3

The Identity Landscape

Bart Priem[1], Ronald Leenes[1], Eleni Kosta[2], and Aleksandra Kuczerawy[2]

[1] Tilburg University
[2] KU Leuven

3.1 Introduction

Many people will have an image of 'who they are' and how their identity is established. Moreover, most individuals will probably relate the concept of identity (and identity management) to their reputation as an individual, how they define themselves, and how others look at them. In this view, identity relates to the personal aspect of identity. However, the term identity is also used in many other ways, for instance in the sense of cultural identity — what makes an Englishman English? —, or in the sense of identity management in IT systems. Because of this, a clear definition of 'identity' is difficult to provide.

One of the developments that influences the notion and the use of the term identity is the development and use of Information and Communication Technologies (ICTs). Especially the creation of the 'online environment' has added complexity to the notion of identity. The online environment, for instance, lacks a clear ID infrastructure. It was designed to identify the endpoints of communication, which typically are devices (such as computers) that are, or were, shared by multiple individuals. Important aspects of identification in the offline world, such as the presence of the body as a means to recognize and identify individuals is lacking online. Instead, internet-facilitated interaction currently relies heavily on information that can be manipulated and that has unclear status to identify and represent human beings and devices. Because of this, several initiatives exist to improve online identity management (IdM). All these initiatives operate in a rapidly evolving field with moving targets and changing issues. Furthermore they need to deal with the diverse interests of the various stakeholders.

J. Camenisch, R. Leenes, and D. Sommer (Eds.): Digital Privacy, LNCS 6545, pp. 33–51, 2011.

To put the PRIME project, its technology, and its vision in perspective, this chapter will provide an introduction to the current landscape of online identity and online identity management. We will discuss some of the meanings of the term 'identity' and describe developments in the identity management field which can be summarized as an evolution from enterprise centric towards user centric solutions. We will conclude the chapter with some complicating developments that illustrates issues to come that need to be incorporated in any comprehensive identity management system. This chapter serves as a foundation for the chapters to come. It does not, however, provide an extensive overview of the philosophical and sociological aspects of identity.

3.2 The Concept of (Online) Identity

Identity is a dynamic and contextual concept. It has several dimensions. It is, for instance, used to represent a person, but is also used to identify and recognise such a person. Thus, identity is used both in descriptive terms and process terms [WP205]. One can furthermore refer to identity as to who a person 'really is' (sometimes called 'ipse identity'), but also as to how a person is characterised or represented by himself or by others (or 'idem identity'). There is thus a difference in the notion of identity from a philosophical point of view (who someone really is) which regards identity as fluid and indeterminate, and the more 'practical' view on identity which relates to the static representation of an individual in a certain context in the form of a set of attributes related to this individual (see [WP205]).

When identity is considered in the context of online identity management, we mainly deal with the static identity of an individual (represented in data) and its composition and deployment throughout online contexts. In the online environment, identity management primarily relates to the composition of an identity out of 'identity information' that relates to an individual or another entity that acts in this environment. In this sense, both human beings and devices can have an online identity; historically, device identity preceded human identity in the online environment because the internet was developed as a computer-to-computer infrastructure [Coy07, Cam05].

Both online and offline, individuals interact with people and organisations in many different relations. All these relations concern the exchange of information and/or attribute-value pairs. Different (kinds of) relationships involve different parcels of information and therefore individuals present different images of themselves in different contexts. A single individual therefore consists of different characterisations tied to the different contexts in which she operates. For example, the co-workers in a work-related context will characterise an individual differently than the friends that interact with the same individual in the context of friendship. The relevant attributes associated to an individual are different in a working environment than in a social environment and individuals may also represent themselves differently throughout such contexts. As we will see later in more detail, this capability to keep the different contexts separated, ‘audience segregation’ [Gof59], is an essential characteristic of modern (western) societies which allows for different kinds of social relationships to be established and maintained [Rac75].

In the online environment, the different manifestations of an individual can be defined as partial identities, or ‘digital personae’ [Cla94], which are constructed from the information people give, or ‘give off’ in a relation [Gof59]. The construction of a partial identity is not solely based on information that is determined and controlled by the individual to whom an identity relates (‘projected’ in Clarke’s terms). Others, the recipients, may construct their own image of the individual by observing them or their behaviour as represented in data and they may add information to an existing partial identity (which leads to ‘imposed personae’ in Clarke’s terms). The information contained in a partial (imposed) persona may not always be known to the individual concerned.

Partial identities in the online world are thus determined both by information known and unknown to the represented individual and this information may be controllable or uncontrollable by the individual. Moreover, the perception of a partial identity can be different between the individual to whom an identity relates and the person or organisation that uses such an identity [WP205].

Identity already used to be a complex concept for the offline environment, but in the online world it is even a more ‘muddled thing’ [Cha06], because the internet provides the possibility of disembodied use of identities (ie. without the individual’s bodily presence) and facilitates the decontextualisation and transfer of identities (and identity data). On the internet, traditional ‘trust tokens’ (e.g., clothing, buildings, driving licenses) are largely absent.

3.3 Asymmetric Perspectives

The field of identity management has many stakeholders with different, and potentially conflicting, interests in the design and use of identity management systems. Consumers, regulators, and enterprises can have different perspectives on the concepts of identity, identity management, the online environment, and the use of identity information. ‘One-size-fits all’ solutions may therefore be difficult to develop and designers need to balance difference perspectives, interests, and requirements. In order to understand these different interests and conceptions of identity and identity management, we will first discuss identity management from the perspective of two principal stakeholders, enterprises (and government) and the individual.

3.3.1 The Enterprise-Centric View on Identity Management

Enterprises and governments have driven the development of identity management systems as a means to know with whom they communicate [OMS^{+}07].

Access control to resources and hence, identification, authentication and authorisation are therefore the key concepts in contemporary identity management. Private *enterprises* that are active in the online environment, make use of identities (e.g., user accounts) to meet strategic objectives, such as ensuring the accuracy of identity information, utilizing the possibilities to store and manage large amounts of data, and the use of information to develop and distribute products and services effectively and efficiently (in a better way than competitors do), and reducing the risk of data loss. The *government* is another major stakeholder. The government needs identity management to provide efficient personalized electronic public services and to prevent citizens from falling victim to fraud and insecurity whilst providing these services. Moreover, the government is a stakeholder in IdM development in general, because IdM promotes the free flow of information in society which can increase welfare, for example.

Identity management developments have been driven by an enterprise-centric view on IdM. Many of the developments that will be described later on in this chapter depart from a perspective that the core function of an IdM system is to manage who has access to certain resources. Online IdM in this view comprises the use of partial identities for identification, authorisation and authentication of individuals to provide them certain services. Central to this kind of identity management are user accounts. These accounts also contain (or link to) data that provides insight in (customer) preferences, purchasing history, and contact data, for example. This information allows the enterprise to create personalised, and customer-oriented services. Most organisations active on the internet keep track of users' purchases, and there is an active market for such customer data [EI06, Tay02].

Enterprise-centric IdM systems focus on facilitating service delivery to the right person, which is 'their' customer or client. The fact that these customers also have accounts at other enterprises which causes inconveniences for these individuals is not a primary concern of the respective enterprises.

3.3.2 A User-Centric View on Identity Management

Individuals are right in the middle of online identity management, because it concerns the management of their identities, and because decisions are made on the basis of these identities. From an individual's point of view, the concept of identity management therefore not only relates to the access control regarding resources. It also, or maybe even rather, relates to how they are manifested and represented, and how this is aligned to their own perception of their identity. Identity management in this sense strongly relates to role playing and presentation of self. The individual should be able to act as an autonomous individual, be able to control their reputation, and have insight in the way they are judged by others in a specific context.

The online environment facilitates the construction and maintenance of projected and imposed personae. Data can easily be collected and combined into rich personae, transcending the context in which individual bits of information were disclosed. The decontextualisation and combination of data from different sources makes it difficult for individuals to control their different digital personae. This undermines their capabilities to control the image they present in different contexts and to segregate audiences online. The need to do so exists online just as it does offline. People engage in different kinds of activities online (e.g., public, commercial, and intimate) and need to be able to construct matching identities that meet the behavioural rules and requirements set by these different environments.

Important values such as reputation, dignity, autonomy, judgement, and choice are closely related to the individual perspective on identity management. When people cannot determine or control their identity, they may become overexposed, confused, or discriminated, for example. Human beings have an interest in naming and sorting themselves [Gan93, Raa05] and to play different roles. Sometimes they may even need to be anonymous and unidentified (e.g., for purposes of emotional release, relaxation, unpunished criticism, and making mistakes). Individuals appreciate to have a diverse and autonomous life, and need to be able to adapt their identities to the environment they engage in. Even though identity management is not usually the primary goal of the individual, which may explain why many people are not eager to invest time and money in IdM systems [DD08], the social values outlined previously warrant the individual perspective to be taken into account in the development of IdM systems.

3.3.3 Combining the Perspectives

Integrating the different interests in online IdM increases its complexity. There is a clear gap between the enterprise-centric emphasis on customer-relations and access, and the user-side approach which, for instance, requires users to be able to choose different partial identities for different purposes — even within the same system — or be able to use the same partial identity in different contexts [Pfi03]. This gap needs to be closed.

It is also difficult to implement the 'personal' perspective on identity in IdM systems because of the business and government requirements of facilitating trustworthy interaction between them and their users/citizens. We need to acknowledge that the processing of some identity related information is part of the online environment and may be considered necessary in several circumstances. To completely renounce the need for the collection and processing of identity information (personal data) would severely hamper the adoption of such a system by enterprises.

A further complication of integrating both views lies in the fact that multiple parties need to subscribe to the model. Individuals can only use the same or different identities in different occasions and for different purposes if the identity system allows for this, and this requires standards and interoperability.

The fact that the enterprise-centric view to identity management is too limited seems to be acknowledged throughout the industry, as online IdM systems are evolving towards federated systems and recent developments even point towards the development of ‘user-centric IdM systems’ (coined ‘Identity 2.0’ by some), which will be demonstrated in the following section. The PRIME-project aims to be at the forefront of these developments.

3.4 Evolving Identity Management Systems

Different models for online identity management have been developed in recent history. Traditionally, identities were managed in so-called corporate identity ‘silos’. In this model one single identity management environment is operated by a single service for a specific group of users [Pat03]. Hence, every (online) service had its own identity management system built to their own requirements for authorisation and identification of individuals. From the perspective of users of multiple systems this means that they have to maintain an identity (account) for each and every service they use, which in practice means several sets of passwords and usernames. The ‘silo-model’ is still a dominant model for identity management on the internet. An obvious drawback of this scheme from the perspective of the users is that it requires them to provide the same (personal) information for every new online service.

The construction of identities in these systems is guided by rules (implicitly) set by the provider of the service. Each account is identified by an identifier. Sometimes these identifiers can be freely chosen, sometimes they have to satisfy certain rules (e.g., at least one number, 8 characters long), or be a valid email address. Individuals are therefore sometimes forced to create different identities (or rather the identifiers that identify the identity) even when they want to use the same identity across domains. Or, in the case of being obliged to use a valid email address, they may have to use identities they don’t want to use for a particular use. As a result of these practices two effects on identity construction are visible: one, difficult to remember identifiers as a result of the rules on identifiers imposed by the service provider, and two a convergence of identities to a limited set of partial identities as a result of the requirement to use email addresses as ‘usernames’. Furthermore, the ‘silo’-approach has resulted in many identity ‘one-offs’ and an ad-hoc nature of internet identity even though the identities in these silos can be managed by, for instance, storing passwords and usernames in software (password-managers) on a local computer or on a server [OMS+07, Cam05].

A next step in the development of IdM systems has been the development of single organisation single sign-on (SOSSO)[OMS+07]. Here individuals gain access to different resources (applications, web sites) within a single entity’s domain once they are authenticated. This kind of IdM slightly alleviates the individual’s burden of having to cope with potentially different identities within such a domain. Usually it also limits the individual’s capabilities to use different identities within a certain domain (e.g., the association of an account to an email address limits the number of accounts an individual can establish without also obtaining new email addresses). Effects of SOSSO are the collapse of different (social) contexts within a given domain controlled by the enterprise and linkability because the IdM provider can recognize the individual access to the various resources. SOSSO makes coping with enterprise centric IdM easier for the individual within a particular domain (e.g., company), but does not help when multiple domains are involved.

Multi-organisation single sign-on (e.g., Microsoft .Net Passport) aims to solve this problem, as well as lessen the burden of implementing and maintaining IdM systems within each enterprise in a federation [OMS+07]. In this model, authentication is outsourced to a trusted identity provider (IdP). The IdP identifies and authenticates the user and provides a credential that can be used to access resources from associated service providers. Drawbacks of this model are that the IdP stores the user's data which creates security vulnerabilities. Furthermore, the attendance of one single IdP in all interactions on the Internet creates linkability because the IdP can trace the user after authentication. It also creates a vulnerability (and convenience) because relying enterprises depend on a single IdP involved in all transactions.

Enterprise centric federated identity management (e.g., Liberty Alliance) addresses the problems related to the dependence on a single IdP in a federation, by allowing any number of IdPs to handle authentication. The user authenticates with any of the IdPs in the federation and subsequently can access resources at each of the entities in the federation (where the user has proper authorisations). Some federation schemes not only handle authentication, but also allow the transfer of attributes between the federates [OMS+07]. Federated identity schemes again limit the burden for individuals of having to cope with multiple identities when they want to use a single identity, but do not address the needs of individuals when they want to use different identities for different activities in the federation. The advantages mainly benefit the enterprises which can achieve costs savings arising from a shared scheme based on a standardised, interoperable architecture, and the outsourcing of authentication and IdM to professional identity providers.

Various initiatives in the landscape of 'federated' identity management can be pointed out. Many of these are 'token' based, whilst some are 'anonymous-credential-based systems' (see PRIME's Framework [PRI08]). The traditional token-based systems rely on identity providers that mediate the transactions. The identity provider distributes tokens to the service providers with which an individual interacts. In a token-based system, the service providers still are relying parties (Rp) with regard to the identity attributes they receive. They depend on the IdP, even though some of their vulnerability can be circumvented by means of contracts.

In recent years, a shift from an enterprise centric view to a user-centric view can be observed. Notions, such as 'Identity 2.0' (Sxip, Microsoft Cardspace, Higgins, PRIME, etc) belong in this sphere. In these initiatives the IdP is no longer in the centre of issuing and creating identities, but rather the user is.

In user-centric identity management, the individual's interests are acknowledged in the sense that they manage their own personal data and obtain credentials from identity providers which they can use in their interaction with service providers. Systems based on anonymous credentials even give the user and relying party the opportunity to use identity attributes without the use of a central identity provider [PRI08]. Such systems make it possible to really put the user in the middle of IdM, and thus indicate a shift from an enterprise-centric perspective to a user-centric perspective. The user-centric model provides the user more control over the way they present themselves to others. If designed properly, they assure the necessary level of *privacy* in the online environment.

Federated IdM systems increase convenience for the user to make use of several different services and make identities portable. Furthermore, they can create opportunities for organisations to ease the process of registration, authentication, and authorisation. In addition, these systems allow for cost saving on the retention and collection of data and can create new business opportunities (see [OMS$^+$07]).

3.5 Existing Identity Management Applications

Multiple competing identity management initiatives have emerged in recent years to deal with the Internet's lacking identity layer. These initiatives range from the aforementioned 'identity silos' and 'enterprise centric SSO systems' to 'federated IdM systems'. We will briefly describe some prominent IdM systems.

3.5.1 Microsoft Passport

One of the early initiatives for a cross service identity management is 'Microsoft Passport' (1999). It featured hundreds of millions of accounts due to the fact that Microsoft used Passport for its MSN and Hotmail services. Passport provides the user the benefit of an SSO-experience, and aims to reduce the time a user needs to register and authenticate for different services on the internet associated to Microsoft by means of contractual agreements [OMS$^+$07, PM03].

Microsoft Passport is a web-based service redirecting the user's browser for the purpose of authentication to a central authenticating server. It makes use of Cookies for maintaining (session) credentials [PM03].

In Passport, personal information is stored in a central location (under Microsoft's control) and therefore websites that participate in the initiative rely on Passport for the authentication of users instead of arranging their own authentication schemes [Opp04]. Individuals register at Passport through

Passport's home page, the Microsoft Windows operating system, or via a Hotmail e-mailaccount.

Passport's centralised model makes it vulnerable to attacks and failure. Also, because the system hardly imposes restrictions on user-selected passwords, many users pick easy to guess passwords which increases vulnerability [Opp04]. Furthermore, Passport is based on a single identity provider (Microsoft) which means that it is involved in customer relations of many other organisations. With 'Microsoft in-the-middle', (potential) users and privacy advocates have voiced concerns that this powerful IdP may acquire significant amounts of data about internet activities of the systems users and organisations that make use of the Passport system [Cam05].

Even though Passport provides a simple solution for identity management, it does not fully comply with user requirements and organisational IdM requirements. Especially the dependence on a single identity provider, Microsoft, seems to have obstructed the adoption of Passport in non-Microsoft services. Microsoft's stake in the centralised Passport system has been considered 'out of context' [Cav06]. Another aspect of a centralised IdM system like Passport that could have negatively affected adoption is that it raises concerns in the fields of control over private information, security, and competition [Cho06].

3.5.2 Liberty Alliance

A more decentralised identity management system is being developed by the Liberty Alliance project. This project was initiated in 2001 and has over 150 members, active in education, government, and including technology vendors, as well as many others. The Liberty Alliance aims to develop a federated identity management system with multiple identity providers. Because of this, identity data does not have to be stored at a central organisation whilst users can still have a web based, SSO-experience.

The goal of Liberty Alliance is to establish an open standard for federated identity management. Its technology makes it possible to form 'circles of trust' between trusted authentication service providers (ASP's) and service providers (SP's) [PM03]. Thus, organisations can make agreements with regard to the authentication of individuals and can provide individuals the possibility to use a specific identity within these circles of trust. This reduces the burden for individuals to cope with different identities within certain contexts. For enterprises, the benefit of Liberty Alliance are cost savings from sharing a standardised and interoperable architecture, and from outsourcing activities to identity providers.

Liberty Alliance, however, still relies on organisations that act as identity providers. It focuses on a business-to-business scenario [Pfi03]. Individuals therefore still need to be aware of linkability risks and need to be cautious when they choose privacy policies [PM03].

3.5.3 OpenID

OpenID is a decentralised SSO system, which chiefly aims at lessening the user's problem of having a multitude of passwords and usernames. OpenID-enabled websites relieve the burden for users to remember different usernames and passwords by only requiring them to register at an OpenID identity provider. The advantage of this is that people do not necessarily need to 'sign up' and 'log-in' for every single service on the internet within one browsing-session, but instead can go from one of the sites in the federation to the next once logged in. OpenID rises primarily out of the blogging community but currently both the amount of users and the number of places where OpenID identities can be used is growing rapidly [PR07].

OpenID works with an URL, owned and provided by the individual, that is used for authentication. Websites that require authentication can request the OpenID URL from the individual. The presenter of the OpenID URL is then authenticated by verifying the URL at the OpenID-URL issuer (the IdP). If the issuer certifies that the user actually belongs to the URL, authentication is complete.

OpenID makes use of credentials which are not stored at one single organisation or server. The users can decide for themselves whom they trust with their credentials. Several different OpenID providers already exist, also due to the ease of implementation of OpenID. In addition, OpenID provides a single individual the choice to develop and maintain several different identities at different OpenID providers. OpenID is therefore in the user-centric corner even though users still need to rely on some identity providers.

The OpenID authentication process depends on the redirection of a user to the identity provider's site. This process of redirection raises concerns with regard to 'phishing' attacks (described later in this chapter), because trusted sites can easily be imitated, resulting in a possible exposure of credentials and login information to distrusted parties. This is especially the case when a username and password are being used to login at the IdP's website. Furthermore individuals are still vulnerable to potential unlawful actions of identity providers that can store, collect, and link their data. Moreover, the real separation of contexts still depends on the creation of different accounts, at several servers, requirering extra effort from the individual. For many services on the internet, OpenID is a feasible solution, but some of its design aspects still make it difficult to apply, especially when it concerns 'sensitive' contexts.

3.5.4 Microsoft Cardspace

Microsoft Cardspace is an identity metasystem developed by Microsoft. It is incorporated in Microsoft's operating system Vista. The system uses the metaphor of 'information cards' for the representation of digital identities to provide the individual with a consistent and comprehensible user experience

[Cha06]. Users can create the information cards they want to use by themselves, but it is also possible to use information cards that are issued by third parties like banks, insurance companies, or government.

The Cardspace system claims to circumvent the widespread problem of 'phishing' that occurs when traditional, easily imitated, password-based, web login screens are being used. Microsoft Cardspace addresses the issue of phishing-attacks by 'taking over the screen' of the operating system. Cardspace manages identities at the end user's machine [Mal06]. Moreover, it is an identity metasystem, which makes it complementary to existing identity architectures, like the aforementioned OpenID system. In addition, Cardspace allows users to have different digital identities, regardless of the kinds of security tokens used by other systems. It is therefore also an 'agnostic' IDM system [Cha06].

The user of Microsoft Cardspace is positioned between the relying parties and the identity providers because the information cards are stored in the user's application, which can pass on the information cards to the relying parties when the user chooses. Thus, instead of having one or several organisations 'in the middle', Cardspace facilitates that the user is in the middle of issuing and constructing identities.

3.5.5 Other IdM Systems

There are many other IdM systems under development, for instance, *Higgins*, *Shibboleth*, *Bandit*, *WS-federation*, *Sxip* and *Kerberos*. The current brief overview of some of the leading systems suffices for the purpose of this chapter.

It is clear that there is no lack of competition in the identity management landscape [CMBG$^+$02]. The individual perspective until recently has received limited attention though. The same conclusion applies to the privacy aspects of identity management systems. Before turning our attention to these aspects in the following chapters, we briefly review some of the factors complicating the identity management landscape.

3.6 Complicating the Online Identity Landscape

The online environment in which individuals interact and maintain their identities is evolving. From a unidirectional source of information, the internet has become a realm in which many people interact with each other, businesses and the government. Enterprises and governments offer personalized services that require users to establish and maintain online identities. People also increasingly use the internet to maintain their social networks, to relax, to play, or to seek relieve. All these developments have an impact on how identities are constructed and used online and affect the risks that people and organisations take when they are online. In this paragraph we will describe some developments that emphasize the need for IdM systems in which both the personal and the organisational perspective on identity are represented.

3.6.1 The Internet as a Social Environment

The Internet is transforming into Web 2.0 [O'R05]). Instead of mainly consuming information provided by (professional) service providers, ordinary users increasingly actively participate in creating online content. Users are active in social environments and the 'bloghosphere', and contribute to wikis. The use of all social media platforms, such as weblogs, photo-sharing websites, social network sites, and chat rooms, has grown significantly over the last years [Uni08].

Social media change the collection and dissemination of news, provides commercial organisations new business opportunities, and influences social life and family situation. For example, the millions of existing blogs cover nearly every topic and dissolve the boundaries between professional journalists and amateurs [Sol07]. Social network sites have an effect on the nuances in social connections, and are likely to influence the amount and quality of ties that an individual can manage [Sol07, DB04, WG99].

Personal information does not necessarily have to be shared to maintain social relations via the internet. Individuals can also act anonymously in online social environments. Many people, however, do disseminate personal information percisely because they have an interest in the creation of social capital and reputation, and because a 'display of connections' is considered important [DB04]. Because many people make use of the internet for 'social purposes', much personal information (text, video, and audio) is therefore uploaded and shared. People leave digital traces everywhere. This does not mean that these individuals upload their personal information to 'the public', in the sense that it may freely be used by others. Context still matters, even in online social media. The ease with which information can be decontextualized and used 'out of context', however, undermines the sense of 'public privacy' and can lead to reputational damage (see for instance: [Sol07]), and identity fraud. In general current web 2.0 applications are not very well tailored to help people to segregate their audiences.

3.6.2 Customer Empowerment

Another aspect of 'Web 2.0' is a change in the way customers and organisations (enterprises and governments) interact. The internet appears to intensify the relation between users and organisations. Dissatisfied consumers post their grievances on discussion fora and blogs that can be read by fellow consumers. Enterprises increasingly monitor these media and actively engage in them in order to try to manage negative scenarios regarding their reputation. Moreover, technologies make it possible to use and process the ideas and suggestions of customers directly into the process of innovation, in line with managerial trends like 'open innovation' and 'democratic innovation'[1].

[1] Terms that were introduced by Henry Chesbrough and Eric von Hippel.

Via the internet, organisations can empower their customers, which creates an incentive to construct business models around (the knowledge of) the user. Hardware and software vendors, for instance, all have knowledge bases that are fed by their own staff as well as by users of their products. Information from users can be a key asset for organisations. The internet makes it possible to apply the 'wisdom of the crowd' to the benefit of the organisation, which means that collective intelligence can provide better insight in the requirements for services and products that need to be developed.

However, customer empowerment can also lead to more personalisation and personal data collection. These data can not only lead to better (tailored) products, but can also be used for the purposes of data mining, targeted advertisement, and discrimination.

Electronic services are provided on a global scale (web browsers need no passport to travel to different countries) and includes anything from health services (like providing medical records and medical information) to online gaming. This means that (personal) data relating to a rich set of activities flows across the globe crossing jurisdictions and policies regarding the collection and use of personal data and involving private and public entities.

3.6.3 Identity-Related Crime and Misbehaviour

The difficulty in properly identifying both individuals and organisations online has also drawn the attention of criminals. Online identities are valuable for criminals and people with harmful intent. Technologies have increased the opportunities for 'identity theft', 'identity fraud', and 'identity deception' (for definitions of the terms see for instance [KL06, KLM$^+$08]), because online identities are used in disembodied environments. The individual increasingly is physically absent when identification or authentication occurs.

Technological developments seem to have made it easier and profitable to abuse identities [MWB$^+$04]. Online financial services, for instance, have become a main target of cybercriminals (see [APW07]). Especially in the United States, identity fraud is a prevalent and fast growing form of crime [WF08, BMK07], and it has been assumed that also for Europe identity fraud is growing, even though less statistics are available for this region [LGM$^+$05]. The economic loss as a result of ID fraud for enterprises is significant [MWB$^+$04], but the negative effects do not stop there. ID-fraud can also seriously affect the trust of consumers in online services.

Identity abuse is, however, even more unpleasant for the individual. The economic loss resulting of ID abuse is often not the individual's biggest concern, but rather reputational damage, confusion, burden of proof, and the restoration of damages done are. The side effects of identity abuse may furthermore extend for years, for instance in the exclusion of services, accusations, or stigmatisation [MWB$^+$04].

One of the most popular methods of committing ID fraud is 'phishing'. Phishing concerns tricking people to reveal their confidential information by luring them to websites that resemble those of genuine entities where the user may have an account, or sending them e-mails ‘on behalf’ of such entities. The collected information can then be used by criminals to make purchases, or launder and transfer money [Oll04]. Especially in the US, phishing costs companies billions, and has led to ‘numerous consumer alerts and the creation of industry working groups’ [EI06, P. 58].

Criminal abuse of identities is not the only form of abuse. Identities can also be abused for activities such as bullying and betrayal. On weblogs and social network sites some people may for instance intentionally reveal another user’s identity or use another user’s identity for the purpose of deception or manipulation [DB04]. With wrong or revealed identities, people can provoke violent reactions, destroy the integrity of an online environment, and intimidate others.[2]

The use of the internet for the purposes of criminal activities, manipulation, or deception highlights a need for thinking about accountability or identifiability of individuals in specific contexts. Moreover, the potential use of the internet for terrorist activities or activism may even further intensify the ‘call for accountability’ on the internet. However, such a call for accountability can also lead to superfluous surveillance and supervision, because technologies also provide instruments for constraint, control, deception, and criminality. IdM systems have a function in the creation of the appropriate levels of accountability and freedom in online contexts. The increasing use and abuse of identities furthers the need for IdM systems which have the features that facilitate such a balancing act. It is a challenge to create IdM systems that allow for accountability, without the possibility of identity abuse, and without eroding the necessary level of privacy.

3.6.4 The Expanding Internet: Always-On and Everywhere

Internet penetration and the amount of households with a computer is increasing rapidly in Europe (see [Soc07]). People also spend more time online. The use of internet already overtakes the use of television amongst young people[3], and a significant amount of users spends more than 16 hours online per week (see:[EIA07]). However, at the same time, many people seem to be concerned about the amount of personal data they leave on the internet. The amount of digital data held on every person, are exploding [Hen08], yet only a minority of internet users employ tools that increase data security [Org08].

[2] Famous is the Megan Meier case on MySpace. Megan Meier committed suicide after a friend, Josh Evans, a false identity allegedly created by Megan’s neighbour Lori Drew, wrote that the world would be better off without Megan. See, for instance: `http://archives.chicagotribune.com/2008/may/15/nation/chi-megan-meier-myspace-080515-ht`.

[3] Which is emphasized by a recent IDC study, see ‘IDC Finds Online Consumers Spend Almost Twice as Much Time Using the Internet as Watching TV’ from 19 Feb 2008 on `http://www.idc.com`.

The increasing use of the internet will lead to a higher dependency on its infrastructure and on the services it facilitates. For some, the internet is a means to be 'always-on'. For mobile phones this is already the case for most users. The boundaries between work and private life diminish, many people leave their computers on and check their (work related) email in the evening and during weekends. Vice versa, private affairs are also conducted in the workplace; workers do visit websites for private purposes during working hours.

Mobile phones no longer are just phones, many are smart phones. They contain proper web browsers and email clients, and judging from the popularity of the Apple iPhone, this addition to appliances appears to be the best thing since sliced bread. Smart phones will likely increase the amount of time people spend online, which potentially means a further increase in the amount data trails people leave online. Given the fact that many smart phones also contain capabilities for determining the location of device (by GPS), which supplements the server side capabilities to locate devices (by GSM/GPRS or by WiFi positioning), the data trails can even be enriched by location data. Therefore, not only the user's behaviour, but also the location where this behaviour is exhibited can increasingly be monitored.

3.6.5 The Internet of Things and the Citizens of Tomorrow

In 2005, the ITU prepared a report on 'The Internet of Things', describing an evolution towards next generation 'always on' communications. We are moving from today's era of people-to-machines communication, from conventional Internet and mobile phones, to the era of machine-to-machine communication: the Internet of Things. In this new type of communication, new technologies, such as RFID, will enable the creation of networks with always interconnected devices. There are innumerous functions these 'things' will be able to perform. They will be able to "direct their transport, adapt to their respective environments, self-configure, self-maintain, self-repair, and eventually even play a role in their own disposal" [RFI08, p.3].

The Internet of Things will have radical effects on the way we interact with technology. Nowadays we are aware that we turn on our laptop or TV, the internet of things changes this. "It is all about making technology ubiquitous" [Sri06]. Ubiquitous computing may make individuals less aware that data is being disclosed and collected, much like many people are increasingly unaware of the camera surveillance that is becoming common in European cities.

In the today's world, the ratio of radios to humans is almost 1 to 1. The vision of the Internet of Things will challenge the very foundation of this landscape. In scenarios where even devices such as toothbrushes indicate electronically to remote devices that they need to be charged or when each light bulb in your house has a unique identifier, the ratio of radios to humans could easily exceed 1.000 to 1 [Sri06].

In the way to the networked era of the Internet of Things, also major changes are taking place with regard to identification documents. Electronic identification documents (ID documents) are seen as a necessary upgrades of important paper ones. RFID chips are chosen by the International Civil Aviation Organization and the European Union as the storage medium for data on the ID document holder. These chips have sufficient storage capacity to store biometric images and they are believed to ease the identity checks and enhance security. The equipment of ID documents with RFID chips is claimed to reduce fraud and prevent identity theft, as the ID document will not be easily tampered with. Furthermore the limiting of human inspection of the documents would help decrease the amount of errors made in the process.

The privacy and security risks that arise from the vast deployment of electronic ID documents are easily neglected. The RFID chips facilitate continuous tracking and tracing of individuals.[4] Unauthorised reading can not be ruled out and enormous databases with sensitive information about the individuals are expected to be created. The European electronic passport is already a reality and a many initiatives are currently ongoing regarding the introduction of electronic identity cards in Europe and several US States.

Besides RFID and similar technologies, the use of biometrics as identifiers is increasing dramatically. There is a transition from the traditional method of identifying yourself via something you have (key) or something you know (PIN) to something you are. A part of ones body is used as the means of identification and is the 'key' that allows her to have access to a restricted area, to operate a machine or to secure information.

3.6.6 Identifying the Individual in the Era of the Internet of Things

The Internet of Things depends on unique identifiers that will allow every-'thing' to communicate. But will every-'thing' qualify also as personal data? Will every-'thing' be linked to an individual? Will our perception of personal data need to change in order to tackle the challenges posed by this new situation?

The European legislation on data protection applies when processing of personal data is entailed. According to Article 2(a) of the Data Protection Directive personal data shall mean "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity." Therefore in order to

[4] The European biometric passports do implement access control security measures, but these are not unbreakable as various studies have shown (see for instance http://www.guardian.co.uk/technology/2006/nov/17/news.homeaffairs for a story about the UK passport).

define whether some information qualifies as personal data, we need to assess firstly if the processed data relate to a natural person, and secondly whether the data relate to an individual who is identified or identifiable [PN07]. The latter question is the one that stimulates vivid discussions.

When information refers directly to an individual, such as his name, age, nationality etc., it is beyond doubt that it qualifies as personal data. The qualification is more challenging when the information can not be directly linked to a natural person, i.e. when the person is only "identifiable". Recital 26 of the data protection directive reads that in deciding whether data could be used to identify a particular person "account should be taken of all the means *likely reasonably* to be used either by the controller or by any other person to identify the said person" (emphasis added). Thus the recital sets two criteria for identifiability: the probability and the difficulty that tend to be interlinked [Byg02].

The national legislation of the European Member States and their interpretation by the national Data Protection Authorities construe the concept of identifiability in different ways. The data protection laws of France, Belgium and Sweden, for instance, have adopted a broad interpretation of the concept of personal data, rendering any information as personal data if an individual can be identified, regardless of the technical or legal difficulties in determining the identity of the individual. The German legislation, on the other hand, has adopted a more pragmatic approach to the notion of identifiability. The German Federal Data Protection Law in article 3(6) defines the notion of 'Anonymisation' as follows: "Rendering anonymous' means the modification of personal data so that the information concerning personal or material circumstances can no longer or only with a disproportionate amount of time, expense and labour be attributed to an identified or identifiable individual". The definition of anonymisation allows the deduction of the following *argumentum a contrario*: personal data are information that can be attributed to an identified or identifiable individual without a disproportionate amount of time, expense and labour.

These issues are not merely semantic battles for cocktail receptions. The 'battle' surrounding the question whether IP addresses are personal data between search engine providers (such as Google) and the European data protection authorities is centered around this issue. The Article 29 Working Party in its opinion on IPv6 sustained that IP addresses attributed to Internet users are personal data [Par02]. The same opinion was supported a few years later, where the Article 29 Working Party confirmed its opinion that IP addresses are personal data and noted that "unless the Internet Service Provider is in a position to distinguish with absolute certainty that the data correspond to users that cannot be identified, it will have to treat all IP information as personal data, to be on the safe side" [Par07].

However opposite opinions have also been expressed, presenting significant argumentation. Google, by means of its Chief Privacy Officer, Peter Fleischer has taken the position that IP addresses are not personal data (most of the time).[5] Fleischer quotes the Secretary for Home Affairs of Hong Kong (Dr Patrick Ho), who maintains that: "An Internet Protocol (IP) address is a specific machine address assigned by the web surfer's Internet Service Provider (ISP) to a user's computer and is therefore unique to a specific computer. An IP address alone can neither reveal the exact location of the computer concerned nor the identity of the computer user. As such, the Privacy Commissioner for Personal Data (PC) considers that an IP address does not appear to be caught within the definition of "personal data" under the PDP."[6]. Although it is obvious that Hong Kong does not fall under European law, the argument expressed by Dr Ho can be valid in the current debate on IP addresses in Europe.

IP addresses will be of seminal importance in the Internet of Things era, as every little 'thing' will have an IP address that will allow it to be networked and interconnected. However it will become even more difficult for an ISP "to distinguish with absolute certainty that the [IP] data correspond to users that cannot be identified" [Par07], as required by the Article 29 Working Party. The example of IP addresses clearly illustrates the difficulties in defining whether a piece of information shall be considered as personal data or not.

3.7 Conclusion

This chapter has provided a first glance at the identity management landscape. It has introduced two different perspectives on identity management, an enterprise view and an individual view. The enterprise view concentrates on access control to resources that is usually implemented as a system of user accounts. Each account specifies which user is entitled to which services. Identity management in this perspective is closely tied to Identification, Authentication and Autorisation. The individual perspective, on the other hand, is based on the way individuals manage their identity in everyday life. Identity in this view relates to the way individuals present themselves to others and how others view them. As people engage in different (kinds of) relationships, they display different aspects of their identity. What is shown in the private setting of the family differs from what is shown in the workplace or during shopping. Identity management in this view is (unconsciously) deciding what image of self to show to others in a specific context. Individuals may present themselves as the same across contexts (I may tell my employer that I am indeed the famous tennis player by the same name) or as different (I may not tell my grocer that I work in Tilburg, even though he has seen a picture of me on the website of Tilburg University).

[5] See for instance his blogspot: `http://peterfleischer.blogspot.com/2007/02/are-ip-addresses-personal-data.html`.

[6] http://www.info.gov.hk/gia/general/200605/03/P200605030211.htm, as quoted on Peter Fleischer's blog (Chief Privacy Advisor of Google).

Context shapes how identity is constructed and maintained. The identity of an individual can be said to consist of the sum of various partial identities displayed in the different contexts. This does not, however, mean that all data related to these various partial identities can be combined into 'one' identity. Data associated to the various partial identities is contextual and therefore the combination of data may lead to seemingly inconsistent pictures.

Identity management developments are until recently, driven by the enterprise perspective. Originally each entity requiring access control developed and maintained their own solution for implementing access control. The result of this has been a plethora of fragmented and incompatible IdM solutions. For individuals the consequence of this landscape is that they have many online identities that are composed of similar data that was disclosed to each and every of the enterprises. Furthermore, the user has little control over the identity they want to present to the various enterprises. Their freedom to present themselves as the same or different is limited by the restrictions imposed on them by the IdM systems.

In recent years, a move towards identity federation can be observed. Enterprises collaborate and design systems that allow interoperable identity provisioning and access control. These developments primarily solve enterprise needs because these systems lower their expenses in setting up and maintaing IdM systems. Also the user benefits from the single sign on functionality offered by federated IdM, but the lack of control over the identities to be used largely remains.

A step further is the move towards user-centric IdM where the individual is at the steering wheel. The individual creates and maintains her online identities and populates these with credentials obtained from the various identity providers. The level of control over the presentation of self can be significant in these systems.

Not only the unification of the enterprise perspective on IdM with an individual perspective is challenging. We have also described a number of technological developments that complicate identity management. Users are changing from consumers to producers of content (Web 2.0). They actively engage in social networks, blogs and wiki's and disclose data on the go. Furthermore, technology is increasingly becoming pervasive and ubiquitous. More and more device are networked and connected. This raises questions regarding the identification of things in what is called the Internet of Things. As things are used by humans, there clearly is a link to the identification of humans and to identity management of humans and things. The developments make clear that the existing concepts on which data protection and privacy regulation is built no longer self evidently adequate.

The IdM landscape is evolving rapidly. Until recently privacy concerns hardly have played a role here. As we will argue in the following chapters, this needs to change and we will show that this is indeed possible.