore, we decided to analyse the problem of unwanted data loss from the perspective of lifelong privacy. We found that current solutions do not provide a sufficient level of data protection when it comes to lifelong extent of time and privacy of the Data Subject holding the data. Based on our findings, we decided to demonstrate that it is possible to cope with problems amplified by the requirements on lifelong privacy when protecting the Data Subject against unwanted data loss.

The proposed privacy-enhanced backup and synchronisation demonstrator focuses on the following problems closely linked together under the light of lifelong privacy:

1. Protection of the Data Subject against unwanted data loss during his lifetime by redundancy and physical distribution of the data;
   Our findings resulted in the conclusion that the problem of unwanted data loss can be solved by redundancy and the physical distribution of multiple copies of the data from the lifelong perspective. As far as backup and synchronisation tools are also dealing with the problem of unwanted data loss, we decided to establish the main conceptual pillars of our demonstrator on the backup and synchronisation functionality. In the demonstrator, we are proposing to solve the problem of unwanted data loss by taking advantages of services provided by online storage providers which are nowadays available on the Internet (for example Dropbox, Apple MobileMe, Windows Live SkyDrive, Ubuntu One and others) and store multiple copies of the data in a distributed environment. Distribution of potentially sensitive backup data in such kind of environment, however, leads to confidentiality problems.

2. Assurance of lifelong confidentiality of the Data Subject's data stored in a distributed environment;
   The problem of data confidentiality in a distributed and untrusted environment can be solved by the encryption of the data. Encryption must assure that only the authorised Data Subject (whom the data belongs to) is able to operate with his data stored in distributed backups by default and nobody else should have implicitly access to it even after the death of the Data Subject. On the other hand, during the lifetime of the Data Subject, unpredictable situations might occur, which might temporarily or permanently limit him in his ability to access his own data (for instance in case of his illness, hospitalisation or death). This might lead to situations that his data, which might be important for other parties relying on it (possibly in a legal relationship with the Data Subject), is not accessible by these parties when needed (for example important work documents) or is permanently lost.

3. Delegation of access rights to the Data Subject's backup data allowing other parties to operate with his data if specific conditions are fulfilled;
   delegation capability of the demonstrator allows other parties authorised by the Data Subject (whom the data belongs to) to access his backup data in case partic-

ular conditions specified by the Data Subject are satisfied. Delegation of access rights of the Data Subject's backup data could in general lead to situations where authorised parties with corresponding access rights are not only able to access the desired data but also other data possibly covering other areas of the Data Subject's life, which they are not authorised to access. This might, however, not be desired by the Data Subject himself.

4. Distribution of the backup data according to different areas of life of the Data Subject and his different partial identities.
   Distribution of the backup data according to particular areas of the Data Subject's life or his different partial identities enables the Data Subject to manage his privacy in such a way that allows him to physically and logically separate his data related to distinct domains of his social interaction.

Besides the above mentioned problems, additional non-trivial issues must be addressed, which are covered by the high-level requirements on prototypes developed within the PrimeLife project. As far as the demonstrator is based on the backup and synchronisation functionality, it also has to address further privacy-related issues amplified by the backup and synchronisation nature. Due to space limitations, we cannot elaborate all the requirements and possible solutions here. The interested reader is recommended to read the document "Towards a Privacy-Enhanced Backup and Synchronisation Demonstrator Respecting Lifetime Aspects" available on the Internet[4].

In order to correctly understand the descriptions in the following sections, it is helpful to be familiar with the following terminology:

Terms:

Primary item: an original item for which one or more backup items are created during the back up action. In a general sense, a primary item can be referred to as any determinate set of data, which has one or more copies called backup items dedicated for backup purposes. A primary item can be a file but it can also be a more specific type of data such as, for instance, an e-mail, a contact, or even settings on the TV.

Backup item: a copy of a primary item stored in the backup. A backup item reflects the data of a primary item at the time when the backup item is created. Note that even if each backup item must belong to one and only one primary item, this primary item may not exist during the entire lifetime of the backup item. A backup item can exist in several versions at a particular point of time.

Backup: a non-empty set of backup items.

Backup task: describes which specific set of primary items should be backed up to which storage provider according to which schedule. The output of a run of a given backup task is a backup.

[4] Document H1.3.6 (http://www.primelife.eu/images/stories/deliverables/h1.3.6_final.pdf)

Actors:

Primary user: Data Subject who owns/holds primary items.
Storage provider: provides storage space for backups.
Delegate: an entity that receives particular rights on the backup from a delegator.
Delegator: an entity that has the privilege to delegate rights to delegates concerning a particular backup. In most applications of this demonstrator, the primary user acts as the delegator.
Delegate candidate: an entity that was selected by delegator to act as a delegate but does not possess particular rights yet.
Delegation request: a request sent to the delegate candidate asking him whether he accepts particular rights from the delegator.
Credential issuer: an entity that issues a credential verifying a certain status of the primary user. This status can for example be: "primary user is ill", "primary user is hospitalised", "primary user is dead" or others.

4.2.1 Overview of the Backup Demonstrator Architecture

After describing the overall goals and visions with respect to the backup scenario in the previous section, we will report on the current state of the design and implementation of the actual backup demonstrator in the rest of this chapter.

The backup demonstrator consists of three main components:

1. **the core**, which offers the main functionality. The core is written in Java and runs as a background process on the machine that contains the data (primary items) that should be backed up. The core makes its functionality accessible using a REST-like interface.
2. **a Web-based user interface** (called "backup console"), which is written using HTML, CSS, JavaScript and Ajax. It can be displayed using an ordinary web browser. It utilises the REST calls offered by the core to accomplish the tasks desired by the user.
3. **a tray icon** shown in the notification area of the taskbar found in many modern operating systems. This tray icon informs the user about important information and status changes related to the backup process. Moreover, it allows the user to launch the backup console.

4.2.1.1 Basic building blocks used by the core

The core is the central place that provides all the methods necessary to create backup tasks and actual backups, to restore them, to manage delegations etc. Moreover, it manages all the data (configuration information, credentials etc.) related to this.

The entire functionality is provided through REST-like calls. Thereby HTTP is used as the underlying application level protocol. Therefore, the core contains an embedded web server (namely Jetty[5]). The binding between the HTTP URLs and the Java methods is done with the help of the "Java API for RESTful Web Service" (JAX-RS[6]). JAX-RS uses Java annotations, which simplifies the process of making a Java method available as web service. Particularly the Jersey[7] reference implementation of JAX-RS is used.

In case the URL itself does not encode all parameters and data necessary for a given REST call, the HTTP body contains the additional data needed as input for the given REST call. The data transmitted in the HTTP body can be encoded either using XML or JSON. The desired encoding type is set by the client. The marshalling/unmarshalling is done using JAXB[8] (in case of XML encoding) and Jackson[9] (in case of JSON encoding). Both APIs allow an automatic mapping between the internal representation and the external one, thus avoiding any need for manually implementing serialisation methods for every data type used.

4.2.1.2 File system and external storage providers

When it comes to the question of where to store the backups, the backup demonstrator follows the trend to store data "online", e.g., by using dedicated online storage providers or more generally spoken "in the cloud". Although storing backup data more "locally" is supported by the backup demonstrator (e.g., on an external hard drive connected to the machine, where data should be backed up), using online storage is the more important use case. The reason for this is clearly not the "following trends" aspect. It is rather driven by the "lifelong" aspects which should be demonstrated.

On the one hand, the backup scenario implies that the data (backups) are available for the entire life of the primary user. Clearly managing a large set of external hard drives would lead to much more burden on the user compared to managing contracts with online storage providers. Moreover, by using online storage, the data is accessible from every place in the world where an internet connection is available. This makes the whole process of delegating access rights to a backup much more feasible. Finally, it is much easier to store the backups as truly redundant copies, which in turn is a precondition for avoiding long term data losses.

On the other hand, using online storage leads to interesting research questions with respect to the "lifelong privacy" aspects. First of all, the backed up data needs to be stored in a way so that only authorised entities can gain access to the actual backup content. Besides this need to achieve confidentiality of the backed up data, the "privacy" of the involved entities (e.g., primary user, delegates etc.) should be assured as well. In our concept, it means that each online storage provider should

[5] http://eclipse.org/jetty/

[6] http://jcp.org/en/jsr/detail?id=311

[7] https://jersey.dev.java.net/

[8] http://jcp.org/en/jsr/detail?id=222

[9] http://jackson.codehaus.org/

learn as little information as possible about the involved parties (e.g., who accesses which backup, at which time, how often etc.).

The demonstrator uses the "Apache Commons Virtual File System"[10] library, which allows access to various different local and remote file systems by a single API. Besides the online storage providers, which are supported by default (like SFTP, WebDAV, FTP etc.), plug-ins for well known online storage providers (like Dropbox or Ubuntu One) were developed. Although these plug-ins were deployed together with the software of the demonstrator, conceptually they could be downloaded from the Internet as well. The general idea is that either some kind of directory service lists available online storage providers and provides the necessary plug-ins or that at least a given online storage provider offers a plug-in for his service on his web site.

A storage provider plug-in would not only implement all the necessary functionality to actually access the online storage but will also provide user interface components which allow a user to create a new account with this storage provider. This in turn comprises e.g., the necessary SLA and the payment.

With respect to the trustworthiness (or more general the properties) of an online storage provider, the main assumptions are related to availability, i.e., it is assumed that some data stored at a given storage provider would be (with high probability) available according to the negotiated SLA. Beyond that, each storage provider is seen as a potential attacker (in the sense of the concepts of multi-lateral security). Especially, it should not be necessary to trust the storage provider with respect to confidentiality or integrity of the stored data. Nor should the storage provider be trusted with respect to the privacy of a given user. Therefore the backup demonstrator needs to implement appropriated measures to achieve these protection goals (e.g., by means of cryptographic mechanisms like encryption, integrity protection etc.).

In order to reduce the linkability between different transactions done by the same user with a given online storage provider (or multiple storage providers), it is assumed that a communication layer anonymisation service is used. If the access to the online storage is based on HTTP (like WebDAV, Dropbox etc.), existing anonymisation services like AN.ON[11] or Tor[12] could be used.

Nevertheless, there usually remains the linkability introduced at the application layer. Because we want to support existing (legacy) online storage providers, we cannot assume that they base their access control on unlinkable anonymous credentials. Rather, the common combination of login/password would be used. In this case, the only way to avoid linkability, e.g., that two backups belong to the same user, a user has to create multiple accounts (ideally using multiple online storage providers). Note that for the sole purpose of demonstration, we plan to develop a special online storage provider that in fact uses anonymous credentials for access control. More concrete, the implementation of this storage provider will be based on

[10] `http://commons.apache.org/vfs/`

[11] `http://anon.inf.tu-dresden.de/`

[12] `https://www.torproject.org/`

the Identity Mixer[13] anonymous credentials, which a part of the PRIME Core[14] developed within the EU FP6 integrated project "Prime"[15] and the PrimeLife project.

4.2.1.3 Backup and Restore

"Backup" and "Restore" are the most prominent functionalities common to nearly any existing backup solution. Moreover, most backup tools operate on a *data* level, that is the user selects files, directories, partitions or whole disk drives. This kind of selection mode is supported in the PrimeLife backup demonstrator as well.

In addition to this common data level based selection mechanisms, the demonstrator offers an *identity* based selection mode. Remember that a basic concept of privacy-enhanced identity management is the separation into different partial identities, areas of life and stages of life. Thus, the user can select from each of these categories, e.g., specify which area(s) of life or partial identities he wants to backup.

Items of different types can be grouped together within a so-called container. A container can be understood as template of the primary items to be backed up. This would ease the related selection during the creation of a new backup task.

Areas of life, partial identities and files are related to each other (see Figure 4.3). An Area of Life typically covers multiple partial identities; likewise a partial identity is related to many files. Note that a file can be related to more than one partial identity and area of life, e.g., a digital photo that shows persons from the "Family Area of Life" as well as the "Working Area of Life".

In the field of privacy-enhanced identity management, one of the basic reasons for differentiating between different partial identities, areas of life etc. is to avoid unwanted linkability between them. Consequently, this unlinkability property has to be preserved for the backups as well. Thus, even if the user can group different partial identities or areas of life together for creating one logical backup task, the actual backup data need to be stored separately. Assume for instance, that the user selects files that belong to two different partial identities. In this case, two different backups will be created, ideally stored on two different online storage providers. Otherwise an attacker might learn that the two partial identities in question actually belong to one and the same user.

You might wonder how the backup demonstrator knows about the existing areas of life, partial identities and the relation among them and the files stored on the given machine. Conceptually, this information is provided as a core functionality of a privacy-enhanced identity management system, which in turn is out of scope of the backup demonstrator itself. Thus, for the backup demonstrator, we simply assume that such an IDM system is in place. However, it does not really exist today in practice. Therefore, we decided to create mockup data (areas of life, partial identities, files and folders) that are used for demonstrating the privacy-preserving

[13] `http://www.primelife.eu/results/opensource/55-identity-mixer/`

[14] `http://www.primelife.eu/results/opensource/73-prime-core/`

[15] `https://www.prime-project.eu/`

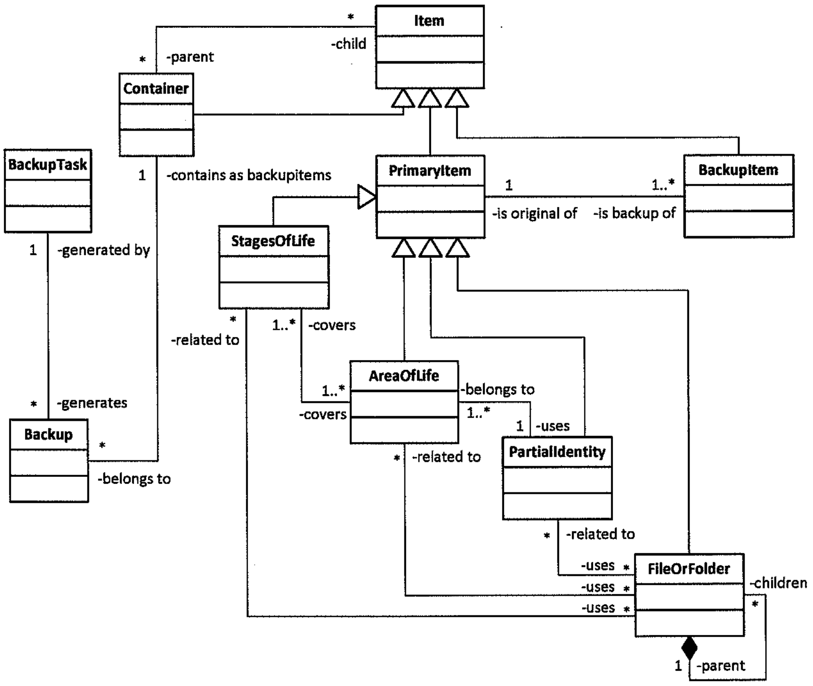

Fig. 4.3: Relations among of Areas of Life, partial identities, files etc.

backup aspects related to them. Nevertheless, the backup demonstrator allows for creating backups of real files and folders, but these items are not associated with any partial identity or area of life. Thus, from a privacy (linkability) point of view, they are treated as being "equal" and therefore are stored within a single backup.

4.2.1.4 Delegation

Delegation is considered to be one of the most important aspects to be shown by the demonstrator. It cannot only be seen as an innovative feature in itself, which is not common in today's backup tools, but it also has many implications with respect to the area of lifelong privacy, which in turn is the underlying motivation for the demonstrator.

From a functional point of view, delegation means that a primary user (the delegator) delegates access rights to a particular backup to some delegates (see Figure 4.4). These access rights are usually bound to a policy describing under which circumstances the access should be granted. A typical use case would be that a user delegates the access rights to the files related to his work partial identity/area of life

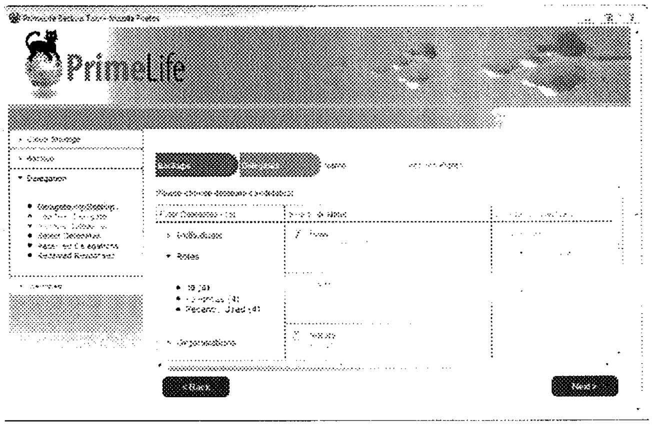

Fig. 4.4: Management console of the backup demonstrator showing the interface for selection of delegates

to his colleagues under the policy that the access is only permitted if the user becomes ill. Moreover, the policy would express the obligation that the user needs to be informed if one of his colleagues really accesses the backup.

Delegation does not only deal with the privacy of the delegator but also with the privacy of the delegates. Therefore, a delegate will be asked if he is willing to accept the delegation request. Thereby in the spirit of informed consents he will be informed that each access to the backup would also be reported to the delegator.

It is currently an open question as to how much additional "meta-data" about a given backup will be communicated to a delegate. On the one hand, a delegate might want to know beforehand which files are included in a given backup, so that he can better decide if he really wants to accept the delegation or access the backup, respectively. On the other hand, this information might already have some negative impact on the privacy of the delegator. For enhancing the privacy of the delegator, it is desirable that a delegate only learns that information if he actually accesses a given backup. Thus, if the conditions defined in the access policy never become true, any unnecessary information flow will be avoided.

In the current version of the demonstrator, the list of possible delegates is predefined as mockup data. There are plans to integrate other sources of information for these address book-like data. Natural sources are social networks such as Facebook, Xing, LinkedIn etc. Moreover, the current demonstrator uses plain e-mail messages for transmitting delegations and the related acceptance responses. Future versions of the demonstrator might use the communication infrastructure offered by the mentioned social networks as well.

The delegation itself is an XML document describing the contents of the backup, the location of the backup and under which circumstances (policy) the backup can be accessed by the delegate. The delegation might also transfer some credentials (e.g., issued by the delegator) to the delegate, which the delegate will need in order to access the backup.

Conceptually, a lot of modern cryptographic mechanisms exist that could be used to construct a privacy-enhanced protocol for the purpose of delegation. Such a solution would require infrastructural support that is not in place today. Examples would be anonymous credentials issued by governmental or public institutions, public key infrastructures, attribute based access control mechanisms etc. Therefore we decided to implement a much simpler scenario, which on the one hand is much closer to the current practice and on the other hand integrates components developed within the PrimeLife project and thus would not only demonstrate delegation itself, but also illustrate the interplay of the various components developed within the PrimeLife project.

Our delegation scenario comprises the following entities/components:

1. the eCV, a component which allows a user to store an electronic version of his CV. This component was developed within the PrimeLife project to demonstrate privacy aspects in service oriented architectures (see Section 21)
2. a trusted third party (TTP) utilising the PrimeLife policy engine (see Section 20)
3. a legacy online storage provider
4. the delegator
5. a delegate

In this scenario, we demonstrate how a delegator can delegate the access to his backup to a delegate who can access the backup in case the delegator is ill. The delegation then would comprise the following steps:

1. The delegator stores the encrypted backup at the legacy online storage provider. The encryption is done using a symmetric cipher and the key k.
2. The delegator generates a random value k_1 and calculates k_2 such what $k = k_1 \oplus k_2$. Note that $k_2 = k \oplus k_1$, i.e., k_2 can be seen as a one time pad encryption of k using the key k_1.
3. The delegator stores k_1 at the TTP. A PPL policy regulates the access to k_1 saying that:
   - only give access within a certain time frame and
   - to someone who can present a credential C_1 and
   - a credential proving that the delegator is currently ill
   - the obligation is to delete k_1 after the time frame is over and
   - to inform the delegator if someone has accessed k_1
4. The delegator sends k_2, C_1, C_2 to the delegate (encrypted under the public key of the delegate). The delegator informs the delegate about the circumstances under which (illness of delegator and valid time frame) and how the delegate can access the backup.

Now let's assume that the delegator does in fact become ill. In this case, the delegator (or his doctor) sends a certification of illness to the eCV of the delegator.

Moreover, the access policy to that certification says that someone who shows credential C_2 is allowed to access the certificate of illness (in this case the delegator should be informed by the eCV).

In case the delegate wants to access the backup of the delegator and thus uses the rights delegated to him, the following steps happen:

1. The delegate downloads the encrypted backup. How the access control to the encrypted backup is done depends on the methods provided by the legacy online storage provider. Usually there exist some means of sharing the stored data with a defined set of users. But a broader discussion of this issue is out of scope here.
2. The delegate shows credential C_2 to the eCV and requests the certificate of illness of the delegator. Note that the delegate has received C_2 in step 4 during the delegation process.
3. The delegate requests k_1 from the TTP. He therefore shows credential C_1 together with the certificate of illness of the delegator to the TTP. The delegate gets k_1 and the TTP informs the delegator about that access to k_1.
4. Now the delegate is able to calculate $k = k_1 \oplus k_2$ and can decrypt the encrypted backup of the delegator.

Of course, the description above shows a very brief overview explaining the general ideas we have with respect to using existing components developed by the PrimeLife project for the backup demonstrator. Further research needs to be done to avoid/remove all the linkage that remains between the different steps (e.g., using different transaction pseudonyms for delegator and delegate etc.). Moreover, the information the various parties learn should be minimised (e.g., the TTP does not need to know that the "certificate of illness" is actually a certification of illness (e.g., we could use "meaningless" credentials here)). As a final remark, please note that all the parties mentioned above can be distributed (in the usual k out of n setting). To some extent, the involvement of the TTP already is a kind of distribution, because the eCV and the TTP can be seen as entities that store information accessible under a given access policy.

4.2.2 Deployment and Usage of the Demonstrator

Due to the platform independent design based on Java and web-technologies, the demonstrator can be installed and run on many modern operating systems including Linux, Mac OS X and Windows. Nevertheless, we decided to create a dedicated virtual machine based on VirtualBox[16] as virtual machine monitor and Ubuntu[17] Linux as guest operating system, which contains all the necessary components preinstalled. This makes the process of "playing" with the demonstrator much easier, especially if it comes to the more complex delegation scenarios. Besides the demonstrator itself, the virtual machine contains a local WebDAV online storage provider,

[16] `http://www.virtualbox.org/`

[17] `http://www.ubuntu.com/`

two separate user accounts (one for the primary user/delegator and one for the delegate), a local e-mail infrastructure (SMTP and IMAP servers) etc. In order to make the installation of the virtual machine itself as easy as possible, a screencast explaining the necessary steps was created.

4.3 Concluding Remarks

In this chapter, the concept of the privacy-enhanced backup and synchronisation demonstrator was presented. It was shown that the objectives of lifelong privacy lead to practical results, which can be applied for solving real-life issues in enhanced ways. Our demonstrator reveals new problems that emerge as soon as lifelong aspects related to the Data Subject are taken into consideration. We presented a new approach, which can help the average citizen to protect himself against unwanted data loss respecting his different partial identities and areas of life. Our approach proceeds in such a way that it takes into account lifelong aspects of a human being and corresponding implications within the scope of the privacy.

Nevertheless, although for many of the envisaged problems that need to be solved with respect to lifelong aspects and here especially lifelong privacy solutions are known in principle. However, sometimes the concrete implementations that need to be realised in the demonstrator are currently still under development. This on the one hand explains why certain aspects and mechanisms were only described at a very high (abstract) level and defines on the other hand the research and development roadmap for the remaining five months planned for the finalisation of the demonstrator.

4.4 Acknowledgements

Many thanks are due to Hagen Wahrig and Katrin Borcea-Pfitzmann for their contribution to the design and implementation of the demonstrator and to Sebastian Clauß for his contribution to the requirements analysis. Our thanks also go to everyone in the PrimeLife project who took part in the scenario, requirements and demonstrator discussions, especially Rainer Böhme for his collection of demonstrator ideas from these discussions.

References

[ACK$^+$09] C.A. Ardagna, J. Camenisch, M. Kohlweiss, R. Leenes, G. Neven, B. Priem, P. Samarati, D. Sommer, and M. Verdicchio, *Exploiting Cryptography for Privacy-Enhanced Access Control: A result of the PRIME Project*, 2009, to appear.

[Ada99] Anne Adams, *The implications of users' privacy perception on communication and information privacy policies*, In Proceedings of Telecommunications Policy Research Conference (Washington, DC), 1999.

[AG06] Allesandro Acquisti and Ralph Gross, *Imagined communities: Awareness, information sharing, and privacy on the facebook*, 6th Workshop on Privacy Enhancing Technologies, 2006.

[BK09] Jo Bryce and Mathias Klang, *Young people, disclosure of personal information and online privacy: Control, choice and consequences*, Inf. Secur. Tech. Rep. **14** (2009), no. 3, 160–166.

[BM09] Suzy Bausch and Michelle McGiboney, *News release: Social networks & blogs now 4th most popular online activity*, `http://en-us.nielsen.com/content/dam/nielsen/en_us/documents/pdf/Press%20Releases/2009/March/Nielsen_Social_Networking_Final.pdf`, March 2009.

[BMH05] Matthias Bauer, Martin Meints, and Marit Hansen, *Structured Overview on Prototypes and Concepts of Identity Management Systems: FIDIS Del. 3.1*, Available from `http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp3-del3.1.overview_on_IMS.final.pdf` (letzter Abruf 09.02.2009), 2005.

[Bur09] Jörg Burger, *Lügnerin! Betrügerin!*, `http://www.zeit.de/2009/53/Internetmobbing?page=all`, December 2009.

[Cha85] David Chaum, *Security without identification: transaction systems to make big brother obsolete*, Communications of the ACM **28** (1985), no. 10, 1030–1044.

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

[CK01] Sebastian Clauß and Marit Köhntopp, *Identity management and its support of multilateral security*, Computer Networks **37** (2001), no. 2, 205–219.

[Con09] G. Conti, *Googling security; how much does google know about you?*, New York, Addison Wesley publishers, p. 91., 2009.

[CvH02] Jan Camenisch and Els van Herreweghen, *Design and implementation of the idemix anonymous credential system*, Proceedings of the 9th ACM conference on Computer and communications security, 2002, pp. 21 - 30.

[Dör08] Nicola Döring, *Reduced social cues / cues filtered out*, Medienpsychologie. Schlüsselbegriffe und Konzepte (Stuttgart) (N. C. Krämer, S. Schwan, D. Unz, and M. Suckfüll, eds.), Kohlhammer, 2008, pp. 290–297.

[DPW09] *Opinion 2/2009 on the protection of children's personal data (general guidelines and the special case of schools)*, Art. 29 Data Protection Working Party, 2009, `http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2009/wp160\_en.pdf`.

[EB09] Lilian Edwards and Ian Brown, *Harboring data: Information security, law and the corporation*, ch. Data Control and Social Networking: Irreconcilable Ideas?, Stanford University Press, 2009.

[EGH08] Anja Ebersbach, Markus Glaser, and Richard Heigl, *Social Web*, UTB, vol. 3065, UVK, Konstanz, 2008.

[EK02] G. Eysenbach and C. Khler, *How do consumers search for and appraise health information on the world wide web? qualitative study using focus groups, usability tests, and in-depth interviews*, 2002, pp. 573–577.

[Eni08] Enisa, *Technology-induced Challenges in Privacy and Data Protection in Europe*, A report by the ENISA Ad Hoc Working Group on Privacy and Technology, European Network and Information Security Agency (ENISA), Heraklion, Crete, Greece, October 2008.

[Fac] *Facebook statistics*, `http://www.facebook.com/press/info.php?statistics`, last access 21 October 2010.

[FC01] J. W. Fritch and R. L. Cromwell, *Evaluating internet resources: identity, affiliation, and cognitive authority in a networked world*, Journal of the American Society for Information Science and Technology **52** (2001), no. 6, 499–507.

[FM00] A. J. Flanagin and M. J. Metzger, *Perceptions of internet information credibility*, Journalism & Mass Communication Quarterly **77** (2000), no. 3, 515–540.

[FSD^{+}03] B. J. Fogg, C. Soohoo, D. R. Danielson, L. Marable, J. Stanford, and E.R. Trauber, *How do users evaluate the credibility of web sites? a study with over 2.500 participants. proceedings of dux2003, designing for user experiences conference*, `http://www.consumerwebwatch.org/dynamic/web-credibility-reports-evaluate-abstract.cfm`, 2003, pp. 573–577.

[GA05] Ralph Gross and Alessandro Acquisti, *Information revelation and privacy in online social networks*, WPES '05: Proceedings of the 2005 ACM workshop on Privacy in the electronic society (New York, NY, USA), ACM, 2005, pp. 71–80.

[Gof59] Erving Goffman, *The presentation of self in everyday life*, Doubleday, 1959.

[Gri08] James Grimmelmann, *Facebook and the social dynamics of privacy [draft version]*, `http://works.bepress.com/james_grimmelmann/20/`, 2008.

[HBPP05] Marit Hansen, Katrin Borcea-Pfitzmann, and Andreas Pfitzmann, *PRIME – Ein europäisches Projekt für nutzerbestimmtes Identitätsmanagement*, it - Information Technology, Oldenbourg **6** (2005), no. 47, 352–359.

[Hou09] Michelle G. Hough, *Keeping it to ourselves: Technology, privacy, and the loss of reserve*, Technology in Society **31** (2009), no. 4, 406–413.

[How08] Jeff Howe, *Crowdsourcing: Why the power of the crowd is driving the future of business*, Crown Business, 2008.

[HPS08] Marit Hansen, Andreas Pfitzmann, and Sandra Steinbrecher, *Identity management throughout one's whole life*, Information Security Technical Report **13** (2008), no. 2, 83–94.

[HRSZ10] Marit Hansen, Maren Raguse, Katalin Storf, and Harald Zwingelberg, *Delegation for privacy management from womb to tomb – a european perspective*, Privacy and Identity Management for Life (Michele Bezzi, Penny Duquenoy, Simone Fischer-Hübner, Marit Hansen, and Ge Zhang, eds.), IFIP Advances in Information and Communication Technology, vol. 320, Springer Boston, 2010, 10.1007/978-3-642-14282-6_2, pp. 18–33.

[Ind08] Irish Independent, *Taxman admits to Facebook 'trawl'*, `http://www.independent.ie/national-news/taxman-admits-to-facebook-trawl-1297118.html`, February 2008.

[Joi08] Adam N. Joinson, *'looking at', 'looking up' or 'keeping up' with people? motives and uses of facebook*, CHI 2008, ACM, 2008.

[LB08] Matthew M. Lucas and Nikita Borisov, *Flybynight: Mitigating the privacy risks of social networking.*, Proceedings of the 7th ACM workshop on Privacy in the Electronic Society (WPES), ACM, 2008.

[Lea08] Charles Leadbeater, *We-think: Mass innovation, not mass production*, Profile, 2008.

[Man10] Farhad Manjoo, *Social networking your way to a new job*, `http://www.nytimes.com/2010/08/26/education/26SOCIAL.html?pagewanted=1&_r=1`, August 2010.

[Met07] M.J. Metzger, *Making sense of credibility on the web: models for evaluating online information and recommendations for future research*, Journal of the American Society for Information Science and Technology **58** (2007), no. 13, 2078–2091.

[MS09] Viktor Mayer-Schönberger, *Delete: The virtue of forgetting in the digital age*, Princeton University Press, 2009.

[MSF09] Ines Mergel, Charlie Schweik, and Jane Fountain, *The transformational effect of web 2.0 technologies on government*, `http://ssrn.com/abstract=1412796`, 009.

[New09] CBC News, *Depressed woman loses benefits over Facebook photos*, `http://www.cbc.ca/canada/montreal/story/2009/11/19/quebec-facebook-sick-leave-benefits.html`, November 2009.

[O'R07] Tim O'Reilly, *What is web 2.0: Design patterns and business models for the next generation of software*, Communications & Strategies **65** (2007), no. 1, 17–37.

[PBP10] Stefanie Pötzsch and Katrin Borcea-Pfitzmann, *Privacy-respecting access control in collaborative workspaces*, Privacy and Identity, IFIP AICT 320 (Nice, France) (M. Bezzi et al., ed.), Springer, 2010, pp. 102–111.

[PD03] Leysia Palen and Paul Dourish, *Unpacking 'privacy' for a networked world*, Computer-Human Interaction (CHI) Conference 2003, 2003, pp. 129–137.

[php] phpBB, *Official website*, `http://www.phpbb.com`.

[Pöt09] Stefanie Pötzsch, *Privacy awareness – a means to solve the privacy paradox?*, Proceedings of the IFIP/FIDIS Internet Security and Privacy Summer School, 2009.

[PRI] PRIME, *Privacy and Identity Management for Europe*, https://www.prime-project.eu.

[PWG10] Stefanie Pötzsch, Peter Wolkerstorfer, and Cornelia Graf, *Privacy-Awareness Information for Web Forums: Results from an Empirical Study*, 6th Nordic Conference on Human-Computer Interaction (Reykjavik), October 2010.

[rfc07] *RFC 4880 - OpenPGP Message Format*, Tech. report, Internet Engineering Task Force, November 2007.

[RG10] Kate Raynes-Goldie, *Aliases, creeping, and wall cleaning: Understanding privacy in the age of facebook*, 2010.

[Rie02] S.Y. Rieh, *Judgement of information quality and cognitive authority in the web*, Journal of the American Society for Information Science and Technology **53** (2002), no. 2, 145–161.

[SGL06] Martin Szugat, Jan Erik Gewehr, and Cordula Lochmann, *Social Software schnell & kompakt*, entwickler.press, 2006.

[Sol07] Daniel J. Solove, *The future of reputation: Gossip, rumor, and privacy on the internet*, Yale University Press, 2007.

[Tad10] Monika Taddicken, *Measuring Online Privacy Concern and Protection in the (Social) Web: Development of the APCP and APCP-18 Scale*, 60th Annual ICA Conference (International Communication Association) (Singapur.), June 2010.

[Tap09] Don Tapscott, *Grown up digital: How the net generation is changing your world*, McGraw-Hill, 2009.

[Tuf08] Zeynep Tufekci, *Can you see me now? audience and disclosure regulation in online social network sites*, Bulletin of Science, Technology and Society **28** (2008), no. 1, 20–36.

[Wik] *List of social networking websites*, `http://en.wikipedia.org/wiki/List_of_social_networking_websites`.

[WP110a] PrimeLife WP1.2, *Privacy enabled communities*, PrimeLife Deliverable D1.2.1 (Ronald Leenes Bibi van den Berg, ed.), PrimeLife, `http://www.primelife.eu/results/documents`, April 2010.

[WP110b] ———, *Privacy-enabled communities demonstrator*, PrimeLife Deliverable D1.2.2 (Stefanie Pötzsch, ed.), PrimeLife, `http://www.primelife.eu/results/documents`, February 2010.

[Yok07] A.C. Yokum, *I lost my job because of Facebook*, `http://www.associatedcontent.com/article/353378/i_lost_my_job_because_of_facebook.html`, August 2007.

[YQH09] Alyson L. Young and Anabel Quan-Haase, *Information revelation and internet privacy concerns on social network sites: A case study of facebook*, C&T '09, ACM, 2009, pp. 265–274.