

## Chapter 2

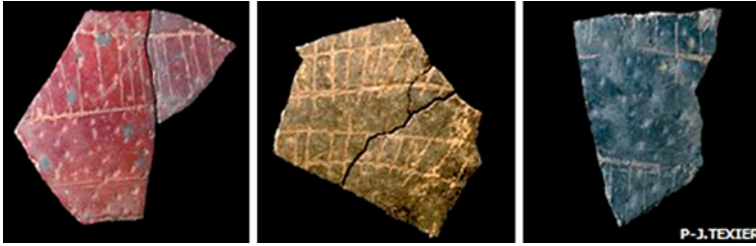
# Digital Identity

*An identity is questioned only when it is menaced.*  
James Baldwin (1924–1987), American novelist  
and civil rights activist

### 2.1 Introduction

Having an identity and expressing it have been of that importance from the early time. Inscribed ostrich shell fragments found in Diepkloof Rock Shelter in Western Cape, South Africa are among the earliest examples of the use of symbolism as a form of expressing identity.

Figure 2.1 shows three over 270 pieces of decorated shells are dated to about 60,000 years ago, 20,000 older than cave painting, which was considered presently the first form of writing in history. At that time the ostrich eggs were used as bottles once engulfed their content. The researchers, who have investigated the material since 1999, argue that the markings are almost certainly a form of messaging—of graphic communication [1–4]. Dr. Pierre-Jean Texier from University of Bordeaux, France explains: “the lines are crossed at right angles or oblique angles by hatching. By the repetition of this motif, early humans were trying to communicate something. Perhaps they were trying to express the identity of the individual or the group” [1]. In the ancient near east excavation brought to light a group of clay tablets and wooden boards, dating to the middle of the third millennium B.C., on which Sumerian and Akkadian inscribed identity information of the collectivity such as geographical names, names of gods, names of rulers, names of exorcist, and hymns, legal documents, medical records, and lists of professions. In addition, they wrote in colophons individuals’ names such as authors and tablets’ collectors [5, 6].



**Fig. 2.1** An old representation of the identity of an individual or a group [1]

## 2.2 Identity: Yesterday and Today

The term ‘identity’, which is firstly known used in 1570, has been used in many different ways in academic research and in popular usage [7]. The term is still of disputed origins, but it’s certainly true that its origin derives from Middle French ‘identité’, from Late Latin ‘identitat-, identitas’, or probably from Latin ‘identidem’ repeatedly, a contraction of ‘idem et idem’ and literally ‘same and same’ [8]. In the American Heritage Dictionary of the English Language, the term ‘identity’ could refer to ‘the collective aspect of the set of characteristics by which a thing is definitively recognizable or known’, ‘the set of behavioral or personal characteristics by which an individual is recognizable as a member of a group’, ‘the quality or condition of being the same as something else’, ‘the distinct personality of an individual regarded as a persisting entity’, or ‘information, such as an identification number, used to establish or prove a person’s individuality, as in providing access to a credit account’ [9].

In the pre-modern times, human identity was defined by geography, community, and family relationships. If an individual was born into a well-known and rich family in London, that is typically the environment in which he or she would remain. If an individual began life in a poor remote community in India, they would typically not be able to change their life pattern or economic status over time. One’s geophysical space and one’s place in society were inextricably linked, the possibility of freedom of movement being severely limited. With modern times there arrived a greater choice for participation in different social circles, and the possibility of social and economic mobility. Today, most people carry some form of identification on them at all times, but this practice is relatively recent in human history. In the past, the declaration of an individual’s name, sometimes accompanied by the name of their city or village, was sufficient to prove their identity. This is no longer the case. Further, the notion of identity today can refer not only to humans, but extends to animals, machines, and other objects or resources. A machine may have an identity which would allow it to access certain information at certain times, or be employed by some individuals, to the exclusion of specified others [10].

## 2.3 Identity Perspectives: Multiple Facets of the Identity

For many centuries, stories of the holy fool Mulla Nasrudin's have been studied in Sufi circles for their hidden wisdom [11]. One of the stories tells of Mulla Nasrudin who traveled to another city. Before he left on his journey, his wife put a sign around his neck with his name on it so that he would not forget his identity. In his way, he spent a night at a caravanserai; while he slept, a joker took the sign and put it around his neck. When the Mulla awoke, he was appalled to find his name tag on the joker's chest. He cried: "It seems that you are me. But if you are me, then who am I?" [12, 13]. The Mulla's dilemma is a ridiculous one, but it illustrates the importance of identity and introduces the multiple perspectives aspect in studying the identity. The Mulla question is one of the key questions in the philosophical debate over identity. The Mulla dilemma touches one of the central identity-related issues in social, cultural, and child and adolescent sciences. In addition, the story could illustrate the importance of losses related to identity theft in digital economy and e-commerce.

Identity concept is seen from different perspectives and applicable into different domains. We describe here multiple perspectives of the identity and mention few major issues from each perspective. The identity debate dates back to ancient world's philosophy. In general, personal identity in philosophy is employed referring to Who am I? It consists roughly of those properties that make the individual unique and different from others [14]. From the same perspective, identity refers to a set of qualities and characteristics that make an entity definable, distinguishable, and recognizable comparing to other entities. In recent times, many philosophers have given attention to the question of change impact over time on the personal identity continuity such as Aristotle that distinguished between 'accidental' and 'essential' identity changes. Accidental changes refer to identity properties changes such as hair color change, while essential changes are radical and don't preserve the identity like someone, who dies. Other concepts arise such as numerical and qualitative identities. Numerical identical is the same one: one thing rather than two, but qualitative identical is exactly similar two things such as twins. The 'personal identity' has addressed the conditions to stay numerical identical throughout time [10, 15]. 'Identity formation' is defined as the process of the fabrication of the distinct personality of an individual in a particular stage of life such as establishment of a reputation. In this context, pieces of identity include a sense of personal continuity, a sense of uniqueness from others, and a sense of affiliation. These pieces could help people to define their selves in the eyes of others and themselves [16]. From the mathematical perspective, the law of identity in logic is upheld by a reflexive relation, states that an object is always the same as itself ( $A = A$ ). In mathematics, the term identity denotes several meaning. Specifically, in algebra, the identity function  $\text{ids}(x) = x$  for all  $x$  in the set  $S$ . The identity matrix includes ones on the main diagonal and zeros elsewhere. In social science, we use the term identity referring to an individual's comprehension of himself as a discrete, separate entity [17]. From the legal perspective, protection

policies of sensitive identity-related information policies are critical, and privacy regulations are on the rise. Although from a technology viewpoint, the priorities may be authorization and control, what seems to be different and evolving is the notion of equipping the end user with the necessary controls to protect his identity information: Users are informed about what data is requested from them and how their personal data is treated, e.g. for what purpose it is used and who can access. Through this process, users can decide whether to provide their data and to consent to the service provider's data handling policies. Ideally, the service provider employs technical components such as access control systems to enforce the consented policies; for instance, to ensure that a user's e-mail address is not used for marketing but only for the consented billing purpose. From the cultural perspective, cultural identity deals with the influence of an individual identity by his belonging to a group or a culture [18]. Other questions and issues arise such as ethnicity, citizenship, nationhood, and how culture could influence on emotion, thoughts and self. In his book 'Culture and Identity' [13], Charles Lindholm states that since the late nineteenth century, psychological anthropology scholars study of the relationship between the individual's identity and culture. The discipline addresses also fundamental questions about the nature of humanity that have become pressing in the present era of multiculturalism and globalization. In social sciences, identity is a modern formulation of dignity, pride, and honor. One of the key question related to identity in social sciences is 'Who is we?' referring to the concept of social identity complexity [19]. It deals with an individual's subjective representation of the interrelationships among his multiple social group identities. The same authors mention that membership in many different groups, multiple social identities, can lead to greater social identity complexity, which can foster the development of global identity. From the economic perspective, particularly in marketing, a corporate identity is visibly manifested by the use of trademarks and the way of branding. Corporate identity is established when there is a common ownership of corporate philosophy, values, and norms that help the attainment of business objectives [20]. In their book titled 'Identity Economics' [7], the authors demonstrates how identities shape the employees' work, wage, and well-being. In psychology, a 'psychological identity' is related to self-image, self-esteem and individuation. It might be defined as a network of values and convictions that structure the individual's life. Moreover, it considered also as a property or a set of properties that an individual might have for a while and then lose, thus, he would acquire a new identity or perhaps carry on without one [10]. The family therapist and child psychiatrist Salvador Minuchin provides psychological definition of identity. He declares "the human experience of identity has two elements: a sense of belonging and a sense of being separate" [21]. From computer science and information technology perspectives, digital identity, online identity and others concepts have emerged. We cover these aspects further on this dissertation. From history, anthropology, and archeology perspectives, identity refer to human origins and identity construction over time. From genetics perspective, major issues are addressed such as genetics and origins of species, and how molecular genetics influence human personalities. From the art perspective, we mention architecture

and identity such as architecture in Islamic culture. We mention the religious perspective to point people may see their identity as defined partly by some moral or spiritual commitment such as Islamic, Catholic, Jewish or anarchist. Or they may define it in part by the nation or tradition they belong to as an Armenian or a Québécois [22]. Finally, from the political perspective, many ongoing debates are over ethnic, race, gender [23], national, and transnational identities [24]. From the sociological perspective, the author [25] provides definitions and the distinctions of ‘identity’ and ‘identification’ concepts. ‘Identity’ denotes the ways in which individuals and collectivities are distinguished in their relations with other individuals and collectivities; and ‘identification’ is the systematic establishment and signification, between individuals, between collectivities, and between individuals and collectivities, of relationships of similarity and difference.

## 2.4 Digital Identity: Definitions, Basics and Nomenclature

Digital identity is composed of two distinct words that we explain each one separately: (1) ‘identity’ is what makes individuals the same today as they were yesterday (sameness), but it is also what makes them different from one another (uniqueness). Though these fundamental concepts have remained the same over time, changes in economic and social structures have affected the determination and perception of identity. Identity is the distinction between the private and the public spheres of human existence, and as such identity and privacy are forcibly linked [10]. As the boundary between the private and the public in the digital age becomes increasingly blurred, the creation and maintenance of secure identities online has emerged as an important priority for businesses and consumers alike. The researchers [26] define ‘identity’ as a set of personal information and identity management system as authentication and attribute management system. While, [27] defines the identity establishment concept as ‘the representation of methods by which, a user, a running process, or a thread of execution is securely associated with a legitimate entity’. The author states that the goal of ‘identification and authentication (I&A)’, which is the process of establishing a user identity, is to provide to the entity access only to authorized computer resources. However, [28] restrict the entity definition to people or organization and define the identity, within a specific application domain, as an entity representation through a generation of a unique key, which combines all the elements of identity information. The researchers [26] define ‘identity’ as a set of personal information and identity management system as authentication and attribute management system. While, [27] defines the identity establishment concept as ‘the representation of methods by which, a user, a running process, or a thread of execution is securely associated with a legitimate entity’. The author states that the goal of ‘identification and authentication (I&A)’, which is the process of establishing a user identity, is to provide to the entity access only to authorized computer resources. However, [28] restrict the entity definition to people or organization and define the identity, within a specific application domain, as an entity representation through a generation of a unique key, which combines all the elements of identity

information; and (2) in the Webster's New Explorer Dictionary, the word 'digital' means 'done with the finger or toe' and narrowly, a 'digital computer' is a mean by which 'provides a readout in numerical digits'. In today's ordinary technological parlance, 'digital medium' refers to machines that are capable of recording, transmitting, or receiving data in binary digit form. In addition, people are getting connected by consuming an increasing amount of digital media and broadband technologies, such as internet and mobile phone. We present in next sections a literature review of the definitions, basics, and preliminaries of digital identity. Digital life is designated to represent a daily life where individuals use digital mediums and technologies to engage activities in online and offline worlds. In the entitled Digital Life Internet Report [29] published by International Telecommunication Unit (ITU), the United Nations specialized agency for Telecommunication, experts in policy and strategy state that today's digital world is transforming individual lifestyles. Always-on internet access has become a global norm and daily lives has brimmed with SMS, e-mail, chats, multiplayer online gaming, virtual worlds and digital multimedia. But what does it mean digital, digital media, digital world, etc.?

Several definitions of the term 'digital identity', from different perspectives, have appeared in the literature. A simple definition is related to one of identity. Thus, identity is defined as a collection of data about subject that represent attributes, preferences, and traits [29], so in parallel, in the digital world a person's identity is typically referred to as their digital identity [29]. The term 'digital identity' has emerged through the evolution of the Internet. Wherever we go, we leave traces of fragmented information about our identity. Leaving a comment in a forum, filling out a form, maintaining a blog, creating a full profile (photo, name, phone number, etc.) in a social network, conducting a parallel existence, we are educating others about what we are, what we do and especially what we think and then constructing 'digital identity'. Internet users are striving to share their digital identity with others to re-enforce their online presence and one of the favorite users' activities on the net is egoGoogling. A 'personhood' means that we recognize that an entity or individual has a person's status and the 'digital personhood' means the person's status projected in digital environment [30]. The authors [31] suggest a conceptual definition of the term 'digital identity'. It refers to two concepts: 'nyms' (called also masks or aliases) and 'partial identities'. In his book [29], Windley defines a digital identity as the data that uniquely describes a subject or an entity and the ones about the subject's relationships to other entities'. The author gives the car title as example of digital identity. The car title contains vehicle identification number that uniquely identifies the car to which it belongs and other attributes such as year, model, color and power. The title contains also relationships such as the set of car owners from the time it was made. From technical perspective, the same author explains that digital identity is built on a set of technologies that includes cryptography, authentication, authorization, identity provisioning, directories, digital rights management, identity federation, and interoperability standards. In contrast, the author [27] does not distinguish between identity and digital identity. He provides a broad definition of identity from a computing perspective as 'a computer representation of an active entity that can be physical (such as human,

a host system, or a network device) or a programming agent'. In the lexicon [32], the authors coincided digital identity and identity definitions as 'a representation of a set of claims made by one party about itself or another data subject' but the authors of Princeton University Wordnet [33] don't distinguish between the two concepts by arguing that either in the real or electronic worlds, an individual may have multiple identities. The same authors point out that identity entails 'individual characteristics by which a person is recognized or known' [33]. The authors of the definitions paper of OECD [30] report insist on the difference between the two concepts by defining the 'identity' as 'a limited notion of set of claims', whereas the 'digital identity' as 'a thing or an artifact that refers to a person'. Adam's speech and Adam's ID card are two claims of the same individual. Based on works of Jenkins [25] and Goffman [34], Professor Shirley Williams of the University of Reading, UK [35] distinguishes also between the identity as 'a social performance' and digital identity as 'performances in digital places', which means the persona that an individual presents across all the digital spaces. He explains that human identity is naturally social and always involves, in addition to agreement and disagreement, convention and innovation, communication and negotiation, a performance, which denotes the activity of an individual which occurs during a period. He highlights that digital reputation and trust are other people's interpretation of the person's digital identity [35]. Moreover, the authors [30] highlight the referential and partiality natures of identity. Referential because claims must refer to a person and partial identity refers to 'a subset of identity information as the thing may not be sufficient to identify a person at different moments in time'. They add that the term 'digital identities' is a synonym of 'partial identities' in which a set of identity attributes are enclosed [30]. Digital identity is considered as an intersection of identity and technology in the digital age [36]. The author of the Digital Identity book [29] points out that identity is crucial to enable the virtual 'place'. He adds that digital identity will ensure that internet infrastructure respond to multiple needs including security, privacy, and reliability.

The world of digital identity has its own nomenclature. The following terms are derived from [27, 28], Windley's book [29], SAML-OASIS glossary [37], Liberty Alliance Technical glossary [26, 38, 39]. An 'entity' represents an active element of a computer/network system. It could be a single person; a group of persons, an automated process, a set of processes, a software program, a subsystem, an entire organization, a machine, a host system, a networking device or in general other thing making a request to access a resource. An entity's access to a system is encapsulated an 'account' and the 'principal' is the internal representation of an active entity in a specific environment. 'Attributes' describes a property associated with the subject such as physical trait, network address, medical record, purchasing behavior, bank balance, credit rating, dress size, and age. Attributes can also include preferences and traits. 'Preferences' represent desires such as preferred seating on an airline, brand of ice cream, and preferred language, and used currency. 'Traits' are like attributes but two differences are noticed between them: traits are inherent rather than acquired, and attributes may change but traits change slowly. Examples of traits are person's blue eyes, hair color, company's location and date when it was incorporated. Since



the distinction between attributes, preferences, and traits rarely makes a difference in the design of an identity system, we will typically use, in this dissertation, attribute to mean all three unless there's a need to distinguish among them. Attributes are often represented as pairs of attribute name and attribute value(s) and might be conveyed through an 'attribute assertion'. An 'Attributes Authority' (AA) manages the identity store and provides to IdP the requested attributes in the desired format such as through an attribute assertion. An 'identity store', 'repository' or 'directory' refer to any technology that could be used to store identity attributes such as the LDAP directories, databases, and files. Attribute 'scheme' or 'schema' represents the definition of the structure and the form of attribute held in a directory or database. 'Enrollment' is the process by which an identity of entity is created in a specific identity system. A 'Service Provider' (SP) interacts with entities primarily via HTTP and provides service to the user through a medium such as a portal (e.g. an online retailer, a financial institution, a government agency). An 'Identity Provider' (IdP) provides identity attributes to other providers (e.g. telecommunication company) and it may act as an authentication service provider. Note that 'provider' can refer to either SP or IdP and could interact and discuss details behind authentication. 'Attribute aggregation' is the ability to collect user attributes from IdP(s). An 'identifier' is used in two senses: (1) one that identifies; (2) uniquely refers to the system entity. Essentially, an identifier is a distinguished attribute of an entity. 'Credentials' are transferred data in order to establish a claimed entity identity and they allow transferring trust between subjects. 'Identification' is the process of using claimed or observed attributes of an individual to infer who the individual is. An 'identifier' points to a subject and it could be a name, a serial number, or some other pointer to the individual being identified. 'Pseudonym' is a name or label that may identify an individual within a system but does not correlate to that individual outside of the system. 'Secondary use' of information represents any use of identity or linked information that is inconsistent with an identity system's purpose. 'Authentication' is the process of establishing confidence in the truth of a number of claims. Finally, the following definitions drawn from the glossary of terms and definitions of the Ofcom research report [40]. 'Avatar' is defined as 'a computer user graphical representation of him or herself. An avatar can be two or three-dimensional'; a 'Profile' as 'the personal homepage on a social network site, usually including information about a user, photos, and their friend list. Profiles form the basis of social networking sites'.

## 2.5 Digital Identity, Security and Trust

Digital identity related mechanisms are the core of modern systems, networks, and applications security. In the book [27], the author considers that anonymity is not a desired computing goal but secure identification of users is the core element of computing security. He adds that the level of security is attached to an authenticated identity associated with it. The ultimate goal is to enable deterministic accountability and lay the foundation for responsible and secure computing [27]. Narrowly,



identities are critical to define access control policies [41]. Identity is having more importance in the online world. In the offline world, anonymous transactions can be conducted successfully, but in the service-oriented online world, we have to know something about the service recipient. Building digital identity infrastructures is an attempt to establish a community of trust, which becomes a requirement for conducting online business [29]. For instance, eBay community of trust lays on users' reputations. Windley [29] points that in order to make use of digital identity; organizations are required to understand other concepts such as trust and privacy. Corporations are considering identity infrastructure to provide security so that interactions with customers, partners, employees, and suppliers become more flexible and richer. The business should not be limited to just transactions, but relationships with customers, employees, suppliers, and partners and identity tends to change this relationship from one-way to a more customized one. Therefore, agile, business-responsive IT infrastructure should have at its core a flexible, interoperable identity infrastructure.

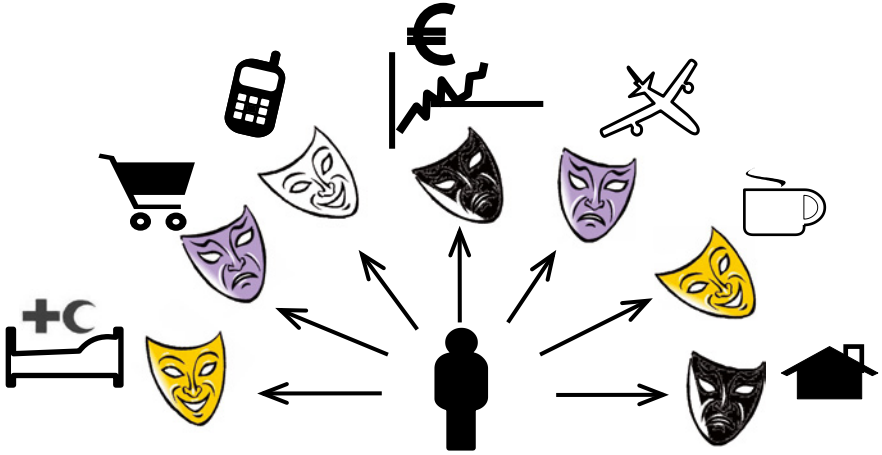
## 2.6 Digital Identity: Major Issues and Complexities

*There is no a single problem of personal identity,  
but rather a wide range loosely connected questions.*  
Stanford Encyclopedia of Philosophy

We do not intend to cover in this section all the issues related to digital identity rather than pointing couple of major issues and complexities.

### 2.6.1 Mutation from One YOU to Multiple YOUS

Currently, people are maintaining multiple identities. From the social science perspective, the recognition of an individual has no one, 'personal self', but rather 'several selves' that correspond to widening circles of group membership. Thus, an important issue that has been addressed is how individuals combine these different identities when they want to define a subjective identity within a social group? [10] Currently, the latter question becomes applicable to the digital/online world and being subject of many studies and researches. The authors [42] mention that the online world encapsulates a growing amount of scattered and unordered fragments of users' identities due to two major reasons. The first is because of the lack of a robust generic identification system and the second is the intentional creation of users' alternate identities. Figure 2.2 is an illustration. Creating more than one identity can be desirable for users depending on the context. A user may wish to be aggressive and egotistical in online multiplayer war game, but sensitive and sociable for virtual encounters and social networks. Thus, the online world represents an ideal nameless and faceless environment for users to easily create multiple representations of their identities:



**Fig. 2.2** Digital masks and partial identities

‘digital personae’ [42]. However, we usually speak of identity in the singular but in reality it is plural because it encapsulates multiple identities, ‘perspectives’, or ‘facets’ [29]. Researchers at Stanford University’s Virtual Human Interaction Lab don’t distinguish between ‘digital you’ and ‘virtual you’ and they consider them as synonyms of digital clone, avatar, nym, personae [42], which strongly influences the ‘real you’ [44]. In contrast, the authors [45] defines the ‘virtual you’ as a representation of a virtual version of the subject in the virtual world.

In ITU 2006 “Digital Life” report [43], the authors mention that ‘nyms’ and ‘profiles’ provide the subjects interacting capabilities with other parties in different environments. For example, nyms enable subjects to exercise their freedom anonymously in digital life by setting up synthetic personae complete with attributes such as age, race or religion. Another example is ‘social profiles’ that are created in popular social Web sites and online networks such as MySpace,<sup>1</sup> Bebo,<sup>2</sup> and Facebook<sup>3</sup> could be useful by allowing the users to post and share content, and staying in touch with others. Actually, the users log-in with pseudonyms in order to preserve anonymity that what make these networks attractive but in the other side anonymous users could engage malicious activities.

Avatars could enable online interaction and business opportunities. An avatar is ‘a graphical personification or incarnation of a user in a shared virtual reality space, more specifically, in online role-playing games and virtual universes (e.g. Second Life<sup>4</sup> and Active Worlds<sup>5</sup>) for a specific objective’ [43].

<sup>1</sup> <http://www.myspace.com>

<sup>2</sup> <http://www.bebo.com>

<sup>3</sup> <http://www.facebook.com>

<sup>4</sup> <http://www.secondlife.com>

<sup>5</sup> <http://www.activeworlds.com>

### Choose a starting look

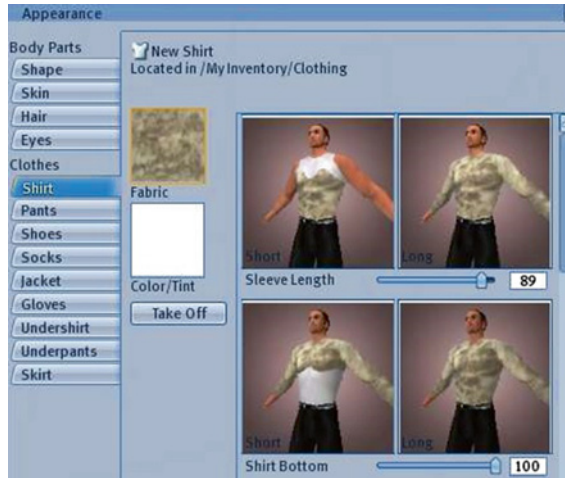
Click on images below to select a starting look. Once in Second Life, you can change your appearance, or shop for a whole new look.



You in Second Life

**Fig. 2.3** Selecting the right ‘you in second life’ (avatar)

In the Second Life, abbreviated “SL”, the user can choose multiple avatars with speech and language capabilities, Fig. 2.4, to participate within different virtual situations, such as virtual meeting, virtual tutoring and virtual commerce using virtual currency Linden Dollar (L\$). The use of avatars has been extended to online social networks and forums and is affecting the identity construction such as the phenomenon of gender switching when the user uses opposite sex avatars [43]. In his book ‘Coming of Age in Second Life’ [46], the anthropologist Tom Boellstorff stresses the important role that avatar plays in everyday activities in SL. He says: “a man spends his days as a tiny chipmunk, elf, or voluptuous woman. Another one lives as a child and two other persons agree to be his virtual parents. Two “real”-life sisters living hundreds of miles apart meet every day to play games together or shop for new shoes for their avatars. The person making the shoes has quit his “real”-life job because he is making over five thousand U.S. dollars a month from the sale of virtual clothing” [46]. Besides providing a comprehensive introduction to social, economic, political, and cultural settings in which the new media operate, the author of the book [47] presents multiple reasons why people might take the opportunity to explore different identities, including: (1) the ability to change character and physical traits at will, as illustrated in Fig. 2.3. This will provide to users the opportunities to explore other forms of existence and change the ways in which they may be perceived by others; (2) the opportunity for shy people or those who are uncomfortable with face-to-face interaction to form relationships and express views freely; (3) the potential to bring geographically and socially disparate individuals together based on common interests, thereby stimulating dialogue and curbing loneliness. An avatar could represent the offline personality of the user or another more desirable personality that the user cannot construct and afford in the offline world. In the online world, contact with strangers is encouraged and expected. It is acceptable to exaggerate, hide, alter or undermine the truth about oneself in order to encourage constructing desirable online impressions or reputations. In the paper titled “the connected identity” [48], the author confirms the presence of a relationship between the image of the visual interface, such as visual pseudo or avatar, and the



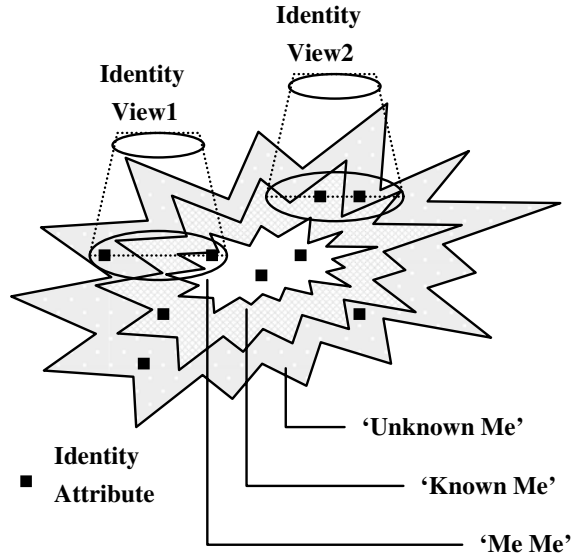
**Fig. 2.4** Customizing the “appearance” of the second life’s avatar

personality of the individual. He explains that the image reflects the personality of the user and shows who he really is. Identity has been building on freedom of expression through various media (videos, photos, blogs, avatars, music) and, so far, it is gaining upper hand over the sense of the community belonging [48].

Distributed attributes, which represent multiple set of different attributes within different environments, is a consequence of a context-based nature of identity concept. Partial identities [49] or digital selves [50] are any subset of attributes associated with entity that the entity itself can select for interacting with other parties. In the real-life, various forms of identity are required for various contexts in which the identity is to be presented in a suitable way and within suitable information by the identity holder [49]. For instance, the person A, as a traveler, is asked to provide a passport at the counter of customs or immigration to proof his identity; the person A, being a car driver, is asked to show his driving license to a police officer, who stopped him in the highway; the person A, as a customer, is asked to provide his credit card with fidelity saving card in a movie store to take advantage of DVD prices reductions; the person A, as a student, is requested to show his student card to have access to computer lab facilities; the person A, as a patient, is asked his medical card in the hospital to receive health services. Thus, for each domain, a specific partial identity is provided for identification. The partial identity can be named or unnamed, which means it might or might not be related to the entity’s true identity. In order to establish trust between parties in the digital world, a subset of digital identity attributes needs to be communicated. Digital identities exist in specific contexts and the contextual relationship between them is crucial to managing transactions and interactions. The context will determine which subset of attributes is required, or which “partial identity” will establish enough trust for the transaction to go forward. At the organizational level, identity attributes are distributed over different environments such as files, enterprise directories, databases, and online

social networks [49]. Similarly, in the online world, the authors [42] mention that users easily create multiple representations of their identities called ‘digital personae’. The author [29] calls multiple identities or ‘personas’ that the subject holds as digital identity ‘perspectives’ or ‘views’, which represent different perspectives on who is the subject is and what attributes he processes. They represent also a set of attributes that other entities have and can access. For instance, a bank sees account attributes, a physician in a hospital sees health record attributes, a district police sees criminal records attributes, and the employer sees other attributes such as full name, social security number, and one bank account number for paycheck deposit. The same author, in his book [29], explains that identity attributes are classified and organized into three sets, called ‘tiers of identity’ [51]. Each identity tier maintains relationships with other and its perceived value by the subjects is different. Tier1, labeled ‘My Identity’, includes attributes and traits associated with the subject such as hair color. ‘Shared Identity’ is the label of tier2, which consists of the attributes that are assigned to the subject by others in the sake of identifying him temporarily within a specific context and based on some kind of relationship. Driver’s license, credit card, health insurance card, library card are all examples of shared identity. Once the relationship that defines the identity is terminated (when the context changes), the attributes associated with it are no longer useful. Tier3 is largely about profiling; it deals with ‘Abstracted Identity’, which establishes abstractly the identity of a group. Marketing companies provides abstracted identity by classifying a subject as a male over 50, a Swiss Air frequent flyer, and a Geneva resident. The same author points that commonly the subject perceives the value and benefit of the tier2 identity relationship, which is usually established with his consent to meet a real need, however, tier3 relationships are usually forced on us and they rarely meet a real need. He states unsolicited commercial email or spam as a tier3 identity issue. The same author stresses that major identity issues that face organizations deal with tier2 identity. He stresses that how employees and customers perceive the effort to build digital identity infrastructure depends on their sensitivity to tier3 identity issues and their satisfaction towards the added-value of tier2 identity relationships. Tier2 relationships are dictated by organizations and consented by the individual. The one-way relationship is likely to change as service-oriented economy emerges. The power-shifts are brought on by increasing available services and improved systems that make it easier for customers to switch their allegiances; and more customized services would make it more likely that customers dictate their terms in their relationships. More specifically, the power-shift is the switch from the world of “take-it-or-leave-it” to “mass customization”. Good business would recognize these shifts. The author mentions two fallouts from the identity power-shift: (1) identity aggregation: multiple tier2 relationships create identity silos. From the user perspectives, multiple identities create inconvenience to maintain these identities but the user is generally willing to have his identities aggregated for more convenience in getting the desired services; (2) convergence of tier2 and tier3 identities: since the world is moving from mass marketing towards individual-specific marketing effort, there are chances that demographic groups related identity, tier3, will converge to tier2 identity [29]. Thus, identity fusion and aggregation takes a place.

Fig. 2.5 Identity views



The authors [52] address multiplicity of ‘views’ or ‘perceptions’ that can exist on subject’s identity in almost the same way to identity tiers. A single view defines a subject’s digital identity that has a context’s validity and appropriateness as shown in Fig. 2.5, which is adapted from [52]. There are three views of subject’s identity: ‘Me Me’ refers to the part of the identity information that the person is aware of and directly controls (e.g. residence address). ‘Known Me’ is the part of identity information that the person is aware of and indirectly controls (e.g. revenue data and the associated tax levels that are under the control of the department). ‘Unknown Me’ is the part of identity information that the person is not aware of and over which the person has no control. This information can be controlled by known parties (e.g. certification authority) or by unknown parties (e.g. credit rating agencies and identity thieves) [52]. We believe that this picture of identity that comprises multiple views, perspectives, or views is derived from a multi-dimensional classification of the human world, and the definition and role of identity in social sciences. It is said that: “identity is to know ‘who’s who’ (and hence ‘what’s what’). This involves knowing who we are, knowing who others are, them knowing who we are, us knowing who they think we are, and so on: a multi-dimensional classification of the human world and our places in it, as individuals and as members of collectivities” [53].

We believe that multiple YOUTs constitute the identity, or overall identity, of the subject. We borrow the words of Amin Maalouf, who grew up in Lebanon and now lives in France. He is the author of the book: “In the Name of Identity: Violence and the Need to Belong”. He shares his perspective and answers the question about identity; is he considering himself half French and half Lebanese? He says “not at all! The identity cannot be compartmentalized; it cannot be split

in halves or thirds, nor have any clearly defined set of boundaries. I do not have several identities; I only have one, made of all the elements that have shaped its unique proportions” [24].

### 2.6.2 *Origins of Fragmented Identity*

Digital identity is bringing a whole new dimension to our existing identities. We leave increasingly digital footprints in cyberspace forming a web of trails. Examples are digital records of our prenatal scans available on Flickr,<sup>6</sup> personal profile within a social networks, death information in FamilySearch<sup>7</sup> historical records, data collected by diverse agencies on our behalf, blogs’ contributions, emails, performed searches with various engines. Visible or invisible, left consciously or not, the data aggregation contributes to the definition of our identity. Editing our personal profile within social networks is different from that carried out by an employer ‘googling’ of a prospective employee, tracking our activities as a citizen, and possibly inferring health problems from our undertaken activities in self-advocacy groups [50].

Friends or other people opinions about an individual are highly affecting his digital identity. For instance, social networks users can tag friends through free online tagging services such as TagMyPals.<sup>8</sup> Such service offers a set of predefined digital representations or avatars based on classification of people personalities. TagMyPals users can tag friends full names on the avatars based on their perception of others’ personalities. The avatars and tags can be easily added to photos section in Facebook and Myspace. Above, in Fig. 2.6, few TagMyPals avatars. Distributed fragmented identity attributes is a consequence of a context-based nature of identity concept. In the real-life, various forms of identity are required to various contexts in which, the identity is to be presented in a suitable way and within suitable information by the identity holder. For instance, the person A, as a traveler, is asked to provide a passport at the counter of customs or immigration to proof his identity; the person A, being a car driver, is asked to show his driving license to a police officer, who stopped him in the highway; the person A, as a customer, is asked to provide his credit card with fidelity saving card in a movie store to take advantage of DVD prices reductions; the person A, as a student, is requested to show his student card to have access to computer lab facilities; the person A, as a patient, is asked his medical card in the hospital to receive health services [54]. The online world encapsulates a growing amount of scattered and unordered fragments of individuals’ identities due to two major reasons [42]. The first is because of the lack of a robust generic identification system and the second

---

<sup>6</sup> <http://www.flickr.com/>

<sup>7</sup> <http://fsbeta.familysearch.org/>

<sup>8</sup> <http://www.tagmypals.com>





**Fig. 2.6** Free tagging service according to friend's personality's classification

is the intentional creation of users' alternate identities. Creating more than one identity can be desirable for individuals depending on the context. A user may wish to be aggressive and egotistical in online multiplayer war game, but sensitive and sociable for virtual encounters and social networks [42].

Different enterprise directories store different pieces of identities. Modern organizations become distributed and maintain multiple identity repositories. This reality promotes spreading identity attributes across information systems and landscaping identity silos. Thus, different pieces attributes of our identity are contained in different environments such as files, enterprise directories, databases, and online social networks. We illustrate identity silos shaping and origins with the following use cases: (1) managing finance and preserving privacy. Rather than using a single credit card for shopping, most of the people prefer to use multiple credit cards to better manage finances and assure anonymity. A man buys a birthday's gift for his spouse with one of his credit cards rather than using the jointly held credit account. Therefore, each credit card issuer maintains a different set of user attributes; (2) managing attributes schema and policies restrictions. The restriction occurs when a number of identity stores do not allow write permission for several reasons, such as technical, governance and political reasons. In addition, the directory schema could be static and cannot be changed without major repercussions on the whole infrastructure. Hence, attributes would be stored only in a limited number of repositories and could not be distributed over all identity stores. We can extend this use case to point out that having identity attributes within different semantics, such as languages and cultural considerations could foster the identity fragmentation; (3) context-based nature of identity and governance issue. Each context requires a specific form of attributes to authenticate an identity holder; (4) technological advent and emergence. The identity management and access control related technologies have evolved within different computing waves that range from mainframes, mid-size systems to personal computing, and from enterprise distributed network infrastructure to the internet and web. The history of computing shows that new fragmented identities are created with the emergence of each discipline; (5) business dynamics. As a consequence of corporate mergers and acquisitions over time is a complex fragmented identity infrastructure; (6) Simple authentication and access management. Often, different lines of business or divisions maintain separate identity repositories in order to easily manage users' access to different and heterogenous business applications such as CRM

and HR; (7) multiple Web subscription. Many web sites require user subscription before providing services. As a result, a growing array of online fragmented identities is maintained by the Web sites' back-ends [54]. Concurrently, in information systems, access control and policies are different within different applications. Each application or service provider requires a specific set of attributes to let the user access the assets. A person may hold multiple credit cards issued by multiple banks that results multiple set of client attributes distributed over multiple repositories and locations. Furthermore, each individual has a couple of static attributes such as date and place of birth and dynamic attributes that may change such as blood pressure, home address, and phone number. Thereby, each person would have multiple sets of different attributes within different environments [49, 54].

### ***2.6.3 Digital Identity and Digital Memories***

All of the person's communications with other people and machines, as well as the images he sees, the sounds he hears, the Web sites he visits, and the Web searches he performs are recorded. US president Barack Obama provided some counsel for youngsters who want to grow up and be president. He replied to a 9th grader at Wakefield High School in the Washington suburb of Arlington, Virginia, who asked how he too could become President one day, saying that: "When you're young, you know, you make mistakes and you do some stupid stuff (...) I want everybody here to be careful about what you post on Facebook, because in the YouTube age whatever you do, it will be pulled up again later somewhere in your life" [55]. In the Digital Life article of the Scientific American Magazine [56], Gordon Bell and Jim Gemmell state that human memory can be maddeningly elusive and the era of digital memories is inevitable. Recently, a team at Microsoft Research Labs has developed a system, called MyLifeBits, to mainly digitally chronicle every aspect of a person life and to provide some of the tools needed to compile a lifelong digital archive. When the person is on the go, the system continually uploads his location from a portable Global Positioning System device. All of these recording are transmitted and stored in a personal digital archive that is both searchable and secure. After 6 years, more than 300,000 records, taking up about 150 GB are amassed. Portable sensors can take readings of things that are not even perceived by humans, such as oxygen levels in the blood or the amount of carbon dioxide in the air. Sensors can also log the three billion or so heartbeats in a person's lifetime. The authors explain of the new systems services by saying: "New systems may allow people to record everything they see and hear—and even things they cannot sense—and to store all these data in a personal digital archive" [56]. The same authors questioned why recording someone's life becomes possible today than before. The author cites three main reasons: (1) the growth of digital storage capacity has been staggering. Today a terabyte (one trillion bytes) hard drive can store everything the person read including emails, Web pages, papers and books, all the music the person purchased and downloaded, 8 h of speech and

10 pictures a day for the next 60 years. The author predicts that if current trends continue, in 20 years, with the same hard drive price, a person can buy a 250 TB of storage. This capacity should be able to satisfy anyone's recording needs for more than 100 years; (2) some of these devices can record a wealth of information about the users; (3) the dramatic increase in computing power has led to the introduction of processors that can efficiently retrieve, analyze and visualize vast amounts of information. Metadata such as the date, place and subject of a photograph or written or spoken comments that the database appends to the file, are easing the retrieve, or recall, process of digital memories. However, the advent of the digital-memories era will not be trouble-free. Many countries currently impose restrictions on recording conversations or photographing people. Moreover, many individuals are equally concerned about recording information for three reasons: (1) information could be used against them in court; (2) information could invade privacy; and (3) fear of access to records by identity thieves, gossipmongers or authoritarian states. In addition, from the security perspective, storing a lifetime of personal data in a single archive is vulnerable. One of the major advantages of digital memories is also mentioned. Digital memories allow vividly reliving an event with sounds and images, enhancing personal reflection in almost the same way that the Internet has aided scientific investigations. Every word one has ever read, whether in an e-mail, a document or on a Web site, can be found again with just a few keystrokes [56]. Emmpanuel Hoog, the CEO of INA, answers the questions of *Le Nouvel Observateur* reporters about the future of the world's digital memory and how to civilize Internet. He explains that years ago, individual or collective memory, is considered as a rare cultural asset and therefore valuable. A 100 years ago, a family life was illustrated by a dozen of pictures. But today we take hundreds of photos in summer holidays with small digital camera and mobiles. We are passing to future generations a huge stock of digital memory. In addition, museums, archives, universities, heritage institutions have long been in charge of sorting and organizing knowledge, but today, nobody can accept this because each digital producer manage by himself his memory with his manner. This would weaken our ability to draw a common destiny. He adds that given the ever growing content available on the Internet, the fundamental issue is how to sort, to make choices. The government has focused so far on the issue of digitization of content but now it should focus on how to make content accessible to more people. He thinks also that authorities should urgently address the issue of access criteria and the hierarchy of knowledge on the web at local and regional levels. Today, the monopoly of access is between the hands of search engines, which are using non transparent criteria for web content indexing. Such content is considered as an economic asset, he urges public authorities to create real spaces for public service, knowledge and expertise on the Internet. Hood calls continuity logic between souvenirs, memory, and history as 'memorial ecosystem'. He adds that the memorial ecosystem is called into a question with the advent of the digital world. Yesterday, there was some continuity between stages and each stage is the pre-cursor of the next one. Today we can remember everything, thus souvenirs and memory are taking precedence over history. And somehow, too much

memory kills the memory-or, rather, too much memory kills the history. The explosion of the memorial bubble may produce two consequences: (1) the resurgence of the wars of memory. Because history can be unfair, every minority can claim its history and identity at a large scale. The excess of such claims may generate identity crisis; (2) amnesia and collective cultural loss; (3) why indeed memorize, since the machine remembers us? Hoog writes “always more memory, but still less marks” to explain that the right to forget he is needed as a requirement for democracy. Today, there is a tremendous privatization of our personal data in the Internet. Companies are drawing profiles on personal information of each of us. Despite the efforts of the National Commission for Informatics and Liberties (CNIL) in France, the situation is not satisfactory. Every citizen is in danger of his past that can reappear at any moment. At the same time, we become producers of memory and we have accepted a regression of our privacy. However, privacy, rights to privacy is the foundation of a liberal society. Hoog adds that the digital native would have the challenge how to search on the internet. In the real world everyone can distinguish with the naked eye a grocery store, a school, a town hall, and a garage. For the Internet, it should be the same thing. I think that civilized Internet is allowing everyone to navigate easily. It is a challenge that calls for new forms of public regulation. Not everything can be left to the search engines that are now the only players in the web, which structure and organize it [57].

#### ***2.6.4 Digital Identity in Social Networks***

Social networking sites are gaining more and more importance on people daily life. They offer people ways to communicate and socialize with each other via the internet through a PC or mobile phone. Individual's friendship chain become part of digital identity. Would you be my online friend? Once the user finds a profile of a friend or someone else, he can add him by sending a message to the other user requesting friendship. If the recipient approves the connection, the relationship is visible through both users' list of friends. The friends' list typically includes a list of links to other friends' profiles. Thus, when participants surf on social network sites, they can jump from one profile to another through a friendship chain. Based on a research results published in the report [40], the average adult social networker has profiles on 1.6 sites, and most users check their profile at least every other day. Part of the digital identity is constructed through the web of trails that individuals are leaving in the online world, especially in social networks. In fact, thirteenth century Mulla's dilemma touches the central social problematic of identity construction [13] and, in the same way, the author of the book [7] explains that digital identity is bounded, not only to identity attributes, but to the individual's behavior. Thus, in a restricted manner, digital identity is bounded to individuals' behaviors in social networks. Trails could be customized profile information, opinion sharing about a subject or other friends, photos posting, and so on. The Ofcom report states also that users create well-developed profiles as the basis of their online presence and such

profiles often contain very detailed individual's information, even though it is not compulsory to provide that much of information [58]. People could easily and simply create their own online page or profile, and construct and display an online network of contacts, often called 'friends' [40]. As examples of well-known social networks: Friendster,<sup>9</sup> MySpace,<sup>10</sup> Facebook,<sup>11</sup> Bebo,<sup>12</sup> Skyrock Blog,<sup>13</sup> Hi5,<sup>14</sup> Orkut,<sup>15</sup> LiveJournal,<sup>16</sup> and CyWorld.<sup>17</sup> In order to join these networks, a user should register and create a social profile by entering a set of static and dynamic user attributes such as their demographics and tastes, a self-description, and often photos that provide a visual image. The participant's social profile is considered in this context as a social persona being a part of his digital identity. Some social networks sites allow participants to articulate and publicly display their relations to others in the system, which, in turn, allow viewers to traverse the network.

Online social profiles and activities is having more visibility and gaining more accessibility through "Universal Social Networks", abbreviated "USN". USN, called also social networking convergence service [58], is basically an application which focuses on making easier for end users to create content independently of the blogging platform usage. It allows updating all the blogs and web services from within one environment. USN permits to make it easy for the end user to let his friends and colleagues around the web know what he's up to and what he's writing. It keeps the end user friends on any network informed about his activities. But what are the consequences of USN usage on our digital identity? We present a list of USNs that are classified into four categories: (1) Social feed aggregator, called also lifestream, or online presence aggregators: MyMashable [59], Profilactic [60], Snag [61], Profileomat [62], Naymz [63], SocialURL [64], PeopleAggregator [65], ProfileFly [66], SocialNetwork.in [67], and Mashable [68]. These services are ready to exploit but others are still in status of work in progress such as ProfileLinker, Upscoop, MyLifeBrand, Tabber, Ex.plode.us, Correlate.us, Istalkr, and SocialStream [69]; (2) desktop aggregator, an application that provide a single access to many social networks and aggregation capabilities: 8hands [70, 71], NoseRub [72, 73], and Minggl [74, 75]; (3) people finder such as Wink [76], a people search over the user profiles of MySpace, LinkedIn and Bebo. Spokeo [77] is another example of people finder that offers a search, by name, email address, phone number and friends; and (4) users' bookmarks aggregator such as SecondBrain [78].

---

<sup>9</sup> <http://www.friendster.com>

<sup>10</sup> <http://www.myspace.com>

<sup>11</sup> <http://www.facebook.com>

<sup>12</sup> <http://www.bebo.com>

<sup>13</sup> <http://www.skyrock.com/blog>

<sup>14</sup> <http://www.hi5.com>

<sup>15</sup> <http://www.orkut.com>

<sup>16</sup> <http://www.livejournal.com>

<sup>17</sup> <http://us.cyworld.com>

### ***2.6.5 Digital Identity, Context-Awareness, and Ubiquity***

Establishing the identity of a person is becoming an important need in context-aware environments. Context awareness originated as a term from ubiquitous computing, called also pervasive computing deals with linking changes in the environment with computer systems such as RFID, GPS, ambient intelligence and other emergent context-aware applications [79]. In criminal cases, psychological profiling has given way to DNA matching. In consumer products, commodity logistics have given way to RFID databases. Genomics are the universal identification of life abstract; biometrics is considered as the universal identification of life in particular; collaborative filters are the universal identification of life in the relational [80]. Biometrics is specified as the science of recognizing an individual based on psychological or behavioral traits. Biometric systems, which rely on the evidence of fingerprints, hand geometry, iris, retina, face, hand vein, facial thermo-gram, signature, voice, etc., are deployed as a means of establishing and validating identity [81]. Privacy issues related to digital identity would inevitably rise as far as coincidence between happening and storage becomes more persistent in the future.

### ***2.6.6 Frauds, Misuse, Fake Profile and Crimes of Identity***

Identity fraud is a profitable enterprise. “Individuals have an asset called their identity”, said Dr Tom Ilube, CEO of a data security company. He adds: “it is valuable to you and valuable to those people that want to abuse it” [82]. Fraud is rising rapidly because people are posting personal facts on the Web as well as government agencies are steadily making databases available online. These databases include birth, marriage and death certificates, credit histories, voter registrations and property deeds [83]. Security, identity theft, incorrect computer records, credit rating destruction, privacy, online purchasing and banking, loss of identity, misuse of personal information, phishing, identity cards, behavioral monitoring and tracking, the list of concerns goes on and on [84]. The Liberty Identity Theft Task Group, defined the three stages of identity theft as: (1) stealing identity data: while the numbers and stories about identity data loss are sensational, companies that suffer this traditionally only faced embarrassment and a bruised reputation; (2) hijacking existing accounts: 80 % of phishing attacks are against financial services<sup>18</sup>; and (3) concocting new accounts: the fraudulent opening of new accounts using another’s identity is more dangerous because valid credentials are given to the criminal. When identity credentials are given to the wrong person, the strength of identity technology is powerless to help [85]. In addition, identity theft and fraud rate is increased due to risks posed by data deluge. Data deluge poses risks such as disks full of social-security data go missing; laptops loaded with tax records left in taxis; credit-card

---

<sup>18</sup> Anti-phishing working group: <http://www.antiphishing.org>.

numbers are stolen from online retailers; and Big Brotherishness of customers' personal information. The consequence is privacy breaches, identity theft and fraud [85]. More than 800 million active users in Facebook, around half of them currently access Facebook through their mobile devices [86]. As far as social networks are attracting more people, digital identity within these networks becomes more fragile and easily fall prey to social engineering traps. The case of Robin Sage experiment [87, 88] illustrate how fragile is digital identity in social networks and how it is easy to create a fake online profile that refers to nonexistent offline person. The Robin Sage experiment was conducted by Thomas Ryan. He created blatantly false identity of a woman claiming to work for in military intelligence and then enrolling on various social networking websites. Ryan deliberately chose an attractive young female's picture to prove that appearance is crucial in trust and people's eagerness to connect with. After a month, Robin has accumulated connections to around 300 online social networks. Contacts included an array of executives at government entities, employees of global 500 corporations and throughout the experiment Robin was offered gifts, government and corporate jobs, and opportunities to speak at various security conferences. Ryan tried to highlight how easily trust is given in these spaces and how much different information gets leaked out through various networks. He recommends social network users to accept only contacts that they know or make a research on people before accepting contacts' requests. See more cyber-criminality for black-markets report [89].

### ***2.6.7 Digital Identity Aggregation Drivers and Issues***

We ascribe “Out of Many, One” from “E Pluribus Unum” [90], which is used in the Great Seal of the United States [91], to underline the idea behind the scene of digital identity aggregation and fusion. Profiles are either unified into one all-encompassing digital dossier or relationships are defined among them to form a single digital identity. Moreover, we use the expression to point out high and urgent societies' expectations and needs for digital identity fusion capabilities that help investigators to identify a terrorist blended in with many people.

Data fusion can drive organizations to make better use of the data they own and provide convenience by creating an information resource that is more powerful, more flexible and more accurate than any of the original data sources [92]. Early in the mid-1800s, Matthew Fontaine Maury of the American navy had the idea of aggregating nautical logs from ships crossing the Pacific to find the routes that offered the best winds and currents. He created an early variant of a “viral” social network, rewarding captains who submitted their logbooks with a copy of his maps. But the process was at that time very slow and laborious [93]. Las Vegas casinos have been pioneers in fusing data from various sources because they face so many schemes to rip them off. Watching Hollywood films such as *Enemy of the State* and the *Jason Bourne* trilogy shows that shadowy organizations have instant and easy access to all the databases for various security purpose, particularly to



identify terrorists. DARPA researchers argued that the World Trade Center bombing of 1993 and the Oklahoma City bombing of 1995 might have been prevented if US public security services could have linked commercial databases to identify large purchases of fertilizer by non-farmers [92]. In addition, the author of the Economist ‘Data Deluge’ article [94] explains the current situation of digital identity aggregation and fusion by pointing out that despite years of large-scale efforts, law-enforcement and intelligence agencies’ databases are still not effectively linked yet. He gives the examples of health care industry in which computerizing health records tend to run into bureaucratic, technical and ethical problems. The digitization of health records could have been helpful to spot and monitor health trends and evaluate the effectiveness of different treatments. We point out that features and tools offered by Naymz [63] such identity aggregator, reputation assessment tool, and reputation score ‘RepScore’ could inevitably help to build trust-based professional community. After 9/11, the American Defense Department launched a program called “Total Information Awareness” to compile as many data as possible: e-mails, phone calls, web searches, shopping transactions, bank records, medical files, travel history and much more. In his article [92] titled “Information of the World, UNITE!” published in Scientific American Magazine, Simson L. Garfinkel explains through a hands-on, real-life experience motivations of digital identity aggregation or fusion. He says: “A few years ago I bought a latte at Starbucks on the way to the airport, parked my car and got on a flight for the U.K. 8 h later I got off at Heathrow, bought a prepay chip for my cell phone and went to buy a ticket for the train into London, when my credit card gave up the ghost and refused to work anymore. Not until I got back to the U.S. did I find out what had happened. Apparently, the small purchase at Starbucks, followed by the overseas purchase of the cell phone card, had tripped some kind of antifraud data-mining algorithm in my credit-card company’s computer. It tried to call me, got my voice mail and proceeded to blacklist my credit card. What I found so exasperating about the entire experience was that the computer should have known that the person using my card in England was me. After all, I had bought my plane ticket with that same card and had flown with a major U.S. carrier. Aren’t all those databases supposed to be tied together?” [92]. In the next sections, we explain that mashing digital identity attributes, from credit-card bills to cell phone logs, poses technical, economic, legal and ethical problems. Below, motivations for security purpose are listed and explained.

### ***2.6.8 Digital Identity Aggregation for Security Use Cases***

A digital identity silos consolidation is considered as one of the current challenges and a critical step to secure access to information systems’ assets [27]. Digital identity aggregation, synonym of ‘digital identity silos consolidation’, establishes relationship between distributed attributes. We use the term ‘silos’ to convey that digital identity attributes are rarely stored in one place but rather in diverse and various stores residing within multiple information systems. As a consequence, the individual

is in one-to-many relationship with his identity. Merriam's dictionary defines 'to consolidate' in the meaning of to strengthen and to unite. Several use cases explain and illustrate the need of digital identity aggregation for security purpose. We detail three of the use cases: (1) applications and services may require more attributes to authorize the user accessing resources. This is reflected in the real world as a person, who is asked to provide more than one identity proof comprising different identity information to get a customized service. For instance, a customer is asked to provide a credit card and fidelity saving card in a movie store to take advantage of DVD prices rebates. Moreover, to get into some mistrusted or restrictive environments, such as national security organizations, a visitor is asked to provide more than one identity card; (2) provisioning an employee who leaves. Consolidating employee identity attributes across information systems and synchronizing them would allow recognizing the validity of his authentication performed inside and outside the information system; (3) online reputation systems are in use to trust parties and conduct secure online business. For instance, eBay reputation mechanism unifies member's transaction feedback history to calculate community members' reputations in the form of colored and shooting stars. In addition, we need not only just a consolidation but an effective attributes because a poor administration and maintenance of duplicated, out-of-date, and low-quality identity attributes may expose enterprise assets and resources at a high risk. From the subject and service provider perspectives, digital identity aggregation becomes a highly used tool to reduce identity theft. Currently, services providers are using advanced tactics, collectively known as identity scoring that allows monitoring online data mining, pattern recognition, even semantic analysis of information about a subscriber that appears on Web pages. Examples of firms that offer such services are Garlik [95] in England and MyPublicInfo [96] in U.S. Garlik offer 'data patrol' service to British residents by combing credit reports, public databases and Web sites for information about customers and presents them with a detailed profile. The profile should show whether criminals may be trying to use their personal facts to apply for credit cards, take out a loan, or register a fake driver's license or marriage certificate. MyPublicInfo pieces together a customer's 'public identity profile' for \$79.95 and alert him or her to dubious changes for \$4.95 a month [83]. Moreover, the subject must be able to combine selected claims made about himself by more than one identity authority into a minimal composite set of claims and be able to present them to relaying party, who could not be able to repudiate the original claims [59].

Many participants have different profiles within multiple social networks. From the user perspective, aggregating profiles would (1) increase convenience of the social experience: the participant can post a message to multiple friends within different social networks; (2) ease access control (identification, authentication, authorization, and accountability); and (3) attributes management. From the organizational perspective, social profiles aggregation would ease (1) participant's reputation management: HR department might aggregate a candidate's social profiles in order to decide whether to hire him or to reject his application. Another example is a student, who wants to know more about his professor, would make a Google search and professor's social profiles aggregation; and

(2) service personalization (profiling): in order to increase market shares, companies might aggregate client's social profiles to know more about their preferences and goals, as a consequence, they can personalize products and services. They might also consider the friends list of a client or business partner as prospect clients.

### ***2.6.9 Economy of Digital Identity Aggregation: Digital Gold Mine***

Today, organizations strive to capture and aggregate digital identities because they are convinced that is the new form of 'rué-vers-l'or'. Such agitation is comparable to the one that is used to be with hundreds of people when searching for gold, panning in the streams and digging mines. 'Gold Rush' (1925), the Charlie Chaplin's movie, is a true illustration of major gold rushes that took place in the nineteenth century in Australia, Brazil, Canada, South Africa, and the United States [76, 97, 98].

Digital identity attributes become publicly available and easy to access. Each person now leaves in cyberspace an increasingly amount of digital footprint when aggregated and unified, contributes to the definition of the subject's digital identity. Visible or invisible, left consciously or not, this set of data can be collected from various sources. The very first digital records of pre-natal scans could be shared on flicker and the obituary information on the Social Security Death Index (SSDI),<sup>19</sup> Find a Grave,<sup>20</sup> and Interment.net.<sup>21</sup> It happens also that other data could be available and collected through the one collected by diverse agencies and organizations on our behalf during our life, the blogs that are kept, the emails sent and the internet searches performed [66–68]. Maintaining and editing personal information in learning digital portfolio or personal profile within social network is much feasible and easier than the personal profile that is carried out kept by an employer 'googling' prospective employee, tracking activities as a citizen, and possibly inferring health problems from the visible activities in self-advocacy online groups. For instance, We Feel Fine [99], Fig. 2.7, is a people feeling aggregation engine that harvests automatically human feelings from a large number of blogs every 10 min. Compiled blog data [100] comes from a variety of online sources, including LiveJournal, MSN Spaces, MySpace, Blogger, Flickr, Technorati, Feedster, Ice Rocket, and Google. The engine scans blog posts for occurrences of the text fragments 'I feel' and 'I am feeling'. The approach was inspired by techniques used in Listening Post project [101].

The value of digital identity increases as much as substantial quantity of digital identity attributes has been collected and aggregated. Many people search engines are evolving to better provide services by aggregating people digital identity

<sup>19</sup> <http://ssdi.rootsweb.ancestry.com/>

<sup>20</sup> <http://www.findagrave.com/>

<sup>21</sup> <http://www.interment.net/>

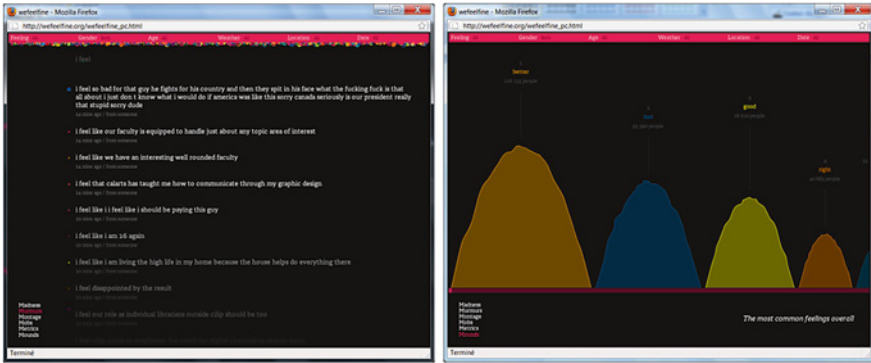


Fig. 2.7 People feelings murmurs and emotions aggregator

attributes. 123People<sup>22</sup> engine provides information on people to learn more about a person, an acquaintance, a colleague, a potential collaborator. Having its roots in Austria, 123people aggregates information of the digital identity of a person on the Web taken from multiple sources such as Web pages, social networks, images, videos, blogs, micro-blogging platforms, and emails [102]. Others such as Spock<sup>23</sup> and USsearch<sup>24</sup> are providers of people search and background checks that work jointly to provide free aggregated digital identity information and paid service to access on sensitive and premium information such as criminal record. Another service that provides obituary information is SSDI Index. The person enters the first and last name, then the SSDI Index resource turns up full name, birth and death dates, last known residence, last benefit, social security number, and state in which the social security card was issued. Other record-related information is available upon order. As an example, we use the Social Security Death Index (SSDI) service provider to look for ‘Abraham Lincoln’ personal information in US public registries. The result is presented in the following screenshot, Fig. 2.8.

Another example of public records aggregator and people finder is Intelius.<sup>25</sup> The system reports genealogy records that comprise phone numbers, address history, birth certificates, death records, marriage licenses and divorce decree. It allows tracing family tree by saving, adding, and joining records together. Moreover, the system provides neighborhood and property information such as home value, sales history, property details and ownership information. In Fig. 2.9, Intelius shows Ghazi Ben Ayed’s public record as he was a resident of Milwaukee, WI from 1998 to 2000. It makes public personal data such as his mother’s full name in the relative column: ‘Zahra Ben Ayed’. When the user heats the View

<sup>22</sup> [www.123people.com](http://www.123people.com)

<sup>23</sup> [www.spock.com](http://www.spock.com)

<sup>24</sup> [www.ussearch.com](http://www.ussearch.com)


<sup>25</sup> <http://www.intelius.com/>

Viewing 1-20 of 21

Name	Birth	Death	Last Residence	Last Benefit	SSN	Issued	Tools	Order Record?
ABRAHAM LINCOLN	17 Sep 1887	Dec 1966	32713 (Debary, <a href="#">Yoknis, FL</a> )	(none specified)	001-12-3083	New Hampshire	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM LINCOLN	22 May 1895	Feb 1970	15108 (Coraopolis, <a href="#">Allegheny, PA</a> )	(none specified)	069-03-6847	New York	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM LINCOLN	22 Jun 1911	Aug 1983	11758 (Massapequa, <a href="#">Nassau, NY</a> )	(none specified)	089-12-8975	New York	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM LINCOLN	23 Jan 1907	Nov 1974	12941 (Jay, <a href="#">Essex, NY</a> )	(none specified)	116-01-1374	New York	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM LINCOLN	15 Nov 1926	May 1979	(not specified)	14729 (East Otto, <a href="#">Cattaraugus, NY</a> )	117-14-3679	New York	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM LINCOLN	24 Nov 1889	15 Nov 1966 (V)	19146 (Philadelphia, <a href="#">Philadelphia, PA</a> )	(none specified)	172-24-6105	Pennsylvania	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM LINCOLN	13 Nov 1900	04 Aug 1990 (V)	17404 (York, <a href="#">York, PA</a> )	(none specified)	178-16-3226	Pennsylvania	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM S LINCOLN	27 Jan 1919	15 Mar 2005 (V)	79925 (El Paso, <a href="#">El Paso, TX</a> )	(none specified)	225-01-5120	Virginia	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a> <a href="#">Search Ancestry.com</a>	<input checked="" type="checkbox"/>
ABRAHAM ELIJAH LINCOLN	08 Apr 1918	26 Aug 2007 (P)	29212 (Columbia, <a href="#">Lexington, SC</a> )	(none specified)	250-09-9182	South Carolina	<a href="#">SS-5 Letter</a> <a href="#">Add Post-em</a>	<input checked="" type="checkbox"/>

Terminé

Fig. 2.8 Abraham Lincoln obituary information in US public records



Sign In – My Intelius  
View My Reports

<< Return to Home

PEOPLE SEARCH RESULTS

Search Again >>

We found 1 person that matches **Ghazi Ben ayed** in the **United States**.  
Click on the **View Details** or **Get a Detailed Report** link for more info.

☒ = Available

See Details on All 1 People!

Expanded Search Results

■ We searched Ghazi Benayed and found 1 records nationwide

Name	Age	Previous Cities	DOB	Phone	Address	Avg. Income	Avg. Home Value	Relatives
<b>Ghazi B Benayed</b>	30	Milwaukee, WI		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Dhazi Benayed Zahra Benayed

View Details

People Search

First Name

MI

Last Name

State

Ghazi

Ben ayed

All States

Search

Advanced Search

What is a People Search?

People Search is great way to find and reconnect with family, old friends, relatives — just about anyone! People Search reports include phone numbers, address history, ages, birthdates, household members, home value, income and more.

Fig. 2.9 Result of ‘Ghazi Ben Ayed’ searching in US public record



Fig. 2.10 Elvis Presley archives available through google news search archives

Details red button requesting to edit the public record-related data located at Wisconsin authorities, the system asks to order and pay, at special or regular prices.

In addition, users of Google news archive search<sup>26</sup> can explore historical archives about events, people or ideas and see how they have been described over time. In addition, users can also see a historical overview of the results by browsing an automatically generated timeline. Search results include content from a number of sources, through content digitized by Google and online archival materials that Google crawled. Search results can include content that is freely accessible as well as content that requires a fee. Articles related to a single story or person within a given time period are grouped together to allow users to see a broad perspective on the topics they are searching [103]. Figure 2.10 shows publicly published Elvis Presley information.

### 2.6.10 Technical Issues of Digital Identity Aggregation

In 2008, the author [92] explains that digital identity fusion is hard because we are drowning in data from a multitude of sources, all with different levels of detail and uncertainty. John Marlan Poindexter, a career naval officer, says that identifying the signatures of terrorist preparations in an ocean of data is much harder

<sup>26</sup> <http://news.google.com/archivesearch>

than finding subs in an ocean of water. In addition, Poindexter argues that oceans may be huge but every spot can be uniquely identified by a latitude, longitude and depth. However, data oceans are not so easily to be categorized. Much of information are spread across millions of individual computer systems and hidden to the authorities. In addition, oceans are not doubling in size every few years like data oceans. Major issues are: (1) data quality. Much of the personal data in databases may not be accurate and they are riddled with errors and meaningless coincidences. A Scientific American editor ordered an US \$80 report from an online consolidator of digital identity, including criminal, real-estate and bankruptcy records. It was riddled with errors such as misspellings and confusion with namesakes. The report showed no signs of identity theft! Currently, new algorithms overcome only some of these hurdles but not all of them; (2) making sense (semantics) of data fusion. Users are sometimes unaware of the digital bread crumbs they leave but companies are increasingly linking isolated databases together into one data scheme could infect a person's entire digital identity and reputation either by stealing data scheme or through attributes aggregation bias, particularly decontextualization of digital identity by data mining algorithms. Yet another problem for data fusers is; (3) identity resolution, which is matching up the various names and account numbers with the right individual by taking into account cultural variation in names and other business-related rules [92, 104]. In online world there may be dozens of people sharing the same name and dozens of names used by the same person, thus the issue deals with ontology and syntax of digital identity attributes. Person's first name may be listed in one database as Robert, in another as Rob and in a third as Bob. A person whose Arabic name is Haj Imhemed Otmame Abderaqaib in West Africa might be known as Hajj Mohamed Uthman Abd Al Ragib in Iraq. Casinos have funded development of a technique called NOonobvious Relationship Analysis (NORA), which combines identity aggregation and resolution with databases of credit companies, public records and hotel stays [92].

Figure 2.11 sums up digital identity aggregation technical issues and illustrates that attributes semantics, ontology, syntax and interoperability issues arise whenever and authority needs to aggregate a multiple digital identity attributes in order to decide whether to provide a service to the subject. For example, how computers could recognize that the short names 'G. Ben Ayed' and 'Ghazi B. Ayed' are referring to the same person with a full name 'Ghazi Ben Ayed'? In addition, names written with typo errors such as 'Gazi Benayed' and 'Ghasi Bennayed', the ones written in other languages and following cultural semantics such as Hispanic, Japanese, Chinese and Arabic, or Arab names written with Latin font could be automatically recognized as being part of the same person's identity? The authors [31] explain that identity management service must support vocabulary definitions of identity attributes. A fundamental assumption is that all parties concerned with identity services share a common ontology and semantic web metadata formats such as Resource Description Format (RDF) and RDF Schema.



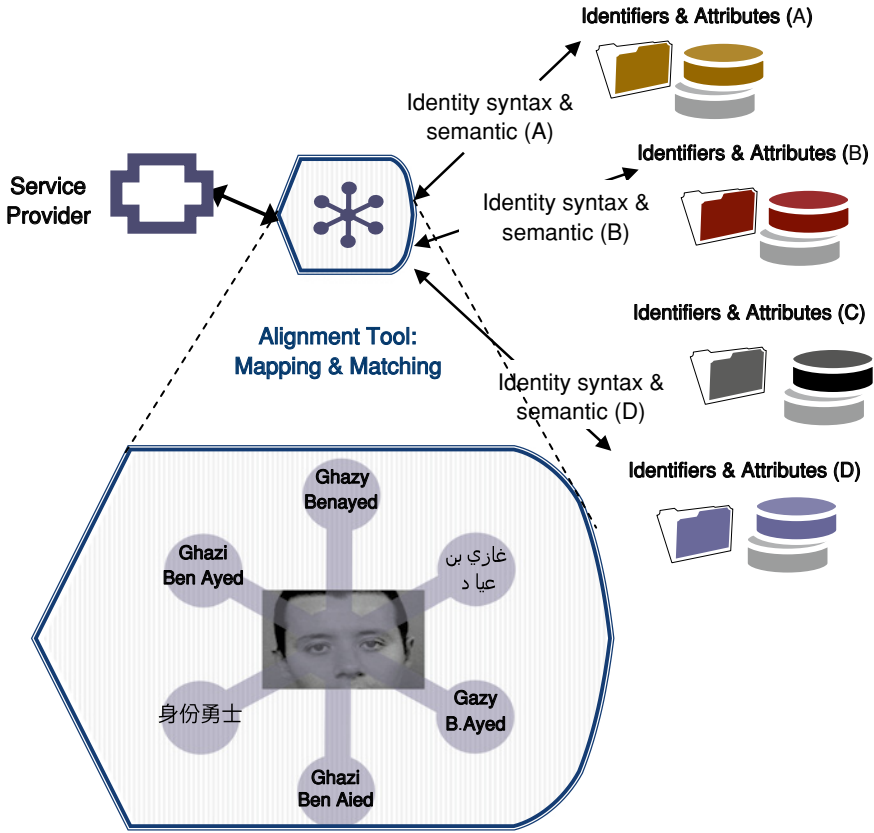


Fig. 2.11 Digital identity aggregation technical challenges

### 2.6.11 Digital Identity Aggregation Systems and Algorithms

Algorithms of data fusion can trace its heritage back to the computerized matching programs of 1970s. US government authorized the creation of the Federal Parent Locator Service that denies a wide range of federal benefits to parents who are behind on their child support. Those data are fused with digital identity of recently employed parents who are not up to date on their payments so that their wages can be garnished [92]. After 9/11, the American Defense Department launched a program called “Total Information Awareness” to compile as many data as possible: e-mails, phone calls, web searches, shopping transactions, bank records, medical files, travel history and much more [105].

The program works by building hypotheses based on existing profiles and then revising these hypotheses as other digital identity attributes become available. In the 1990s software engineer Jeff Jonas developed a system that could match the names in a casino’s computers with other sources of information. Figure 2.12, which is adapted

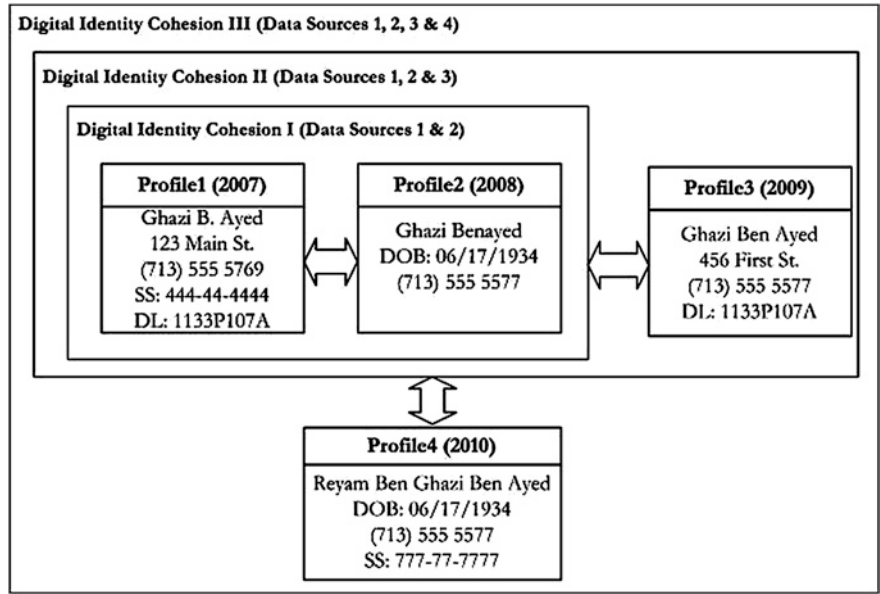


Fig. 2.12 Casino’s digital identity fusion algorithm developed by Jeff Jonas in 1990s

from [92], shows that four of the profiles reside in different locations and have been collected in different periods of time. Digital identity aggregation I combines profile1 and profile2 and each of them holds different attributes, so the system provisionally assumes they represent different individuals. In aggregation II, the system infers that profile3 holds attributes common to both previous profiles: the driver’s license number from one and phone number from the other. So the system reassigns all three to the same individual. Finally, digital identity aggregation III shows that profile4 includes a birth date matching with profile2, thus, the system deduces that the four profiles actually represent two individuals. The program guesses that the two may be father and son since they share the same surname and phone number. In 2005, Jonas sold the system and his company to IBM, which has added a feature called anonymous resolution. Two organizations can determine whether they share the digital identity of an individual in their databases without revealing digital identities of all people who do not match. The technique works by comparing cryptographic hashes instead of digital identity attributes. Currently, most algorithms of data fusion have some kind of sensitivity adjustment. Tipping the scale to the right, and the system fails to find genuine matches; tipping it to the left, the system turns out to be wrong because too many predictions are achieved. Another important issue raised by data aggregation is to find an algorithm that it never confuses original data with a conclusion inferred from those data [92].

Economic gains should justify fusion costs. In 1994, Roger Clarke of the Australian National University in Canberra studied computerized matching programs maintained by federal and state governments in the U.S. and Australia. These systems

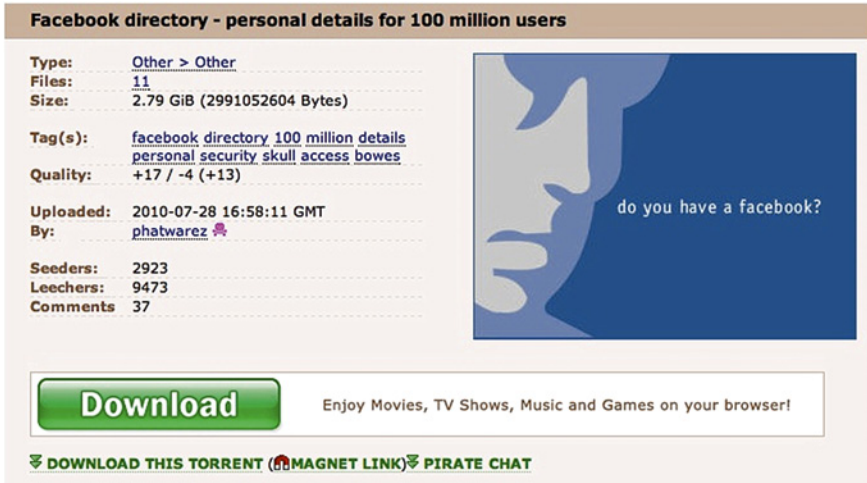


Fig. 2.13 A ready-to-download file comprising details of 100 million Facebook users

scanned millions of records and flagged thousands of potential “hits.” But most of them turned out to be false positives. The benefits did not justify the costs of collecting data, training personnel and chasing down the false positives. The same author argues that many people feel that if a data-fusion program could anticipate and stop a major terrorist attack, it would be worth whatever it cost [92]. However, from the ethical and legal perspectives, linking together databases into a single profile through the process of data fusion is still the *bête noire* of privacy advocates. They advocates still considering that identity data aggregators use personal information for purposes other than the ones for which it was originally acquired [92]. The author of *How To Be Invisible* book [89] states: “Do not, as long as you live, ever again allow your real name to be coupled with your home address”. This is to point out that preserving privacy is a matter of conscience. Privacy issues are detailed in Chap. 4.

Many use cases illustrate the dangers of maintaining digital identities at a poor security level. Almost 3 GB file that contains 100 million Facebook users has been made available on a torrent site downloadable by absolutely anybody in July 2010 see Fig. 2.13. Ron Bowes of Skull Security created a script [106] that harvested user information from Facebook’s user directory [107].

Ron’s idea was to spider and generate first-initial-last-name list and once he had the name and URL of a user, he aggregated users’ pictures, friends, and information about them, with some other details. He wrote a Ruby script to download the full Facebook users’ directory and link personal details to the corresponding first, last, and usernames. The results were 171 million names (100 million unique) [107].

The file, Fig. 2.13, contains the URL of every searchable Facebook user’s profile, the name of every searchable Facebook user, both unique and by count, and processed lists, including first names with count, last names with count, and potential

**Table 2.1** Top Facebook usernames’ lists

A first initial and last name-based list	A first name and last initial-based list	A first name dot last name-based list	A first name-based list	A last name-based list
129369 jsmith	100225 johns	17204 john.smith	977014 michael	913465 smith
79365 ssmith	97676 johnm	7440 david.smith	963693 john	571819 johnson
77713 skhan	97310 michaelm	7200 michael.smith	924816 david	512312 jones
75561 msmith	93386 michaels	6784 chris.smith	819879 chris	503266 williams
74575 skumar	88978 davids	6371 mike.smith	640957 mike	471390 brown
72467 csmith	85481 michaelb	6149 arun.kumar	602088 james	386764 lee
71791 asmith	84824 davidm	5980 james.smith	584438 mark	360010 khan
67786 jjohnson	82677 davidb	5939 amit.kumar	515686 jason	355639 singh
66693 dsmith	81500 johnb	5926 imran.khan	503658 robert	343220 kumar
66431 akhan	77800 michaelc	5861 jason.smith	484403 jessica	324972 miller

usernames with count, as presented in Table 2.1 [107, 108]. Even if the user opts out of inclusion in the search, he could still appear on the directory page of a searchable friend. The statistical lists don’t pose any security threat to Facebook users; however, data could be useful for building automated account cracking software. Lists of the most common names can be used to assemble a good dictionary of potentially popular usernames for use in tools that attempt to identify and crack user accounts [109].

**2.6.12 Digital Native’s Perception of Identity**

What is the impact of digital identity and privacy on the “digital native”? The term “digital native” means the generation that grew up with Internet and new information technologies. “Digital Natives” were born after 1980, when social digital technologies came online. They carry mobile devices all times not just to make phone calls but also to send text messages, surf the Internet, and download music. They’ve been living with mobility, speed access to information, learning with media, participatory action, co-creation of value, etc. [110–112] In the shadow of the daily growth of the global population in general and particularly the digital native one [111, 113], digital identity would play major role in the next few years. Digital Natives live much of their lives comfortably online, without thinking of their digital identity and their real-space identity as separate things. They just have an identity, which is a representation in two, or three, or more different spaces. Digital Natives are constantly connected. Even as they sleep, connections are made online, in the background; they wake up to find them each day. They connect to social networks, IM, and share photos with friends all over the world. Digital Natives are creating parallel worlds on sites like Second Life. And after they do, they record parts of that world and post a video of it on YouTube or Daily Motion

in a ‘machinima’ art form. Digital Natives perceive information as something they can control and reshape in new and interesting ways. They edit a profile on MySpace or encyclopedia entries on Wikipedia, make a movie or online video, or download a hot music track—whether lawfully or not. Digital Natives can rework media, using off-the-shelf computer programs. Research means Google search and particularly Wikipedia before diving deeper into a topic. Most Digital Natives don’t buy newspapers ever but they get it in new ways and in a wide variety of formats. In the process of spending so much time in this digitally connected environment, Digital Natives are leaving more traces of themselves in public places online. With every hour they log online, they are leaving more tracks for marketers—and pedophiles, for that matter—to follow. Digital Natives’ ideas about privacy, for instance, are different from those of their parents and grandparents but how? The repercussions of these changes in the near future will be profound for all of us. The Digital Natives has global culture in scope and nature whether physically based in different cities, countries and continents [111].

### ***2.6.13 Issues and Concerns Associated with Handling the Digital Afterlife***

In his article preparing for the digital afterlife [114], Duncan Jefferies questions how should we handle digital legacy? How should we deal with online accounts such as Facebook and PayPal logs off for good? It might depend on the law, but by default digital assets are “the property of the estate, even if they’re property with no value”. Some assets, such as blogs and photographs, may also be subject to intellectual property law. “People aren’t very aware of what you might call their living online legacy—potential employers looking at their Facebook accounts, for example. The issue of what happens to that information after their death is an extension of that” says Yorick Wilks, a senior research fellow at the Oxford Internet Institute. Facebook puts the profile of deceased person into a memorial state upon notification of their death. Their status is removed, they are withdrawn from any groups and access is set to “friends only”. Couldn’t his descent being part of his social circle of friends? Donna Rawling lost her husband and she says: “I managed to wrap up his affairs, but the area that I was left with was his presence on the web”. Several companies aim to help people to better handle digital legacy by providing Digital deposit accounts playing the role of “electronic safe deposit box”, where people can easily upload login details for digital assets and specify who will receive them posthumously. Examples are LegacyLocker,<sup>27</sup> SlightlyMorbid<sup>28</sup> and Deathswitch.<sup>29</sup> Deathswitch provides an automatically

---

<sup>27</sup> <http://legacylocker.com>

<sup>28</sup> <https://www.slightlymorbid.com>

<sup>29</sup> <http://www.deathswitch.com>

prompts people for their password on a regular basis. If nothing is received after several prompts, the system deduces that the user is already dead or critically disabled, thus, messages are sent to pre-selected recipients. As they are large repositories of passwords, does the hacking community perceive these systems as a virtual El-Dorado? Could these systems not expire before its customers do? “People aren’t very aware of what you might call their living online legacy—potential employers looking at their Facebook accounts, for example. The issue of what happens to that information after their death is an extension of that”, says Yorick Wilks, a senior research fellow at the Oxford Internet Institute [114].

#### ***2.6.14 Digital Identity, Online Reputation and Metadata***

As data become more abundant, the main problem is no longer finding the information but accessing it easily and quickly. What is needed is metadata, which is information about information, to organize the cornucopia of information provided by the internet. In Assyria around three millennia ago clay tablets had small clay labels attached to them to make them easier to tell apart when they were filed in baskets or on shelves. The idea survived into the 20th century in the shape of the little catalogue cards librarians used to note down a book’s title, author, subject, and so on before the records were moved onto computers. The actual books constituted the data, the catalogue cards the metadata. Bar coded and RFID package labels are other examples of metadata. Today, metadata are undergoing a virtual renaissance since many companies are using it to organize information. Google’s search engine creates PageRank metadata to organize web pages by structuring the information, ranking it in order of its relevance to the query. Google handles around half the world’s internet searches, answering around 35,000 queries every second. Metadata are a potentially lucrative business. “If you can control the pathways and means of finding information, you can extract rents from subsequent levels of producers,” explains Eli Noam, a telecoms economist at New York’s Columbia Business School [115].

Metadata could directly affect the digital identity and online reputation since metadata are increasingly become available on the net. Photos uploaded to the website Flickr contain metadata such as when and often where they were taken, as well as the camera model, which could be useful for future buyers. But with the advent of Web 2.0, internet users tag web sites, documents, photos and videos helping to label unstructured information so it can be easily found through folk-minds such as Delicious,<sup>30</sup> Diigo,<sup>31</sup> and Technorati.<sup>32</sup> For any reason, such as for having fun or creating a buzz on the net, Internet users could also instead labeling

---

<sup>30</sup> <http://www.delicious.com>

<sup>31</sup> <http://www.diigo.com>

<sup>32</sup> <http://technorati.com>

a photograph of Barack Obama as “president”, they might bookmark it “sexual harassment”. Thus, this phenomenon would have a negative side affecting the people’s digital identities and reputations [115].

### ***2.6.15 Digital Identity Issue with Cyborg Enhancement***

Identity and privacy issues are immediately important with enhancing implant technology, even in the case of relatively straightforward identification devices. A ‘Cyborg’ is a cybernetic organism, part human, part machine, and is formed by the direct connection between human and technology. In 2002, an implant experiment was carried out through an online collaboration between Columbia University and Reading University. It consists of linking the nervous system of a human with the internet. Intents and purposes the body of that individual does not stop as is usual, but rather extends as far as the Internet takes it. In this case, the human brain was able to directly control a robot hand on a different continent-the Cyborg body extended across the Atlantic Ocean. In this respect, by linking the mental functioning of a human and machine, a hybrid identity is created. By connecting the human nervous system with technology, this not only affects the nature of an individual’s identity but also raising questions as to a new meaning for ‘I’. Who are we if our brain/nervous system is part human part machine? Privacy issues are also pertinent when considering signals being sent into and out of the brain. Feelings, emotions, and even inter-thoughts could potentially be modified by electronic signals alone. Network hacking is far more serious if your brain is permanently connected into the network. Software viruses and biological viruses become, effectively, the same thing. Hence, security, screening and anti-viruses take on much more importance [116].

### ***2.6.16 Digital Identity in Big Data Era***

Information has gone from scarce to superabundant and the quantity of information in the world is soaring and becomes astronomic. Joe Hellerstein, a computer scientist at the University of California in Berkeley, calls it “the industrial revolution of data”. Scientists and computer engineers have coined a new term for the phenomenon: “big data” [94, 117]. Authors provide many examples to illustrate the importance of data deluge. Headquartered in Hong Kong, Li and Fung Ltd.,<sup>33</sup> a major global distribution service company, saw during 2008 one hundred gigabytes of information flow through its network each day; but today the amount has increased tenfold. During 2009, US army’s aircraft flying over Iraq and Afghanistan sent back around 24 years’ worth of video footage. The same author predicted that new aircraft models that

---

<sup>33</sup> <http://www.lifunggroup.com/front.html>



are being deployed in 2010 will produce ten times as many data streams as their predecessors, and those in 2011 will produce 30 times as many. He adds that according to one estimate, mankind created 150 EB (billion gigabytes) of data in 2005 and it will create 1,200 EB in 2010 [94]. In 2000, the telescope of Sloan Digital Sky Survey SDSS collected more data in the first few weeks than had been collected in the entire history of astronomy. Today's SDSS archive contains a whopping 140 TB (240 bytes) of information. A new generation of telescope, the Large Synoptic Survey Telescope, will acquire that quantity of data every 5 days. The retail giant Wal-Mart handles more than one million customer transactions every hour, feeding databases estimated at more than 2.5 PB—the equivalent of 167 times the books in America's Library of Congress [117]. Photobucket, an online photo-sharing service, claims to host more than 4.7 billion digital photographs as of 2008. Facebook reports more than 3 billion photographs, less than 4 years into its existence [111], and reached to home 40 billion photos in 2008 [117]. The author of the Economist article [93] explains that the amount of information is growing at a terrific rate. He adds that experiments at the Large Hadron Collider at CERN generate 40 TB every second and, in 2008, U.S. households were bombarded with 34 gigabytes per person per day [93]. YouTube manages video uploads of 5 h/min early in 2007 to more than 35 h/min in 2010 [118]. The author points that several reasons are driving digital information explosion. He estimates that amount of information increases tenfold every 5 years for the following main reasons: (1) technology is the obvious one. Digital devices soar such as sensors and gadgets are digitizing lots of information that was previously unavailable; (2) there are now many more people who interact with information. Between 1990 and 2005 more than 1 billion people worldwide entered the middle class. As they get richer they become more literate, which fuels information growth [117].

Companies could prosper by gasping new opportunities around big data. The author says that companies could 'pluck the diamond from the waste' by exploiting big data opportunities. Analyzing data could help to spot business trends, prevent diseases, and combat crime. Effective data management could unlock new sources of economic value, provide fresh insights into science and hold governments to account. For instance, exploiting and mining crime figures, maps, details of contracts and statistics that public services are putting into the public domain, or provide the tools for others to do so [94, 117]. Many businesses are providing services based on the access to government data, which recently are made available online. The state is a big generator, collector and user of data. It keeps records on every birth, marriage and death, compiles figures on all aspects of the economy and keeps statistics on licenses, laws and the weather. Until recently all these data have been locked tight and even if they were made publicly accessible they were hard to find, and aggregating lots of printed information is notoriously difficult. Today, things have changed "Government information is a form of infrastructure, no less important to our modern life than our roads, electrical grid or water systems," says Carl Malamud, the boss of Public.Resource Group<sup>34</sup> that puts government data online.

---

<sup>34</sup> <http://public.resource.org>

He was responsible for making the databases of America's Securities and Exchange Commission available on the web during 1990s [119]. The author of the "Clicking for Gold" article [120] explains that the trail of clicks that internet users leave behind from which value can be extracted is becoming a mainstay of the internet economy. "What we are seeing is the ability to have economies form around the data" says Craig Mundie, head of research and strategy at Microsoft. Data are becoming the new raw material of business. Farecast,<sup>35</sup> a part of Microsoft's search engine Bing, can advise customers whether to buy an airline ticket now or wait till the price to come down by analyzing 225 billion flight and price records. Amazon.com is not only tracks the books the user purchases, but also keeps a record of the ones the user only browses in order to recommend other books to him. Information that would be gathered from Amazon's e-book, the Kindle, is probably even richer, how long a user spends reading each page, whether he takes notes and so on [117, 120]. Business intelligence and analytics, which is performing statistical operations for forecasting or uncovering hidden correlations, may allow to firms to gain pay-offs by operating more efficiently, picking out trends and improving forecasting. "Torture the data long enough and they will confess to anything" is a humorous quip made by statisticians to encourage making the most of data. A few years ago business intelligence technologies were available only to big companies, but today the technology has moved into the mainstream. This is due to the fall of the price and better performance of hardware, software and storage. In addition, companies are collecting more data, which in the past they were kept in different systems that were unable to talk to each other, such as finance, human resources or customer management. Now the systems are being linked, and companies are using data-mining techniques to get, "a single version of the truth", which means a complete picture of their operations. Best Buy,<sup>36</sup> an international electronics retailer, found that 7 % of its customers accounted for 43 % of its sales, so it reorganized its stores to concentrate on those customers' needs. The author highlights that data torture depends the accuracy of the information that companies hold. In a study by IBM, half of the managers that are quizzed did not trust the information on which they had to make decisions. Currently, many businesses are increasingly moving to capture accurate data by analyzing real-time information flows instead of stored information about past transactions. Two technology trends are helping to fuel these new uses of data: cloud computing and open-source software. Cloud computing allows organizations to lease on-demand computing power, rather than having to acquire expensive equipment. A free programming language called R<sup>37</sup> lets companies examine and present big data sets, and free software called Hadoop<sup>38</sup> now allows ordinary PCs to analyze huge quantities of data that previously required a supercomputer.

---

<sup>35</sup> <http://www.bing.com/travel>

<sup>36</sup> <http://www.bestbuy.com/site/index.jsp>

<sup>37</sup> <http://www.r-project.org>

<sup>38</sup> <http://hadoop.apache.org>

Two major issues/difficulties that faces data deluge: (1) information storage capabilities: the current situation is a result of a rapid collection of data in a short time and an amount of data that exceeds the available storage space. Based on the forecast of IDC, in 2011, global information will reach around 1,750 EB and available storage about 800 EB. The flood of data from sensors, computers, research labs, cameras, phones and the like surpassed the capacity of storage technologies in 2007; (2) analysis and extraction capabilities of useful information: Alex Szalay, an astrophysicist at Johns Hopkins University, notes that the proliferation of data is making them increasingly inaccessible and he points that we should be able to make sense of them. Only few industries have developed such capabilities. Credit-card companies monitor every purchase and can identify fraudulent ones. They found that stolen credit cards are more likely to be used to buy hard liquor than wine for many reasons such as it is easier to fence. Insurance firms combine clues to spot suspicious claims. They found that fraudulent claims are more likely to be made on a Monday, since policyholders who stage accidents tend to assemble friends as false witnesses over the weekend. Mobile-phone operators, meanwhile, analyze subscribers' calling patterns to offer them customized attractive promotions. Also, retailers, offline as well as online, can tailor promotions to particular customers' preferences. The oil industry uses supercomputers to trawl seismic data before drilling wells [93, 94, 117]. In addition, another concern as the torrent of information increases is energy consumption. Processing huge amounts of data takes a lot of power. "In 2–3 years we will saturate the electric cables running into the building," says Alex Szalay at Johns Hopkins University. "The next challenge is how to do the same things as today, but with ten to one hundred times less power". The NSA in 2006 came close to exceeding its power supply, which would have blown out its electrical infrastructure. Both Google and Microsoft put some of their huge data centers next to hydroelectric plants to ensure access to enough energy and at a reasonable price [105].

Ensuring data security and protecting privacy is becoming harder as the information multiplies and is shared widely around the world. According to Cisco, by 2013, the amount of traffic flowing over the internet annually will reach 667 EB and the quantity of data continues to grow faster than the ability of the network to carry it all [117]. A researcher of the University of California in San Diego says: "information created by machines and used by other machines will probably grow faster than anything else". He adds that "this is primarily 'database to database' information—people are only tangentially involved in most of it" [93]. The author of the article "new rules for big data: regulators are having to rethink their brief" [121] points that current information flows in an era of abundant data are changing the relationship between technology and the role of the government. He adds that many of today's regulations are not brought up-to-date such as privacy laws, which they were not designed for networks, and rules for document retention presume paper records. Now information becomes interconnected and that's why nations are increasingly in need of global rules. The same author mentions that new information-related principles should cover the following broad areas: information privacy, security, retention, processing, ownership, and integrity to reduce risks posed by the

age of big data sets [121]. More details are given by the author [94] to explain the consequences of only two data deluge risks, which are identity theft and fraud, and privacy breaches. He explains that they are consequences of stolen databases, such as disks full of social-security data are missed, laptops loaded with tax records are left in taxis, credit-card numbers are stolen from online retailers. Privacy infringements are encountered in daily basis. For instance, we can witness the periodic fusses when Facebook or Google unexpectedly change the privacy settings on their online social networks, causing members to reveal personal information unwittingly. A more sinister threat is encountered when governments compel companies to hand over personal information about their customers. In order to deal with the drawbacks of data deluge, the author suggests that people should have greater ownership, access, and control over their digital identity. For instance, Google allows users to see what information it holds about them, and lets them delete their search histories or modify the targeting of advertising. Secondly, organizations should be required to disclose details of security and privacy breaches to encourage managers to take information security and privacy more seriously. Finally, organizations should be subject to an annual digital identity and privacy audit to encourage organizations to keep their security measures up to date [94].

## References

1. J. Amos, Etched ostrich eggs illustrate human sophistication (2010), Available: <http://news.bbc.co.uk/2/hi/science/nature/8544332.stm>. Accessed 29 Aug 2010
2. P.J. Texier, *Une découverte remet en question l'origine de "l'écriture"* (Science & Vie, 2010)
3. Wikipedia, Diepkloof Rock Shelter (2010), Available: [http://en.wikipedia.org/wiki/Diepkloof\\_Rock\\_Shelter](http://en.wikipedia.org/wiki/Diepkloof_Rock_Shelter). Accessed 29 Aug 2010
4. P.J. Texier et al., A Howiesons Poort tradition of engraving ostrich eggshell containers dated to 60,000 years ago at Diepkloof Rock Shelter, South Africa, (2010), Available: <http://www.pnas.org/content/107/14/6180.full.pdf+html>. Accessed 29 Aug 2010
5. L. Casson, *Libraries in the Ancient World* (Yale University Press, 2002)
6. A. Whitaker, Nippur, Available: <http://www.ancient-wisdom.co.uk/iraqnippur.htm>. Accessed 25 Sept 2010
7. G.E.A. Akerlof, R.E. Kranton, *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being* (Princeton University Press, 2010)
8. Merriam-Webster Online Dictionary, Definition of identity (2010), Available: <http://www.merriam-webster.com/dictionary/identity>. Accessed 20 Sept 2010
9. The American Heritage Dictionary of the English Language, Definition of identity (2009), Available: <http://www.thefreedictionary.com/identity>. Accessed 20 Sept 2010
10. H. Noonan, Identity, in *Stanford Encyclopedia of Philosophy* (2009)
11. Wikipedia, Nasreddin (2010), Available: <http://en.wikipedia.org/wiki/Nasreddin>. Accessed 23 Sept 2010
12. I. Shah, *The Exploits of the Incomparable Mulla Nasrudin—The Subtleties of the Inimitable Mulla Nasrudin* (Octagon Press, London, 1989)
13. C. Lindholm, *Culture and Identity—The History, Theory, and Practice of Psychological Anthropology* (Oneworld Publications, 2007)
14. E.T. Olson, Personal identity, in *Stanford Encyclopedia of Philosophy* (2008)
15. Identity over time, in *Stanford Encyclopedia of Philosophy* (2005)

16. G. Group, Identity/identity formation, in *Gale Encyclopedia of Psychology* (2001)
17. Wikipedia, Identity (philosophy) (2010), Available: [http://en.wikipedia.org/wiki/Numerical\\_identity#Qualitative\\_versus\\_numerical\\_identity](http://en.wikipedia.org/wiki/Numerical_identity#Qualitative_versus_numerical_identity). Accessed 20 Sept 2010
18. Wikipedia, Cultural identity (2010), Available: [http://en.wikipedia.org/wiki/Cultural\\_identity](http://en.wikipedia.org/wiki/Cultural_identity). Accessed 23 Sept 2010
19. S. Roccas, M.B. Brewer, Social identity complexity. *Pers. Soc. Psychol. Rev.* **6**, 88–106 (2002)
20. J. Balmer, Corporate identity: the power and the paradox. *Des. Manag. J.* **6**, 39–44 (1995)
21. Poem Hunter, Quotations from Salvador Minuchin (2010), Available: <http://www.poemhunter.com/quotations/famous.asp?people=Salvador%20Minuchin>. Accessed 25 Oct 2010
22. C. Taylor, *Sources of the Self: The Making of the Modern Identity* (Harvard University Press, 1989)
23. G. Warnke, *After Identity: Rethinking Race, Sex, and Gender* (Cambridge University Press, Cambridge, 2007)
24. Facing History and Ourselves Foundation, *Stories of Identity: Religion, Migration, and Belonging in a Changing World* (2008)
25. R. Jenkins, *Social Identity: Key Ideas*, 3rd edn. (2008)
26. T. Miyata et al., A survey on identity management protocols and standards. *Oxford Journals: IEICE Trans. Inf. Syst.* (special section on new technologies and their applications of the internet III) **89**, 112–123 (2006)
27. M. Benantar, *Access Control Systems: Security, Identity Management and Trust Models* (Springer Science+Business Media, 2006)
28. A. Jøsang, S. Pope, User-centric identity management, in *Proceedings of the AusCERT Asia Pacific Information Technology Security Conference* (2005), pp. 1–6
29. P.J. Windley, *Digital Identity: Unmasking Identity Management Architecture (IMA)* (O'Reilly Media, 2005)
30. Organization for Economic Co-operation and Development (OECD), At Crossroads: Personhood and Digital Identity in the Information Society. *The Working Paper series of the OECD Directorate for Science, Technology and Industry* (2008), Available: [http://www.oecd.org/LongAbstract/0,3425,en\\_2649\\_34223\\_40204774\\_119684\\_1\\_1\\_1,00.html](http://www.oecd.org/LongAbstract/0,3425,en_2649_34223_40204774_119684_1_1_1,00.html). Accessed 21 May 2010
31. E. Damiani et al., Managing multiple and dependable identities. *IEEE Internet Comput.* (IEEE Comput Soc) **7**, 29–37 (2003)
32. Identity Gang Group—Working Group of Identity Common, Identity Gang Lexicon, Available: <http://wiki.idcommons.net/Lexicon>. Accessed 10 May 2010
33. Princeton University Wordnet—Lexical Database for English, Identity definition [Online], Available: <http://wordnetweb.princeton.edu/perl/webwn?o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&s=identity&i=1&h=0000#c>. Accessed 10 May 2010
34. E. Goffman, *The Presentation of Self in Everyday Life* (1956)
35. S. Williams, This is Me digital identity and reputation on the internet, Available: <http://www.slideshare.net/shirleyearley/this-is-medigital-identity-and-reputation-on-the-internet>. Accessed 27 May 2010
36. Center for Democracy and Technology, Privacy principles for identity in the digital age (Draft for Comment—Version 1.4) (2007), Available: [http://www.cdt.org/files/pdfs/20071201\\_IDPrivacyPrinciples.pdf](http://www.cdt.org/files/pdfs/20071201_IDPrivacyPrinciples.pdf). Accessed 28 May 2010
37. J. Hodges et al., *Glossary for the OASIS Security Assertion Markup Language (SAML) V2.0* (OASIS, 2005)
38. J. Hodges, Liberty technical glossary. *Liberty Alliance Project* (2006)
39. National Research Council—Committee on Authentication Technologies and Their Privacy Implications, Who Goes There? Authentication through the lens of privacy (2003), Available: [http://books.nap.edu/catalog.php?record\\_id=10656#toc](http://books.nap.edu/catalog.php?record_id=10656#toc). Accessed 7 June 2010
40. Ofcom: Office of Communication, Social networking: a quantitative and qualitative research report into attitudes, behaviours and use (2008), Available: <http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/report1.pdf>. Accessed 27 Aug 2010

41. D. Gollmann, Identity and location, in *Security Protocols*, ed. by B. Christianson et al., vol. 3957 (Springer, 2006), pp. 246–250
42. J. Carroll, J. Murphy, Who am I? I am Me! Identity management in a networked world, in *Proceedings of the 4th International We-B Conference* (2003)
43. International Telecommunication Union, *Digital Life*. ITU Internet Report (2006), Available: <http://www.itu.int/osg/spu/publications/digitalife/docs/digital-life-web.pdf>. Accessed 21 May 2010
44. Scientific American Podcast, When the Virtual You Changes the Real You, in *Scientific American Magazine* (2007)
45. D. Teten, S. Allen, *The Virtual Handshake: Opening Doors and Closing Deals Online* (Amacom, 2005)
46. T. Boellstorff, *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*, (Princeton University Press, 2008)
47. T. Flew, *New Media: An Introduction* (Oxford University Press, 2002)
48. N. Reichenthal, Une identité connectée. Bulletin HEC—Le Magazine des Gradués (2007), Available: <http://www.gradueshec.ch/bulletins/documents/75nadine.pdf>. Accessed 13 May 2010
49. S. Clauß, M. Köhntopp, Identity management and its support for multilateral security. *Comput. Netw.* **37**, 205–219 (2001)
50. Organizing Committee of Digital Identity and Privacy (Human Capital and Social Innovation Technology Summit), Call for contribution to managing digital identities for education, employment and business development (2007), Available: <http://events.eife-l.org/HCSIT2007/overview/dip/dip2007>. Accessed 11 May 2010
51. P. Windley, Digital identity perspectives (2005), Available: <http://www.windley.com/stories/2004/04/20/digitalIdentityPerspective>. Accessed 5 Oct 2010
52. J. De Clercq, J. Rouault, An introduction to identity management (2004), Available: [http://devresource.hp.com/drc/resources/idmgt\\_intro/idmgt\\_intro.pdf](http://devresource.hp.com/drc/resources/idmgt_intro/idmgt_intro.pdf). Accessed 12 July 2007
53. R.D. Ashton et al., An organizing framework for collective identity: articulation significance of multidimensionality. *Psychol. Bull.* **130**, 80–114 (2004)
54. G. Ben Ayed, Consolidating fragmented identity: attributes aggregation to secure information systems. *IADIS Int. J. Comput. Sci. Inf. Syst.* **4**, 1–12 (2009)
55. J. Goldman, K.A. Brower, Obama's advice to aspiring politicians: be careful on facebook (2009), Available: <http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aL6GJ25zYajY>. Accessed 4 Oct 2010
56. G. Bell, J. Gemmel, A digital life. *Sci. Am. Mag.* **296**, 58–65 (2007)
57. E. Hoog, Comment civiliser Internet? Le nouvel observateur (2009), pp. 20–21, Available: <http://hebdo.nouvelobs.com/sommaire/les-debats/087207/comment-civiliser-internet.html>. Accessed 18 May 2010
58. MyLifeBrand: social network aggregator playing catch up, Available: [http://www.readwriteweb.com/archives/mylifebrand\\_social\\_network\\_aggregator.php](http://www.readwriteweb.com/archives/mylifebrand_social_network_aggregator.php). Accessed 25 May 2010
59. Mashable: the social media guide, Available: <http://my.mashable.com/>. Accessed 25 May 2010
60. Profilactic: a social media aggregator/life streaming service, Available: <http://www.profilactic.com/>. Accessed 25 May 2010
61. Snag: a social networks aggregator. Available: <http://www.dapper.net/dapplications/Snag/>. Accessed 25 May 2010
62. Profileomat: a shareable profile aggregator, Available: <http://www.profileomat.com/>. Accessed 25 May 2010
63. Naymz Features, Available: <http://www.naymz.com/about.action?section=compare>. Accessed 26 May 2010
64. SocialURL: showcase all your web profiles with a single, Available: <http://socialurl.com/default.aspx>. Accessed 26 May 2010
65. PeopleAggregator, Available: <http://www.peopleaggregator.net/>. Accessed 26 May 2010
66. ProfileFly: control and promote your entire online identity, Available: <http://profilefly.com/>. Accessed 26 May 2010



67. SocialNetwork.in, Available: <http://socialnetwork.in/index.cfm>. Accessed 26 May 2010
68. Mashable, Available: <http://mashable.com/>. Accessed 26 May 2010
69. SocialStream, Available: <http://www.hcii.cmu.edu/M-HCI/2006/SocialstreamProject/index.php>. Accessed 26 May 2010
70. 8hands Intelligent Social Network Aggregator, Available: <http://www.logiagroup.com/socialNetworks.html>. Accessed 26 May 2010
71. 8hands Software (version 0.9.135): download page, Available: [http://download.cnet.com/hands/3000-12941\\_4-1066775.html](http://download.cnet.com/hands/3000-12941_4-1066775.html). Accessed 26 May 2010
72. NoseRub, Available: <http://noserub.com/>. Accessed 26 May 2010
73. NoseRub application (Version 0.8.2): download page, Available: <http://noserub.com/download/>. Accessed 26 May 2010
74. Minggl, Available: <http://doyou.minggl.com/>. Accessed 26 May 2010
75. Minggl Download Page (Firefox add-on), Available: <http://doyou.minggl.com/welcome/install/>. Accessed 26 May 2010
76. G. Ben Ayed, S. Ghernaoui-Hélie, Digital identity attributes cohesion: major issues and challenges for e-services access control, in *Presented at the International Conference on Information Technology and E-service (ICITeS'2011)*, Sousse, Tunisia, 2011
77. Wink: people search engine, Available: <http://wink.com/>. Accessed 25 May 2010
78. Secondbrain: save, share and discover great bookmarks, Available: <http://secondbrain.com/>. Accessed 26 May 2010
79. Wikipedia, Context awareness (2011), Available: [http://en.wikipedia.org/wiki/Context\\_awareness](http://en.wikipedia.org/wiki/Context_awareness). Accessed 23 Aug 2011
80. A.R. Galloway, E. Thacker, *The Exploit—A Theory of Networks* (University of Minnesota Press, 2008)
81. A. Ross, A.K. Jain, Biometrics: when identity matters, in *Advances in Biometric Person Authentication*, vol. 3338/2005 (Springer, 2005), pp. 137–149
82. J. Fildes, Taking control of your digital id (2006), Available: <http://news.bbc.co.uk/2/hi/technology/6102694.stm>. Accessed 22 May 2010
83. M. Fischetti, Scoring your identity: new tactics root out the false use of personal data. *Sci. Am.* 27–28 (2007)
84. P. Cochrane, Forward of the book, in *Digital Identity Management: Perspectives on the Technological, Business and Social Implications*, ed. by D.G.W. Birch (Gower Publishing Limited, 2007)
85. A. Henderson, Practical action: federation and mobility, in *Digital Identity Management: Perspectives on the Technological, Business and Social Implications*, ed. by D.G.W. Birch (Gower Publishing Limited, 2007), pp. 72–88
86. Facebook, Facebook statistics (2011), Available: <http://www.facebook.com/press/info.php?statistics>. Accessed 13 Nov 2011
87. J. Goodchild, The robin sage experiment: fake profile fools security pros (2010), Available: <http://www.networkworld.com/news/2010/070810-the-robin-sage-experiment-fake.html?t51hb>. Accessed 8 Nov 2011
88. B. Ferran, Une fausse chercheuse dupe des experts en sécurité (2010), Available: <http://www.lefigaro.fr/web/2010/07/24/01022-20100724ARTFIG00481-une-fausse-chercheuse-dupe-des-experts-en-securite.php>. Accessed 8 Nov 2011
89. G. Tissier et al., Les marchés noirs de la cybercriminalité. Compagnie Européenne d'Intelligence Stratégique (CEIS) (2011)
90. Wikipedia, E Pluribus Unum (2011), Available: [http://en.wikipedia.org/wiki/Out\\_of\\_Many\\_One](http://en.wikipedia.org/wiki/Out_of_Many_One). Accessed 25 Oct 2010
91. Wikipedia, Great seal of the United States (2010), Available: [http://en.wikipedia.org/wiki/Great\\_Seal\\_of\\_the\\_United\\_States](http://en.wikipedia.org/wiki/Great_Seal_of_the_United_States). Accessed 25 Oct 2010
92. S.L. Garfinkel, Information of the World. *UNITE! Sci. Am. Mag.* 82–87 (2008)
93. K. Cukier, *All Too Much: Monstrous Amounts of Data* (The Economist, 2010), Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557421](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557421). Accessed 13 May 2010



94. K. Cukier, *The Data Deluge: Businesses, Governments and Society are Only Starting to Tap its Vast Potential* (The Economist, 2010). Available: [http://www.economist.com/opinion/displaystory.cfm?story\\_id=15579717](http://www.economist.com/opinion/displaystory.cfm?story_id=15579717). Accessed 13 May 2010
95. Garlik, Garlik's DataPatrol (2010), Available: <http://www.garlik.com>. Accessed 6 Aug 2010
96. Mypublicinfo, Mypublicinfo: identity protection services (2010), Available: <http://www.mypublicinfo.com>. Accessed 26 Aug 2010
97. G. Ben Ayed, S. Ghernaoui-Hélie, Digital identity attributes cohesion to access e-services: major issues and challenges in digital society. *J. E-Technol.* **2** (2011)
98. G. Ben Ayed, S. Ghernaoui-Hélie, Claim-based digital identity fusion to access e-services: major issues and challenges in digital society, in *International Conference on Information and Computer Applications (ICICA 2011)*, Dubai, UAE, 2011
99. J. Harris, S. Kamvar, We feel fine project: an exploration of human emotions, in six movements, Available: <http://wefeelfine.org/>. Accessed 15 May 2010
100. J. Harris, S. Kamvar, we feel fine project: blogs data [Online], Available: <http://www.wefeelfine.org/data/files/feelings.txt>. Accessed 25 May 2010
101. M. Hansen, B. Rubin, Listening Post Project, Available: <http://www.earstudio.com/projects/listeningpost.htm>. Accessed 25 May 2010
102. S. Billard, Référencement, Design et Cie (2008), Available: <http://s.billard.free.fr/referencement/?2008/10/27/521-123people-moteur-de-recherche-de-personnes>. Accessed 25 Oct 2010
103. Google, About News Archive Search, Available: <http://news.google.com/archivesearch/about.html>. Accessed 19 May 2010
104. H. McCallum-Bayliss, Identity resolution in a global environment: fishing for people in a sea of names. *IEEE IT Prof.* **6**, 21–26 (2004)
105. K. Cukier, Cupo (The Economist, 2010), Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557507](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557507). Accessed 13 May 2010
106. R. Bowes, A Quick Hack to Download Facebook URLs from Facebook Directory (2010), Available: <http://www.skullsecurity.org/blogdata/facebook.rb>. Accessed 29 Aug 2010
107. R. Callow, 100 million Facebook users added to a publicly available torrent file (2010), Available: Sync-Blog <http://www.sync-blog.com/sync/2010/07/100-million-facebook-users-added-to-a-publicly-available-torrent-file.html>. Accessed 29 Aug 2010
108. R. Bowes, Return of the Facebook snatchers (2010), Available: SkullSecurity Blog <http://www.skullsecurity.org/blog/?p=887>. Accessed 29 Aug 2010
109. R. Paul, Leaked. Data of 100 M Facebook Users Came from Public Info (2010), Available: <http://arstechnica.com/security/news/2010/07/leaked-data-of-100m-facebook-users-came-from-public-info.ars>. Accessed 14 Oct 2010
110. Wikipedia, Digital native, Available: [http://en.wikipedia.org/wiki/Digital\\_native](http://en.wikipedia.org/wiki/Digital_native). Accessed 18 May 2010
111. J. Palfrey, U. Gasser, *Born Digital: Understanding the First Generation of Digital Natives* (Basic Books, 2008)
112. J. Howe, *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*, Crown Business, (2008)
113. United Nations Population Division—Department of Economic and Social Affairs (1999) *The World at Six Billion*. Available: <http://www.un.org/esa/population/publications/sixbillion/sixbillion.htm> Accessed: May 18 2010
114. D. Jefferies, Preparing for the Digital Afterlife, in *The Guardian Newspaper*, ed, 2009
115. K. Cukier, Needle in a Haystack: The uses of Information about Information. *The Economist* (February 23–March 5). Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557497](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557497) Accessed: May 13, 2010
116. K. Warwick, *Cyborg identity in digital identity management: perspectives on the technological, business and social implications*, ed by D. G. W. Birch (Gower Publishing Limited 2007), pp. 227–238
117. K. Cukier, Data, data everywhere. *The economist* (February 23rd–March 5th). Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557443](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557443) Accessed: May 13, 2010

118. D. Durand, Nouvelles vidéos sur Youtube: +50'000 heures chaque jour ! évolution depuis 2007. Available: <http://www.zdnet.fr/blogs/media-tech/nouvelles-vidéos-sur-youtube-50-000-heures-chaque-jour-evolution-depuis-2007-39756042.htm> Accessed: Nov. 12, 2010
119. K. Cukier, The open society: governments are letting in the light. The economist (February 23–March 5). Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557477](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557477) Accessed: May 13, 2010
120. K. Cukier, (2010) Clicking for gold: how internet companies profit from data on the web. The economist (February 23–March 5). Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557431](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557431) Accessed: May 13, 2010
121. K. Cukier, (2010) New rules for big data: regulators are having to rethink their brief. The economist (February 23–March 5). Available: [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557487](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557487) Accessed: May 13, 2010
122. Spokeo: people finder, Available: <http://www.spokeo.com/>. Accessed 26 May 2010
123. J.J. Luna, *How to Be Invisible: A Step-by-Step Guide to Protecting Your Assets, Your Identity, and Your Life* (Farrar, Straus, and Giroux, New York, 2000)

Architecting User-Centric Privacy-as-a-Set-of-Services  
Digital Identity-Related Privacy Framework

Ben Ayed, G.

2014, XIX, 177 p. 47 illus., 4 illus. in color., Hardcover

ISBN: 978-3-319-08230-1