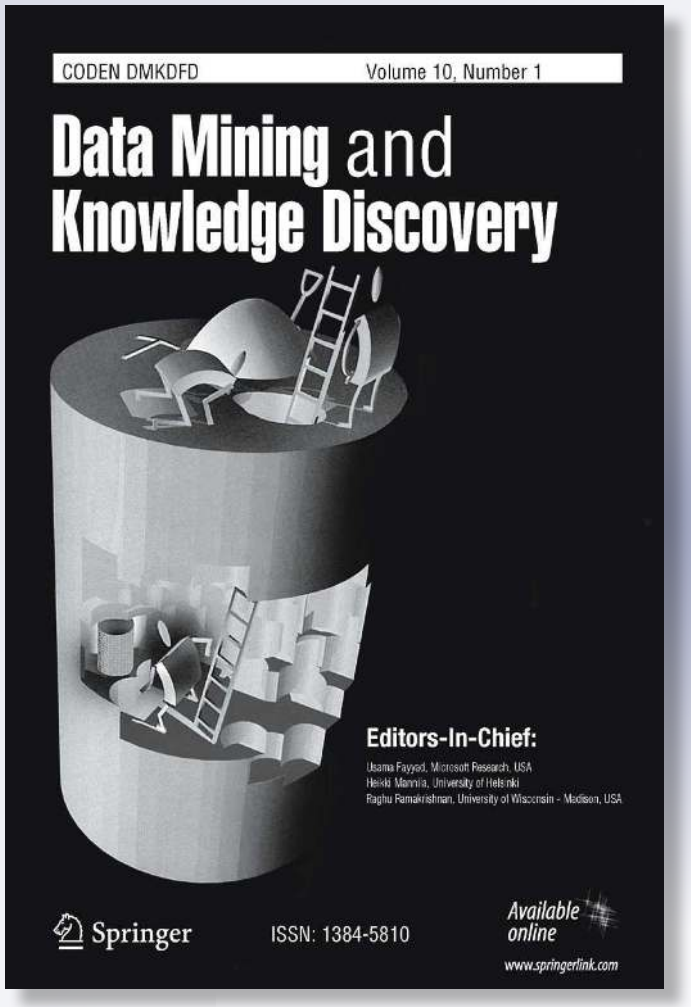
More than modelling and hiding: towards a comprehensive view of Web mining and privacy

Bettina Berendt

Data Mining and Knowledge Discovery

ISSN 1384-5810
Volume 24
Number 3

Data Min Knowl Disc (2012) 24:697-737
DOI 10.1007/s10618-012-0254-1



Springer

Data Min Knowl Disc (2012) 24:697–737
DOI 10.1007/s10618-012-0254-1

More than modelling and hiding: towards a comprehensive view of Web mining and privacy

Bettina Berendt

Received: 11 January 2010 / Accepted: 20 January 2012 / Published online: 15 February 2012

Abstract Over the last decade, privacy has been widely recognised as one of the major problems of data collections in general and the Web in particular. This concerns specifically data arising from Web usage (such as querying or transacting) and social networking (characterised by rich self-profiling including relational information) and the inferences drawn from them. The data mining community has been very conscious of these issues and has addressed in particular the inference problems through various methods for "privacy-preserving data mining" and "privacy-preserving data publishing". However, it appears that these approaches by themselves cannot effectively solve the privacy problems posed by mining. We argue that this is due to the underlying notions of privacy and of data mining, both of which are too narrow. Drawing on notions of privacy not only as hiding, but as control and negotiation, as well as on data mining not only as modelling, but as the whole cycle of knowledge discovery, we offer an alternative view. This is intended to be a comprehensive view of the privacy challenges as well as solution approaches along all phases of the knowledge discovery cycle. The paper thus combines a survey with an outline of an agenda for a comprehensive, interdisciplinary view of Web mining and privacy.

Keywords Web mining · Privacy · Knowledge discovery cycle · Web usage (mining) · Social network analysis

Responsible editor: Myra Spiliopoulou, Bamshad Mobasher, Olfa Nasraoui, Osmar Zaiane.

B. Berendt (✉)
Department of Computer Science, K.U. Leuven, Leuven, Belgium
e-mail: bettina.berendt@cs.kuleuven.be
URL: http://people.cs.kuleuven.be/bettina.berendt/

1 Introduction

Over the last decade, privacy has been widely recognised as one of the major problems of data collections in general and the Web in particular. Hardly a day goes by without new reports of data leaks and the privacy violations arising from such data and the inferences drawn from them. On the Web, this concerns in particular data arising from Web usage (such as querying or transacting) and social networking (characterised by rich self-profiling including information about one's relations to others).

Highly publicised examples include (1) the AOL search logs from whose anonymized query collections two New York Times journalists re-identified at least one person (Barbaro and Zeller 2006; Reuters 2006); (2) Facebook Beacon, an alerting system that informed a user's "friends" on the social network of what this user had bought on affiliated sites (EPIC 2011b); and (3) repeated incidents of employees or students being penalised by their superiors for utterances or deeds publicised on their private blogs or social-network profiles (cf. Smith-Spark 2006). The privacy violations in these examples arise from the retrieval and re-purposing of data (2 and 3) and from inferences from anonymous identifiers to personal identities (1). In addition, example (2) illustrates some possible consequences of recommender systems based on social relations, which has been investigated in data mining as "viral marketing". The data mining community has been very conscious of these issues and has addressed in particular the inference problems through various methods that are often collectively referred to as "privacy-preserving data mining" (cf. Verykios et al. 2004; Aggarwal and Yu 2008b). Results include powerful methods for ensuring desired degrees of anonymity, for preventing the inference of certain undesired patterns and/or for preventing inferences on individuals while retaining the potential to infer aggregate patterns. For example, in a banking context the inference of rules that derive "low creditworthiness" from an individual's nationality might be prevented; alternatively, individuals' nationalities may be swapped such that the rules can still be inferred, but any given individual's nationality cannot. These methods can be complemented by cryptographic and architectural solutions for keeping communications confidential, by the distribution of data, and by access restrictions. Together, these methods have high potential for more secure and selective data and pattern retrieval and inference. So does this mean we have a solution to the problem, does it mean we can (as the name of the field suggests) have the cake of privacy and data-mine it too?

We argue that there are two basic problems with this approach: the underlying notions of privacy and of data mining.

First, the approach relies on the basic assumption that privacy is preserved when certain information is hidden, privacy-as-confidentiality. This assumption underlies much of the current work on privacy in computer science. But why are implementations of these theoretical results not adopted by users on a large scale? Why is it that even though users have professed to valuing their privacy throughout (and even before) this decade of Web mining (e.g., Teltzrow and Kobsa 2003), they have kept sharing highly sensitive data with peers (as in social networks, see the examples above) and others (as in e-Commerce, cf. Berendt et al. 2005)?

Second, the approach relies on a view of data mining in a narrow sense: The term "data mining" is sometimes used to denote the whole knowledge-discovery cycle, as

described, for example, in the CRISP-DM (2000) model, sometimes to denote only the "modelling" phase of it (cf. Kohavi and Provost 1998). Most current approaches to "privacy-preserving data mining" focus on alterations in the modelling phase, i.e. on the discovery of patterns from the given data or from data modified in an additional data preparation phase. These approaches have proven results such as guaranteeing the re-identifiability of an individual only as one of k persons. So why is it that data mining remains being seen as a major foe to privacy? What is the use of a privacy-preserving association-rule mining method when a classifier is learned from the modified data coupled with another information source—where, in addition, any of the sources may be from, say, a stolen laptop or from a CD-ROM with sensitive information that was mistakenly mailed to the wrong address (e.g., BBC News 2008)? What are the (dis)advantages of not being personally re-identifiable when aggregate patterns can be applied to a random social network profile and relationships to predict, e.g., a person's political affiliation (Lindamood et al. 2009)?

To address these shortcomings, we propose to draw on wider notions of the two key components, emphasising that privacy is more than hiding and that data mining is more than modelling.

The contribution of the paper is twofold: First, we present a conceptualisation of privacy that draws on work in computer science, law, economics and sociology. Second, we develop a more comprehensive view of privacy in Web mining by investigating, for each phase of the CRISP-DM knowledge discovery cycle, privacy threats, opportunities and solution approaches. In this sense, the paper is a combination of survey and agenda for an interdisciplinary research programme on privacy in, for and through Web mining. We focus on examples from research on Web usage and social networks, draw on other examples where necessary, and sketch generalisations to other areas of Web/data mining. The reason for this focus is twofold: first, the context of this Special Issue on Web mining; and second, the specifics of the Web as a source of data and an environment for its analysis.

In Sect. 2, we will explain the wider notion of privacy and how it relates to data and (Web) data mining, and conclude with a more specific description of the goal of this article: how to extend the notion of privacy beyond hiding (confidentiality) to also encompass privacy as control and privacy as practice, how to extend the notion of data mining as modelling to data mining as knowledge discovery, and how to examine their interrelationships. In Sect. 3, we will then investigate these interrelationships, by stepping through the different phases of knowledge discovery and investigating their specific threats to and opportunities for privacy, and we will give overviews of solution approaches specifically addressing the problems arising in these phases. Section 4 describes external effects as a key current challenge and new research direction, and Sect. 5 summarises with an outlook.

2 What is privacy?

In this section, we will start from the data that are usually regarded as being implicated in privacy debates, data protection or privacy breaches on the Internet (Sect. 2.2). We then proceed to describe three relevant groups of privacy approaches (Sects.

Table 1 The three privacy definitions used in this paper

Privacy as hiding: Confidentiality
"the right to be let alone", the right to a private sphere—which is potentially threatened by the disclosure of (personal) data
Privacy as control: Informational self-determination
"the right of the individual to decide what information about himself should be communicated to others and under what circumstances"
Privacy as practice: Identity construction
"the freedom from unreasonable constraints on the construction of one's own identity", be it by strategically being able to reveal or conceal data

2.3–2.5). The aim of this framework is to abstract from and complement existing privacy definitions, as explained in Sect. 2.1. In Sects. 2.6–2.9, we illustrate the various faces of privacy using the example of the AOL logs, highlight different notions of identification, extend the discussion by investigating *whose* privacy may be at stake, and characterise specifics of the Web relevant to our questions. We summarise with the resulting goal of this article, to be then presented in the subsequent Sect. 3.

2.1 Aims and non-aims of this section

There are a myriad of definitions of privacy, not only in computer science, but also in legal, social and other sciences. It is not our purpose here to list them all or investigate all of their details. What we aim to do instead is to present a high-level taxonomy of approaches by asking what the global focus of definitions is. We will refer to selected individual definitions to illustrate details and to selected surveys to illustrate ranges of definitions.

The taxonomy distinguishes between three views of privacy: privacy as hiding, as control, and as practice. These are summarised in Table 1 and will be detailed in Sects. 2.3–2.5. (These sections are closely based on Gürses and Berendt 2010b and parts of Gürses and Berendt 2010a; an extended version can be found in Gürses 2010. The definitions have been put into the context of the present article; see in particular Sect. 2.3.1.) We will show how these three views cut across individual definitions' differences with regard to the subjects of privacy and the type of data and knowledge.

While the aim of Sect. 2 is thus to generalise beyond the multitude of privacy definitions in order to define a general framework for the analysis in Sect. 3, we do make one choice throughout the remaining paper: We focus on the privacy of persons. This is in line with the majority of current scientific and popular treatments, but the choice was also made to respect the complexity of the topic. We therefore begin this section with a definition of "personal data" as the focus of persons' privacy. We complement this general focus of the paper by a discussion of some issues of business and state secrets in Sect. 2.8. A thorough analysis of these questions would require the space of another article.

2.2 Privacy and personal data

Since computers are about data and data processing, any concept of privacy in computational environments will concern data, in particular "personal data". Personal data

is "any information relating to an identified or identifiable natural person […]; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity" (EU Directive 95/46/EC, EU 1995), Art. 2 (a); this is defined in a similar way in other legal contexts). Notice the emphasis on identity, which is assumed to be unique and identifiable for one natural person; in line with this emphasis, US terminology talks about *personally identifiable information*. The standard types of personal data are profile data describing individuals, including name, address, health status, etc.

An important question for Web mining is whether IP addresses are personal data. Privacy advocates have long argued that they are (e.g. Privacy International 2007). The European Union Article 29 Data Protection Working Party (2009) has recently adopted this position. If this becomes official policy, it is likely to have strong effects on how organisations can collect and process Web usage data (cf. the argument of Google's Global Privacy Counsel that IP addresses are not personal data, Fleischer 2008).

Thus, it is not the content of a piece of data that defines it as protection-worthy, or the fact that the data were produced by an individual, pertain to an individual, or describe an individual; the only relevant question is whether the data can be linked to a natural person. European law permits the analysis of data if records are identified by a pseudonym. Under US law, a record holder may assign a code to a de-identified record in order to permit the original record holder to identify the record (45 CFR 164.514(b)(2)(a), 45 CFR 164.514(c), 65 Fed. Reg. at 82818).

2.3 Privacy as hiding: confidentiality

In a classical article, privacy has been defined as "the right to be let alone" (Warren and Brandeis 1890). Although originally formulated as a right that protects individuals against gossip and slander, this construct has since then acquired a wider meaning. Namely, it refers to an individualistic liberal tradition in which an intrinsic pre-existing self is granted a sphere of autonomy free from intrusions from both an overbearing state and the pressure of social norms (Phillips 2004).

That privacy encompasses this sense of a protected sphere is widely accepted in sociology (cf. the first of four types of privacy in Phillips 2004: "freedom from intrusion"), and legal scholars, courts and regulators have recognised its data-dependency: The private sphere is something which is potentially threatened by the disclosure of (personal) data. This notion is also popular in computer science and has been interpreted as an autonomous (digital) sphere in which the data about persons is protected, such that outside of this sphere the data remains confidential.

Data confidentiality—the protection of data from unauthorized access—is a strong and useful translation of such privacy concerns into digital space. A key reason is that once data about a person exists in a digital form, it is very difficult to provide individuals with any guarantees on the control of that data. Data collected using current technologies represent activities of individuals in social life that for many are assumed to be private. To preserve privacy is then to keep this data private, in other

words confidential from a greater public. Not exchanging any data would preserve privacy but is inconvenient and probably also not desirable. Therefore, a lot of the privacy research in computer science is concerned with weaker forms of data confidentiality such as anonymity.

Anonymity is achieved by unlinking the identity of the person from the traces that her activities leave in information systems. Anonymity keeps the identity of the persons in information systems confidential, but it is not necessarily concerned with how public the traces subsequently become. This is also reflected in data protection legislation, which by definition cannot and does not protect anonymous data (Guarda and Zannone 2009).

Anonymity can be based on different models. In communications, anonymity is achieved when an individual is not identifiable within a limited set of users, called the *anonymity set* (Pfitzmann and Hansen 2006–2010). An individual carries out a transaction anonymously if she cannot be distinguished by an observer from others in that set. The observer (*adversary*) may obtain some additional information. This is typically information about the likelihood of different individuals having carried out a given transaction. The observer may be the service provider or some other party with observation capabilities or with the ability to actively manipulate messages. Depending on the observer's capabilities, different models can be constructed with varying degrees of anonymity for the given anonymity set. What degree of anonymity is sufficient in a given context depends on legal and social consequences of a data breach and remains an open question (Díaz 2005).

In databases and *privacy-preserving data mining* (PPDM), the conditions for establishing anonymity sets and the targeted objectives are somewhat different than in communications. Anonymity is a popular requirement when (Web or other) data are to be analyzed (e.g. data-mined), especially when this is done by third parties. One difference to communications anonymity is that PPDM methods aspire to protect the utility of the anonymized data for analysts (see Bertino et al. 2008 for an overview of utility metrics).

2.3.1 Privacy as hiding/confidentiality as the focus of PPDM

The setup of PPDM as a class of mining methods clearly shows its focus on privacy as confidentiality. PPDM is motivated by privacy as "the right of an entity to be secure from unauthorized disclosure of sensible information that are contained in an electronic repository or that can be derived as aggregate and complex information from data stored in an electronic repository" (Bertino et al. 2008, p. 3).

A look at further definitions of privacy from the PPDM literature illustrates the underlying idea that 'privacy obtains when certain data are hidden'—regardless of whether the data are general or, for example, from social network sites or query logs: For example, Clifton et al. (2004) focus on the "freedom from unauthorized intrusion" and demand "solutions that ensure data will not be released", and they observe that the "disclosure of knowledge about an entity (information about an individual) [is] a potential individual privacy violation", and that the analogous holds for the disclosure of knowledge about sets of data of other entities such as corporations. "[S]et[ting] up the data to protect the privacy of individual users while preserving the global network

properties [...] is typically done through anonymization, a simple procedure in which each individual's 'name'—e.g., e-mail address, phone number, or actual name—is replaced by a random user ID" (Backstrom et al. 2007). Poblete et al. (2007) distinguish a "general adversary [who is] trying to discover any useful information" from a "particular competitor [who] tries to disclose information", where the information is or includes that about a specific website or business.

PPDM methods address the *database inference problem*: "The problem that arises when confidential information can be derived from released data by unauthorized users" (Verykios et al. 2004), and the *objective of PPDM* is to "develop algorithms for modifying the original data in some way, so that the private data and private knowledge remain private even after the mining process" (Verykios et al. 2004). Here, "private data" are the given inputs to the mining process that are supposed to remain confidential, and "private knowledge" is that part of the knowledge inferred during mining that is supposed to remain confidential. Put more simply, the objective is to learn what we are allowed to learn from data that we are not allowed to see. To achieve this, PPDM methods must solve the *(data mining/publishing) anonymization problem*: to "produce an anonymous [table] that satisfies a given privacy requirement determined by the chosen privacy model and to retain as much data utility as possible" (Fung et al. 2010, p. 5). Key concepts of PPDM are defined in Clifton et al. (2004). Following security-research terminology, adversaries are also called "attackers" who perform an *attack*: a "sequence of activities that result in the disclosure of confidential information" (Poblete et al. 2010).

A closely related area is *privacy-preserving data publishing (PPDP)*. This takes a wider look at a setting often found in PPDM: the publishing, by a data publisher (e.g., a hospital), of—at least in parts—sensitive information on data subjects (e.g., patients), for an audience of data recipients. The latter are in general not known a priori, could be ill-intentioned, and may perform arbitrary data mining tasks. The *objective of PPDP* is then that "access to published data should not enable the attacker to learn anything extra about any target victim compared to no access to the database, even with the presence of any attacker's background knowledge obtained from other sources" (Dalenius 1977). Due to the impossibility of this in the face of arbitrary background knowledge (Dwork 2006), one usually assumes limited and specific background knowledge of the attacker, or requires that probabilistically, the posterior beliefs after looking at the published data are not much different from the prior beliefs ("uninformative principle", Machanavajjhala et al. 2006).[1] The same idea lies behind differential privacy that "ensures that the removal or addition of a single database item does not (substantially) affect the outcome of any analysis" (Dwork 2008). Informally, this could be considered to not hide data, but to avoid that information can be gleaned from them.

A wide literature exists on PPDM and PPDP, which cannot be covered here. For details of specific algorithms and method groups, see also Aggarwal and Yu (2008a), Ciriani et al. (2008) (PPDM), Zhao et al. (2009) (PPDP), Zhou et al. (2008), and Wu et al. (2010) (graphs/networks).

[1] Another difference to PPDM is that sometimes in PPDP, truthfulness at the record level is important: each modified record must still correspond to an entity of the original database. This precludes certain forms of data modification used in PPDM.

Taking a closer look at these data-centric definitions of privacy, one sees that alongside the focus on confidentiality (not seeing data, not learning about an entity, protection from disclosure), there is also the recognition that data need not be kept confidential in every case, but could be disclosed as long as someone entitled to do so "decides" or "authorizes" disclosure/communication. This someone is often the data subject, but may also be unspecified (cf. the definition from Bertino et al. 2008)—we will return to this question in Sect. 2.8. This move away from unconditional hiding, or "privacy-as-confidentiality", leads to the notion of privacy as control, to be discussed next.

2.4 Privacy as control: informational self-determination

A wider notion of privacy, appearing in many legal codifications, defines the term not only as a matter of concealment of personal information, but also as the ability to control what happens with it. This notion does not call for strict data parsimony. One reason is that the revelation of data is necessary and beneficial under many circumstances—and that control may help to prevent abuses of data collected in this way.

This idea is expressed in Westin's (1970) definition of *(data) privacy*: "the right of the individual to decide what information about himself should be communicated to others and under what circumstances" and in the term *informational self-determination* first used in a German constitutional ruling relating to personal information collected during the 1983 census, and highly influential in Europe and beyond since then: "the protection of the individual against unlimited collection, storage, use and disclosure of his/her personal data is encompassed by the general personality rights of the [German Constitution]. This basic right warrants in this respect the capacity of the individual to determine in principle the disclosure and use of his/her personal data. Limitations to this informational self-determination are allowed only in case of overriding public interest" (BVerfG 1983).

Informational self-determination is also expressed in international guidelines for data protection such as the OECD's Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (OECD 1980), the Fair Information Practices (FIP) notice, choice, access, and security (FTC 2000), or the principles of the EU Data Protection Directives (EU 1995, 2002). As an example, consider the principles set up in the OECD Guidelines: collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, and accountability.

In sociological accounts, privacy as control is closely tied to the ability to separate identities, which allows individuals to selectively employ revelation and concealment to facilitate their social performances and relationships (a second type of privacy in Phillips 2004). Computer science has applied these ideas in systems for identity management and access control.

Although informational self-determination principles are desirable, relying only on them when building systems can be misleading. Collection limitation in one system does not protect against the aggregation of those data in many systems. Openness may be overwhelming in current ubiquitous-technology environments, where the numbers

of data controllers increase exponentially. A user may be overwhelmed by the difficulties of individual participation and unable to judge the risk of revealing information or using automated agents for such decision-making. Even if all these principles were implemented, it would be very difficult to identify violations. In the case of trusted parties, system security violations (i.e. hacked systems), design failures (i.e. information leakages), or the linking of different sources of safely released data may cause unwanted release of information. Furthermore, data protection focuses on individual and identifiable data. It therefore offers little protection with respect to aggregation of anonymized data, profiling based on correlations and patterns found in this aggregated data, and the consequent desirable or undesirable discriminations. Finally, privacy as control is an abstract concept that does not consider how people actually do and want to construct their identities. This is the topic of privacy as practice, to which we turn next.

2.5 Privacy as practice: identity construction

Despite interesting research results in the area of privacy preserving methods and tools, individuals are confronted every day with the collection of massive amounts of data about them. This has many reasons. Some services require identification (e.g. hospitals or employment situations). Commercial interests in collecting information often extend beyond such contexts. Popular and usable privacy enhancing technologies are rare to non-existing. Surveillance technologies collect information on a mass level without consent. Furthermore, people often simply desire to reveal information about themselves with their names etc. By privacy as practice, we refer to the definition of the right to privacy as the freedom from unreasonable constraints on the construction of one's own identity,[2] which includes the abilities to strategically reveal or conceal data. This approach requires domain-specific and sociological analysis of users' and communities' information revelation and concealment needs as in the examples given in Hancock et al. (2009), Lipford et al. (2008), and Binder et al. (2009). The diversity of user concerns that is stressed here is often not emphasized in privacy-as-confidentiality and privacy-as-control approaches.

Privacy as practice demands the possibility to intervene in the flows of existing data and the re-negotiation of boundaries with respect to collected data. These two activities rest on, but extend the idea of privacy as informational self-determination: they demand transparency with respect to aggregated data sets and the analysis methods and decisions applied to them. In this sense, these approaches define privacy not only as a right, but also as a public good (Hildebrandt 2008).

Sociologists have investigated the idea that privacy is (social) practice from various viewpoints. Phillips (2004) distinguishes two further types of privacy in addition to the above-mentioned right to be let alone and the possibility of separating identities. The third type is the construction of the public/private divide. This distinction concerns the social negotiation of what remains private (i.e. silent and out of the public discourse) and what becomes public. For instance, the decision by individuals to

[2] The phrasing is due to (Agre and Rotenberg, 2001, p. 7), see also Hildebrandt (2006).

Fig. 1 What revealing search data reveals (adapted from Barbaro and Zeller 2006)

keep their voting choices private is generally accepted today; while in the case of domestic violence, interest groups and individuals have successfully lobbied over the past decades to redefine the "domestic" as a public issue. The fourth type in Phillips (2004) is the protection from surveillance. Here, surveillance refers to the creation and managing of social knowledge about population groups. This kind of privacy can easily be violated if individual observations are collated and used for statistical classification. When applied to individuals, such classifications make statements about their (non)compliance with norms, their belonging to groups with given properties and valuations, etc. Arguably, such processes may pose unreasonable constraints on the construction of identities. Market segmentation is an example of the classification of population groups. In computer science accounts of privacy in networks and in particular social network sites (SNS), similar ideas have been expressed for example by Palen and Dourish (2003) and boyd and Ellison (2007).

These definitions emphasize that confidentiality and individual control are part of privacy, but not all. Privacy includes strategic concealment, but also revelation of information in different contexts, and these decisions are based on—and part of—a process of collective negotiation. Tools should therefore support data concealment and revelation to help individuals practice privacy individually and collectively (see Sect. 3.4).

2.6 Example

These different faces of privacy are neither exhaustive nor mutually exclusive. For example, Fig. 1 shows an excerpt of the search history of the re-identified AOL searcher that the NYT published, along with her explanations of what this search query meant (Barbaro and Zeller 2006). The examples illustrate the different motivations or kinds

of privacy very well: the searcher's different roles in life, her intimate private sphere, the wish to be let alone (expressed in her overall reaction: "I had no idea somebody was looking over my shoulder"), and implicitly also an objection to being profiled.

2.7 Privacy and identification

The legal formulations as well as the AOL case show that (re-)identification of people is a fundamental issue in privacy. However, what does identification really mean? One can distinguish three types (Phillips 2004, p. 700): (a) *lexical identification* links a name to an entity; (b) *indexical identification* points to an entity; (c) *descriptive identification* assigns attributes to an entity (so that it can be compared to and distinguished from other entities).

Examples of lexical identifiers are a person's proper name, but also a pseudonym or an anonymized name (e.g., the hash value of the proper name). Examples of descriptive identifiers are "people who bought book X", "a person who issued the queries $q_1, \ldots, q_n$", "heavy users of music download sites" and "online pirates". Indexical identification is identification that allows one to contact or otherwise (usually physically) reach that person.

Many identifiers are of multiple types. For example, a proper name (lexical) often makes it possible to reach a person (indexical), as the visit by the NYT reporters to the home of the AOL searcher showed, who was thought to be only lexically identified by the anonymized number. The IP number or cookie carried through one or several Web sessions (lexical) can make it possible to reach (indexical) an 'only' descriptively identified individual. The description of a person by her current geographic coordinates (descriptive) can enable a suitably-placed third person to reach the located person (indexical).

Laws, by their definition of personal data or personally-identifiable information, focus on indexical identification. They recognize that proper names and a few other lexical identifiers allow third parties to construct indexical identifiers, but they generally ignore the reconstructability from descriptive identifiers (such as search queries). Beyond this shortcoming concerning individuals, privacy laws generally do not preclude the building of profiles, i.e. the descriptive identification of (real or fictitious) population groups. However, as seen in the section on privacy as practice, such descriptive identifications may well be privacy-relevant.

2.8 Whose privacy?

The answer to this question appears to be straightforward: Privacy protection is the protection of individuals, natural persons, as reflected in the many different definitions given above. However, a closer look at real life as well as at the PPDM (etc.) literature reveals that other entities too may be concerned about confidentiality, control and practices.

Consider, for example, the problem setting in the original "privacy protecting data mining" paper (Agrawal and Srikant 2000, p. 439): "Since the primary task in data mining is the development of models about aggregated data, can we develop accurate

models without access to precise information in individual data records?" This question is as relevant for a hospital that is concerned about the welfare of its patients whose data it publishes, as it is for a company that outsources data analysis procedures and is concerned about the trade secrets in its data. The latter case has been explicitly addressed for example by Boyens and Fischmann (2003). In this setting, the agent that wants to protect its privacy is the data owner; to the extent that the data are about people (data subjects), their privacy may also be affected as a side effect.

This idea has been elaborated in more detail by Domingo-Ferrer (2007), who asked *whose* privacy is to be protected, what the intended action and the unwanted breach are, and which group of computational techniques address this problem. He distinguished: (a) Data subject[3] privacy (DSP): a data owner wants to make a database available to third parties, but wants to avoid that data subjects are re-identified. Solutions mainly come from statistical disclosure control. (b) Data owner privacy (DOP): two or more data owners want to compute something and agree to the results being shared, but do not want to disclose the data. Solutions mainly come from PPDM and from secure multi-party computation. "Hippocratic databases" (Agrawal et al. 2002), negative databases (Danezis et al. 2007), and many recent forms of PPDM are a solution approach to address both DSP and DOP. (c) Data user privacy (DUP): a user wants to query a database, but does not want to be profiled based on these queries. Solutions mainly come from private information retrieval (cf. Chor et al. 1998; Camenisch et al. 2009). Importantly, Domingo-Ferrer (2007) showed that each of (a), (b) and (c) can be protected while the other ones are not. Thus, for example, some forms of "privacy protection" may in fact only protect DOP, but violate DSP and/or DUP.

Aïmeur et al. (2008) investigated what may be considered a special case of (b) coupled with (a): In a recommender system, both customers and merchants are in principle both data subjects and data owners, who wish to analyse these data (preferences in the case of customers, product-catalogue information in the case of merchants) to reach a common goal. The privacy problem is that customers should be able to keep private their personal information, including their buying preferences, and they should not be tracked against their will. The commercial interests of merchants should also be protected by allowing them to make accurate recommendations without revealing legitimately compiled valuable information to third parties. The authors propose an architecture relying on a trusted third party to protect both privacy interests.

Poblete et al. (2007) investigated "website privacy" (renamed "business confidentiality" in Poblete et al. 2010), a specific form of DSP that occurs in, for example, the AOL log. The problem that started the "AOL scandal" was a form of DSP that concerns an individual searcher, whose (anonymized) identity could be recovered by using the query data linked to this searcher. Poblete et al. (2007) showed that the search results may be equally problematic: If the search result is a website, then the queries (aggregated over all searchers) that lead to this website may contain rich information

[3] Called "respondent" in Domingo-Ferrer (2007)

on customer-site interactions and therefore material for competitive market analysis or even industrial espionage. In the AOL log, these search results were partially masked (only the domain was given), but as the authors showed, even fully anonymized websites would be easy prey to an intersection attack that uses the live search engine that released the logs. An interesting feature of this problem is the plausible assumption of a well-informed attacker: a competitor who knows the market, the industry structure, the search terms, etc. very well and therefore has extensive background knowledge for de-anonymization. The authors describe various attacks and anonymization methods to counter them.

These associations of data with different stakeholders and their respective "privacies" makes conflicts unavoidable, for example when one actor's (e.g. the DS's) privacy is another actor's (e.g. the DO's) business intelligence. As a consideration particularly of privacy as practice shows, it may often be the case that a piece of data may be considered private by some actors but required to be public by others, and that this relationship may change over the course of processing. For example, individuals consider their consumption patterns as private data, and a company considers marketing intelligence mined from such patterns its private property, which in turn is requested to be public by the individuals. States consider that state secrets on military operations, the secret service, etc. should not be revealed, citing reasons such as civil security and safety.

Conflicts will arise when people believe things should be public (and make them public) that businesses or states consider should be kept secret. Examples include pointing out deficiencies in the care provided in a foster home or uploading classified military or diplomatic documents onto a whistleblower Web site. If the whistleblower is or becomes non-anonymous, such behaviour may lead to job dismissal or prison sentences (in the first example, the whistleblower brought a criminal complaint against her employer; in the second, a soldier was accused of being the whistleblower based on indirect evidence). As these examples and their public discussion shows, legal as well as political assessments of the admissibility of such privacy/secrecy violations differ (European Court of Human Rights 2011; Nakashima 2011). Other examples include the disclosure of negative information about a company's financial standing that may prevent this company from obtaining further credit and forcing it into bankruptcy—but at the same time protect potential lenders from risky new investments in a failing company.

We would like to emphasize the need for a careful use of the term "privacy". From a computer-science standpoint, it may well be argued that "the difference between such corporate privacy issues and individual privacy is not that significant" (Clifton et al. 2004). However, from other standpoints such as law or philosophy, the privacy of natural persons is different from business secrets and state secrets. For example, unlike persons, neither businesses nor states can have personality rights that the state must protect and that have priority over other rights (such as property rights). For a critique of trade-secret law, see Bone (1998), for a discussion of state secrets and the case of Wikileaks, see Sagar (2007, 2011).

Taken together, these observations call for paying detailed attention to whose privacy/secrecy a given technological solution threatens or protects, what this means for background knowledge and attacker models, and what interdependencies may arise.

2.9 What is special about the Web?

Web data and its analysis are a particularly interesting area for investigating privacy for a number of reasons. First, Web information has rich relational structures: Entities with an already interesting inner structure (such as people, institutions, documents, …) are linked in multiple ways that may themselves be further annotated. While Web data continues to include standard tabular data, increasingly also more structured forms are used. Graph and network structures come to mind first: people are "friends of" other people in social network sites, they "rate" movies and "issue" queries, queries "lead to" search results, etc. However, often these relations are in fact (at least) ternary or otherwise associated with further information: what did the friends exchange, how is a movie rated, etc. A relation like the ternary searched-and-clicked-on(user,query,URL) may in fact be even richer: a query log can reveal statistical aggregations that show the strength of the association between a query and a URL in terms of the number of people who issued this query and then clicked on the URL; and the ranking of each URL relative to a query may be recorded.

This very richness of the data, coupled with the strength of methods for the mining of graphs, hypergraphs and their projections, makes the data a potential source of many interesting findings. On the other hand, the large information content may be overlooked and also make such data vulnerable to a large number of attacks. One example is the observation that query log data not only contain information about users (mostly derived from the projection of searched-and-clicked-on onto users and queries), but also about businesses or other typical entities described by Web materials (projection onto queries and URLs) (Poblete et al. 2010). Rich relational structure means breadth and depth: for example, query logs show not only what users do, but more importantly also what they want, and their analysis may also yield competitive intelligence such as which keywords work and which pages attract traffic and paying customers.

A second key characteristic of the Web is its participatory architecture. This can make it easy to create the background or external knowledge that can be integrated with other data. For example, one can easily create accounts and "friendship links" in an online social network, or create entries in a search engine's log by querying the live search engine. These data-creating activities can then be used for intersection attacks (see Sect. 3.5 for more details).

The Web's participatory architecture is of course not only a threat to privacy, but also an opportunity for many people to utilize a previously unknown range of information and services, and it has the potential to democratize lives and societies. One development associated with these opportunities is that people perform more and more of their lives and its different roles on the Web. This leads to a third key characteristic of the Web: a hitherto impossible coverage of life in terms of data. One disadvantage of this breadth and depth is the risk of panoptic surveillance.

Summary and the problem scope of "privacy-protecting knowledge discovery" In sum, privacy is not only about *hiding* certain information, but also about *controlling* information and its uses (e.g., by constructing different identities), and is finally a dynamic *practice* involving negotiations and tradeoffs between hiding and disclosing/sharing.

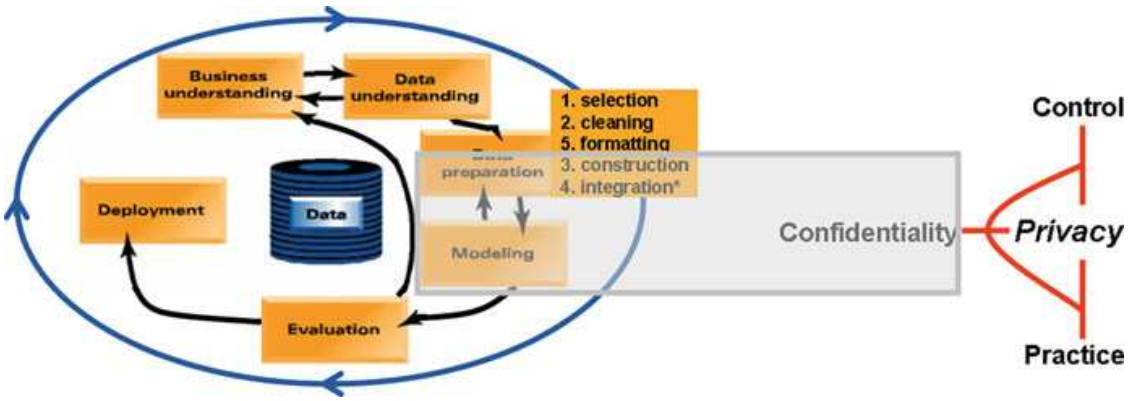


Fig. 2 Privacy-preserving data mining and privacy-preserving data publishing address only a part of the effects of knowledge discovery on different kinds of privacy (shaded box in the middle). The aim of this paper is to outline the effects of all KD phases on all kinds of privacy (The CRISP-DM diagram on the left-hand side is adapted from CRISP-DM (2000))

This leads to a broadening of the scope of the second important term: data mining (in its wider sense of knowledge discovery) is not only about association-rule discovery, decision-tree learning, or similar modelling steps. It extends over all phases of knowledge discovery (KD). This will be explored in the following section.

3 The KD phases and privacy: threats, opportunities, and solution approaches

The term "data mining" is used in different meanings, ranging from the application of a certain algorithm to some data to find, e.g., association rules, to the whole knowledge-discovery cycle. In addition, the knowledge-discovery cycle itself has been defined in more or less comprehensive ways. The main point of this article is to argue that the whole process of knowledge discovery must be investigated in order to understand the ramifications for privacy better.

We understand data mining as extending over all phases of knowledge discovery (KD) as identified in the CRISP-DM model (Shearer 2000; CRISP-DM 2000): *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, and *deployment*. Each one of them consists of a number of sub-phases, see the left-hand side of Fig. 2 for the main phases and the sub-phases of data preparation.

3.1 Aims and non-aims of this section and overview

In the following, we first explain the motivation for choosing CRISP-DM and important consequences of this choice (Sect. 3.2). In the remaining subsections, we then explain how the activities of the different phases may pose threats to, but may also offer opportunities for, the three kinds of privacy. The *threats* are partly familiar, while others only become obvious in the light of the new wider notion of privacy. The *opportunities* will arise from two reasons. The first is that new technologies may afford a reversal of roles (see "whose privacy") with interesting opportunities for knowledge and power shifts between actors—a typical case of privacy-as-practice. We will focus

on the perspective of typical end users—which will inevitably lead to these opportunities being advantageous for them, but possibly threatening for others. The second reason for opportunities is that technologies may change the way a KD step is performed, thereby relieving some privacy concerns. For each phase, we then describe key technological *solution approach(es)* to ward off threats and enable opportunities. It will be seen that approaches from within PPDM/PPDP are only some of these solutions; we will argue that other "privacy-enhancing technologies" (PETs) play an equally important role.

To support also non-sequential reading, we outline the (nearly) factorial design of this section: In the spirit of CRISP-DM, we treat the phases *business understanding* to *data preparation* as including their own respective applications, i.e. we assume that for example the integration of two data sources is not only modelled, but also performed. For these phases, we will show how each phase poses threats and/or opportunities for all three types of privacy. We will concentrate on relevant subphases of data understanding and preparation and order them in a slightly non-standard way, as explained at the beginning of Sect. 3.4. We make a sharper distinction between the obtaining of a data model and its application by following CRISP-DM in separating the phase *modelling* from the *deployment* of the model. This perspective led us to conclude that deployment poses threats and/or opportunities for all three types of privacy, while modelling mainly affects privacy-as-practice. Finally, the *evaluation* phase will be described as a meta-activity with only indirect impact. Throughout, examples from various areas related to Web data (mining) and other areas of KD will be used.

3.2 Why CRISP-DM?

The aim of CRISP-DM was to create (a) an industry- and tool-neutral process model that (b) views the analysis of data in its full application context. (b) can be understood if one compares CRISP-DM with the "KDD process" (Fayyad 1996). The first three phases of this process, selection, pre-processing and transformation, correspond to CRISP-DM's data understanding and data preparation. Its fourth phase, data mining, corresponds to modelling. Its last phase, interpretation/evaluation, corresponds to evaluation. The five phases of SAS's SEMMA[4] correspond to those of the KDD process (sample, explore, modify, model, assessment) and thus also to the four middle phases of CRISP-DM (see Azevedo and Santos (2008)). By means of its first and its last phase, CRISP-DM is more comprehensive than these models. CRISP-DM is also being used as a comprehensive reference model by proponents of more specific standards such as Java DM, a Java API for developing data mining applications and tools. This illustrates property (a) of CRISP-DM: As Hornick et al. (2007) show, Java DM supports the four middle phases of CRISP-DM in various ways.

One activity that remains implicit in this process model is that of data storage and data management. Data management in particular must be done in a secure way in order to avoid certain privacy violations. Unfortunately, security breaches caused by

[4] http://www.sas.com/offices/europe/uk/technologies/analytics/datamining/miner/semma.html

stolen laptops, data carriers erroneously mailed to the wrong people or put into the household garbage, deficient authentification routines or leaked passwords remain a major source of privacy breaches. From the viewpoint of KD, however, secure data storage and management are an external issue. Therefore, and for reasons of space, we abstract from security questions in this article and refer the reader to Anderson (2008) for an overview.

3.3 Business understanding

Business understanding consists of determining business objectives, assessing the situation, determining the data mining goals and producing the project plan. In the following, we also subsume the setting of the business model under these activities. None of these activities in themselves already violate privacy, but the choice of objectives, goals and plans may imply later business practices that are prone to do so.

Knowledge discovery and data mining may be performed by or on behalf of two types of actors: those who have some other core business model and 'only' perform data mining in order to support it (e.g., a supermarket that analyses data to optimise its store layout), and those whose core business model rests on mining. Since Web-based services tend to fall into the latter category (recommender systems, search engines, …), we concentrate on the latter category. Thus, it is the *choice* of a business model that may affect privacy, and the phase of business understanding will help to understand what effects can occur. (In contrast, the 'mere' understanding of a business model of the first category is likely to have no effects on privacy.)

Threats Consider three examples: (a) the business model of DoubleClick's profile database that was revealed in 2000; (b) the 2004 introduction of Google's Gmail service (www.googlemail.com); and (c) The 2007 Google-DoubleClick merger.

(a) DoubleClick provides Internet ads to buyers of ad space and to sellers of ad space, it counts click-throughs and other indicators of an ad having been seen as a basis for pricing, and it tracks the individual Web users who receive ads served through DoubleClick. After its acquisition of Abacus, a leading provider of specialized consumer information and analysis for the direct marketing industry, in 1999, DoubleClick began to offer an "Intelligent Targeting Service", which allowed marketers to target ads based on a database of about 100 million profiles. This was based on combining surfing records with detailed personal profiles contained in a national marketing database. This kind of business model uses and reinforces 'customer profiles', which may violate privacy-as-practice. In addition, the subsequently necessary data collection and storage may violate privacy-as-confidentiality, and the re-purposing may affect privacy-as-control.

In 2000, a complaint was filed with the US FTC that pointed out the unlawfulness of such purpose-changing re-use of data and omissions in the clear communication of relevant business details and activities to users ("deceptive trade practices") (EPIC 2000). The FTC investigation was closed after DoubleClick offered to make a number of modifications to its site and services and to end the Intelligent Targeting Service.

In sum, it turned out that this business model was not consistent with data protection laws.

(b) Gmail started operations in 2004; its business model is to offer free Webmail with virtually unlimited storage space in return for the user's attention to personalized ads. The ads are personalized by "matching [them] to content".[5] In the search engine, this content consists of the queries the user enters; in Gmail, it consists of the emails she sends or receives. The business model thus rests on a violation of the privacy of correspondence, an instance of privacy-as-confidentiality. This was pointed out by a complaint filed by Privacy International to the privacy and data protection regulators of a number of European countries and the EU (Privacy International 2004). The complaint was largely ignored (Privacy International 2007), and Gmail continues to operate.

(c) Google Search and Gmail are two out of more than ten services (EPIC 2007), most of which link gathered data not only to a user's potentially pseudonymous "Google ID" but also to a user's IP address (which may be a lexical identifier and as such personal data, see Sect. 2.2). The services provide Google with self- and third-party descriptions of the user and with behavioural data; the latter including requests (e.g., YouTube), queries (e.g., Search), and activities relating to content creation (e.g., email sending in Gmail; photo uploading in Picasa). In early 2007, Google announced its plan to acquire DoubleClick. This plan was challenged by privacy advocates. The reasoning in the complaint (EPIC 2007, 2011c) rests on (legally actionable) deceptive and unfair business practices including re-purposing of data shown in the past by the two companies, on the expectation of more of the same, and on an expected scenario of panoptic surveillance (privacy-as-practice). The Google-DoubleClick Merger was allowed by the Federal Trade Commission in late 2007 and the European Commission in early 2008 (EPIC 2011c).

Opportunities All data analysis options have different uses, thus, by a reversal of roles, the integration of identities may also be beneficial for privacy-as-practice. In the case of end users, one example is the use of mining against identity theft, a service previewed by Rowe and Ciravegna (2010). Identity theft may occur when technology supports the separation of identities: person A may be misrepresented and defrauded by someone else who poses as A and creates information on A somewhere on the Web without A's knowledge. The authors propose to monitor the Web presence of a given individual by obtaining background knowledge to support automated disambiguation processes. They generate this background knowledge by exporting data from multiple Web 2.0 platforms as RDF data models and combining these models together for use as seed data. They present two disambiguation techniques. The first uses the semi-supervised machine learning technique of Self-training; the second uses the graph-based technique of Random Walks. The semantics of data support the intrinsic functionalities of these techniques.

A de-separation of identities may in general be something that users want in the sense of privacy-as-practice: when different identities, accumulated over time for

[5] http://mail.google.com/mail/help/intl/en/about.html [13 August 2008]

various reasons, are seen as too burdensome to manage. Schemes that answer this perceived need include OpenID[6] and services such as Google Mail Fetcher[7] that enables users to bundle different email accounts. (However, Gmail also performs its own data mining on its access logs and recommends bundlings of mail addresses that were never requested by the user.)

In addition, business models may offer novel ways of protecting privacy (of all three kinds) or changing privacy in ways that empower individuals or groups (fostering the negotiation of the public-private divide in the sense of privacy-as-practice). Particularly interesting in our context are settings that exceed the corresponding options in the offline world. These are basic schemes aimed at avoiding the generation of data (as opposed to PPDM that first collects and then transforms).

One important kind of such (Internet) business models are portals that allow users to input information anonymously. Due in part to the prevalence of public/private divides, it is often easier for people to externalise sensitive information when this can be done anonymously. Such information may or may not be input to KD processes on top of the mere collection. For example, medical doctors need to report—anonymized—cases of certain diseases like tuberculosis or HIV, from which basic statistics, but also prediction models could be derived. The latter could lead to intervention and ultimately prevention of the disease.[8] In the online world, a recent example is the discussion of errors in medicine and care. A number of "Incident-Reporting-Systems" exist in medicine, including the "Patient Security Information System",[9] where users may, but do not need to, report their role in the incident, the portal "Each error counts"[10] for general practitioners, or "Critical Incidents"[11] for elderly care. Other systems are dedicated to Election Incident Reporting.[12] A recent development whose content-related, political as well as technical details are undergoing controversies and rapid development are general-purpose whistleblower sites.[13]

Solution approaches Anonymization proxies and in particular onion routers can be regarded as technologies underlying a business model that addresses the above-mentioned problems. However, anonymity and (full) traceability cannot coexist. Therefore, current controversies (e.g. Feiler 2008, German Working Group on Data Retention (AK Vorrat) n.d.) deal with the extent to which anonymization proxies conflict with the EU Data Retention Guideline Directive (EU 2006) which requires the comprehensive logging of all communications traffic data, whether they will be required to

6 http://openid.net/

7 http://mail.google.com/support/bin/answer.py?ctx=gmail&hl=en&answer=21288

8 In Germany, this is governed by the 'Law for protection from infection' (Infektionsschutzgesetz 2001). For a review of the difficulties of enforcing similar registration regulations worldwide, see Gostin (2004) on the WHO International Health Regulations.

9 www.pasis.de

10 www.jeder-fehler-zaehlt.de

11 www.kritische-ereignisse.de

12 www.voteprotect.org

13 The most famous of these sites is www.wikileaks.org. Due to the fast developments and frequent re-hostings in this area, it was decided not to list further URLs in this journal article.

keep comprehensive logs of the mappings they effected, or whether their continued operation may enable people to circumvent data retention.

The technical foundations of anonymization will be discussed in relation to data collection (Sect. 3.4). The challenges facing anonymization will be discussed in relation to anonymization's main enemy: data integration (Sect. 3.5).

3.4 Data understanding I: data collection

The data understanding phase of CRISP-DM comprises *data collection*. Implicit in the KD model is that data, once collected, will also be stored.

This is followed by the sub-phases *describing data*, *exploring data*, and *verify[ing] data quality*. In these subphases, the data could be analysed to test whether they satisfy properties desirable for privacy—or not, in which case transformations may become necessary. Many of these properties have been described in relation to problems that result from data integration. We therefore interleave the descriptions of the data understanding and data preparation phases as follows: after collection, often an integration with other datasets occurs. The result of this should be explored and then, if necessary, transformed by constructive operations.

Threats Activities for collecting data may, regardless of what later happens to those data, be very intrusive, for example because they involve physical intrusion into a private or identity-separated sphere for setting up the collection equipment (violate privacy-as-confidentiality and privacy-as-control) and create a climate of panoptic surveillance (violate privacy-as-practice). In the same way, storing data may, regardless of what later happens to those data, be intrusive.

It may be argued that factors that support the disclosure of data on the Web are in themselves a threat. One factor is the widely lamented ignorance of (not only) teenagers that data put on SNSs are subject to overly loose privacy policies (thus, people perform privacy-as-control but overlook the consequences), cf. the collection of problems observed in Facebook's practices and references at EPIC (2011a). Another factor (which may lie behind the first) is the difference between people's self-professed privacy attitudes and behaviour (Berendt et al. 2005; Acquisti and Gross 2006). A third factor is that many users may just not understand the consequences of data releases.

Opportunities Activities for collecting data may also be designed in a privacy-protecting way and serve to respect privacy-as-practice by helping to shift the public/private boundary. An example is the anonymized data collection described in Sect. 3.3.

Solution approaches Today's Privacy-Enhancing Technologies (PETs) mainly focus on limiting data collection and/or on making the process understandable.

PETs are tools and mechanisms which, when integrated or used in conjunction with online services or applications, allow users to protect their data provided to and handled by such services or applications. They provide encryption (to prevent eavesdroppers from seeing the content of transferred information), and/or anonymization/pseudonymization (to prevent the identity of the communication partners from being released).

A current overview of the field is given in Goldberg (2007); links can be found at EPIC (n.d.).

Three types of PETs are particular interesting for our purposes: anonymizers, identity management tools, and privacy negotiation techniques that combine technological and industry self-regulation ideas (e.g. P3P (W3C 2006a), which however does not enforce privacy and is therefore not treated here in more detail).

Anonymization proxies work between a user and a Web site; they replace this user's identity (IP address) by an anonymous marker. Onion routers such as Tor[14] break the linkability between a requesting person's (even pseudonymous) identity and the sequence of requests. Onion routers chain several anonymization proxies (to avoid the need of *one* trusted third party that is highly vulnerable), and they make the different user actions in one session seem to come from different places. For overviews of anonymization technology, see (Danezis and Diaz 2008; Acquisti et al. 2011).

As an alternative to anonymity, *pseudonymity* is often preferred (for example, to be able to generate a persistent identity in a certain context). A subject is pseudonymous if a pseudonym (= an identifier of a subject other than one of the subject's real names) is used as identifier instead of one of its real names. *Identity management* systems tools (IDMS) allow users to manage different pseudonyms, providing for a separation of identities and thus providing more privacy-as-control (Pfitzmann and Hansen 2006–2010). There are different types of identity management systems, but here we refer to mechanisms that support separation of context-dependent virtual identities represented by pseudonyms of varying strength. IDMS allow individuals to establish and secure identities, describe those identities using attributes, follow the activities of their identities, and delete identities. They are often based on credentials and access control methods. SNS are interesting as identity management systems with access control mechanisms based on relationships. A simple form of such access control defines access based on the path distance from the node that owns the data. Popular access models in current SNS comprise "friends" (only nodes one hop away from the data owner may see that profile) and "friends-of-friends" (only nodes at most two hops away may see the profile), or in some cases, friends at a longer path length (Bonneau and Preibusch 2009). Overviews of current work related to identity management in social networks and detailed descriptions of different access-control models that protect a user's profile or relationships are given by Carminati et al. (2010) and Acquisti et al. (2011).

A close look at Web-based SNSs shows that these are one of the first massively adopted IDMSs. SNSs provide numerous and rich examples of user-provider-negotiated privacy practices (boyd and Ellison 2007). This can be anywhere from the privacy settings which have evolved immensely in the last years, via the introduction of usable and integrated privacy policies, to the introduction of some simple forms of "privacy mirrors" or privacy awareness tools as a standard feature, etc. Thus, SNSs have arguably become privileged spaces for privacy-as-practice.

Privacy awareness tools aim at fostering understanding and reflection: For example, Lederer et al. (2004) suggest improving privacy sensitivity in systems through

[14] www.torproject.org

feedback that enhances users' understanding of the privacy implications of their system use. This can be coupled with control mechanisms that allow users to conduct socially meaningful actions through them. These ideas have led to suggestions like the identityMirror (Liu et al. 2006) which learns and visualises a dynamic model of the user's identity and tastes. A similar approach is suggested in the concept of privacy mirrors (Nguyen and Mynatt 2002). The authors criticize purely technical privacy preservation solutions that do not take the social and physical environments in which the technical systems are embedded into consideration. Making the data visible would make the underlying systems more understandable, enabling users to better shape those socio-technical systems, not only technically, but also socially and physically. A first implementation of privacy mirrors exists in Facebook, through which users can set controls on their profile information and then check how their profile is seen by their friends, but not by non-friends.

The Hansen (2008) proposal for linkage control in identity management systems is a further example of these ideas. The author suggests mechanisms that provide information about collected data to individuals and the general public. These mechanisms include informing users on possible and actual linkages, as well as de-linking options; communicating privacy breaches to individuals concerned; documenting the sources of data and algorithms used by data controllers as well as the recipients of analyzed data; making accessible personal data and also other data suitable to affect individuals; and providing effective tools to intervene in data linkages in order to execute corrections or deletions. For an overview of privacy awareness tools, in particular concerning social networks, see Acquisti et al. (2011).

3.5 Data preparation I: data selection and integration

Data preparation contains the subphases data selection, data cleaning, data construction, data integration, and data formatting. Cleaning is intended to get the 'real information' out of the raw data by removing noise etc., and formatting is a mostly syntactic step. Data selection and data integration, in contrast, are two steps that determine the semantics of what data are analysed or published, and what data may be used as background knowledge in the sense of an attack model.

Threats Given the increasing availability of more and more data, selection and integration become key levers for subsequent KD. Deployment decisions based on matchings of data to purposes can easily become privacy-violating (consider the case in which a pharmaceutical company marketing department sent out, in the mail, free samples of anti-depressant drugs based on addressees' medical histories, see Burton 2002).

The integration of data is highly problematic because—even before "proper data-mining modelling" sets in—it can enable *inferences* towards more information about a person, inferences that can substitute for data collection and therefore have similar impacts on privacy as data collection—but without the data subject's consent or even knowledge of this process.

Technically, important threats are posed by *record linkage*, where a combination of attributes identifies a small set of records (in the worst case 1), such that a join with another data table may add more and sensitive information. Record linkage can lead to an (indexical) re-identification of a previously anonymous record. The matching of common attributes in diverse data sources leads to more information about an instance (such as an individual) and therefore limits the number of people in the real world that this record could be about. This inference problem has been described by Sweeney (2002) who demonstrated how to re-identify 85% of the people in an (easily available) anonymized medical dataset by combining it with an (easily available) voting register: names can reliably be predicted from gender, birth date and ZIP code. From this, Sweeney derived the notion of *k-anonymity*: To be privacy-preserving in the sense of precluding re-identification, a dataset should be such that every attribute-value combination describes at least k instances. Owad (2006) extended this to show re-identification based on a mashup of freely accessible Web data, creating a visual map of the addresses of people who wanted to read a given book next.

Both cases are examples of an inference towards a personal data item that the data subjects or the data holders did not want to disclose in the given context, and in this sense a violation of privacy-as-control easily bordering on a perceived intrusion into private space (privacy-as-confidentiality). Frankowski et al. (2006) used a movie ratings database (they considered releasing) and an open Web forum to show how easy it is to re-identify a rater based on the comments s/he has written in the forum—even if the person used different pseudonyms in the two spaces, or if one of the datasets is anonymized. Whereas Sweeney operated via matching on equal attributes, Frankowski et al. leveraged the sparsity of real people's universe of discourse (who each only talk about a limited set of movies, even in their different identities). This probably also explains why it is relatively easy to de-anonymize search-engine users (see Sect. 2.6): when one searches for "house prices in …", it will usually be in the city where one lives or where one wants to move. Narayanan and Shmatikov (2009) described an extended case of the intersection attack described in Frankowski et al. (2006) and more robust attacks. While the former study concerned the researchers' own site and data[15], the latter concerned a dataset published by Netflix for a competition to improve its recommendations. After the AOL case had put a stop to search-engine log publication, the Narayanan and Shmatikov study contributed to Netflix withdrawing its ratings dataset from the public domain in early 2010 (EPIC 2010b); further consequences for data publication for research remain to be seen.

Most personal data in these studies were standard Web self-profiling and/or usage data: Amazon wishlists, movie ratings, and forum posts. Additional personal data in Owad (2006) were probably gleaned by the mashed-up services from offline data (address/telephone directories).

There are further linkage attacks in addition to record linkage. In *attribute linkage*, one does not necessarily re-identify an individual, but infers sensitive values from the group that the victim belongs to. This can be seen as an instance of profiling and surveillance, with its negative implications on privacy-as-practice. In *table linkage*,

[15] movielens.umn.edu

one infers the presence or absence of a victim's record in a released table, which gives extra information to the extent that the presence in the table is an attribute in itself (e.g., being one in a list of people who visited a racist website). *Probabilistic attacks* are probabilistic variants of the above—here, the inherent uncertainty of probabilistic information and a disregard for the differential costs of misclassification may pose severe threats when the information is reinterpreted as "true" (rather than "more probable") and/or when actions are taken based on it (see Sect. 3.10).

Richly structured relational data present further opportunities for de-anonymization by integration. For example, Backstrom et al. (2007) showed that even from a single anonymized copy of a social network and limited knowledge about some individual profiles, it is possible to learn whether edges exist or not between specific targeted pairs of nodes. The attacker creates fake profiles and relationships to real users (or uses those that he knows) in the social network of interest. By intersecting this known graph structure with a few known end-points (user profiles) with an anonymized copy of the whole network, he can infer the identities of further nodes and thereby effectively de-anonymize the data. Another example are the methods for de-anonymizing a query log described in Poblete et al. (2010). Here, the attacker generates an extra client-side log with queries and URLs by using the search engine of interest "live". By intersecting this known association with anonymized queries and URLs, he can infer the identities of the masked queries and URLs in the query log, thus de-anonymizing it. The de-anonymization of such data may be interesting because it can be used to derive competitive intelligence. Finally, "human error" presents dangers by data integration. The person made responsible by the U.S. Government for the disclosure of state secrets to Wikileaks is said to have talked about this in a chat room, thereby de-anonymizing himself (Domscheit-Berg 2011).

Opportunities Again, opportunities may arise from a reversal of roles. The same techniques of attribute matching for record linkage that were discussed above, are used in a tool that caused considerable attention soon after its launch in 2007: the Wikiscanner[16]. A Wikiscanner allows Web users to de-anonymize Wikipedia edits from unregistered users' entries by matching the IP addresses recorded and displayed alongside these entries against Whois/RIPE databases that map addresses to owners such as corporations. This led to discoveries of many interest-driven edits, such as energy companies deleting critical remarks about nuclear energy (see the overviews on the Wikiscanner homepage). The Wikiscanner has also been used for scientific purposes (Borra 2007).

Interestingly, Wikiscanners have scarcely been discussed as infringing on privacy, possibly because the discovered edits were perceived as attacks on the integrity of Wikipedia and the "perpetrators" as non-deserving of privacy protection. Viewed as a privacy-related tool, the Wikiscanner can foster discussions about privacy-as-practice, in particular regarding the public-private divide.

16 wikiscanner.virgil.gr

Solution approaches Technological solution approaches rest on ascertaining that the data have certain desirable properties or (more often) transforming them so they obtain these properties.

3.6 Data understanding II: data exploration

Exploring the data is to "[a]nalyze properties of interesting attributes in detail (e.g. basic statistics [of] interesting sub-populations) [and] [i]dentify characteristics of sub-populations" (CRISP-DM 2000, p. 45). While in standard (non-privacy-oriented) KD, this is usually linked to the core data mining goals, but addresses them by querying, visualizations, etc., this sub-phase can gain new importance in privacy-oriented KD. Specifically, the data—in particular if several data sets have been integrated—can be tested for relevant privacy properties. The *opportunities* of this stage lie in detecting problematic configurations (e.g. a dataset that looks anonymous, but is vulnerable to simple attacks like those described in Sect. 3.5), the *threats* in using the wrong test (e.g. a property like k-anonymity that was designed for relational tables but cannot capture properties of network/multi-relational data, see below). Note, however, that the properties typically used all relate to anonymization and confidentiality and therefore focus on privacy-as-hiding.

Solution approaches: testing for desirable properties of data Privacy models for record linkage include k-anonymity as described above and (X-Y)-anonymity. Models for attribute linkage include l-diversity (Machanavajjhala et al. 2006), confidence bounding, (X,Y)-privacy, and t-closeness (which improves on l-diversity when the overall distribution of the sensitive attribute is skewed) (Li et al. 2007). Models for table linkage include δ-presence. Similar privacy models have also been described for probabilistic attacks, whose focus lies on whether the attacker would change her probabilistic beliefs on the sensitive information of a victim after accessing the data; the privacy models in this case aim at achieving the uninformative principle (Machanavajjhala et al. 2006). One problem of these properties is that they are insufficient for rich relational structures, cf. for example the problem analyses and solution proposals concerning query logs (Baeza-Yates et al. 2010) or network data (Hay et al. 2010). For simplicity, we will continue to refer to the basic concepts here. For more details of these properties and of how to test datasets for them, see the cited references.

3.7 Data preparation II: data construction

The *construction of data* includes "constructive data preparation operations such as the production of derived attributes, entire new records or transformed values for existing attributes" (CRISP-DM 2000, p. 24). When it comes to privacy, two types of such operations with widely differing consequences are relevant: the construction of new categories that focus on the semantic description of populations, and construction and transformation operations that focus on formal privacy metrics of data.

Threats The construction and naming of new attributes may create controversial psychological or social categories. The intentional or unintentional reification produces a

social category or norm that may be offensive per se and/or lend itself to abuses such as further privacy-relevant activities (privacy-as-practice).

These categories include criteria such as gender or age bracket (Liu and Mihalcea 2007; Hu et al. 2007), and socioeconomic or other categories deemed important by the provider of the personalisation service. For example, many marketing people in Germany work with the category of *DDR-Nostalgiker*, people who are "nostalgic for the former East German State", because their consumption patterns and attitudes differ from those of other target groups [Sinus Sociovision (n.d.)]. Marketing also has a tradition of classifying by ethnicity, skin colour, or sexual orientation (Wardlow 1996). As a fictitious example from the Web usage domain, consider the classification of Web users with certain behavioural properties (such as frequent visits to music portals) as "(potential) music pirates". Even categories that most people would argue exist can be offensive in the wrong context. An example is the "terror risk score" reputedly assigned to all travelers into and from the US (e.g. Pilkington (2006)). This was soon renamed "terrorist score" in many blogs and other Web sources, while the Department of Homeland Security maintains that no such score exists (Department of Homeland Security 2007). For an assessment by privacy advocates, see (EPIC 2010a).

Opportunities Again, a reversal of roles could present interesting opportunities: the creation of categories that point to "good" behaviour. As a fictitious example, consider the creation of a category of "altruists that help others by answering their questions in online forums or social networks".

A different type of opportunity arises from data construction operations that are designed and done in the interest of privacy protection, for example to counter the record-linkage problem mentioned in the previous section. In the following, we will describe such data construction operations as technical solution approaches.

3.7.1 Solution approaches I: anonymization operations

Technological solution approaches rest on four pillars: privacy models (desirable properties of the data to avoid unwanted inferences), anonymization operations, information metrics (to determine whether the results are still useful), and anonymization algorithms (to find a good or optimal operation to satisfy privacy model and information metric). For a detailed overview and references, see (Fung et al. 2010).

Fundamental privacy models have been described briefly in Sect. 3.6; here, we focus on the data-preparing anonymization operations. Anonymization operations comprise generalization, suppression, anatomization, permutation, and perturbation. Generalization and suppression replace values of specific description, typically the quasi-identifier attributes that are used for record linkage, with less specific description. Anatomization and permutation remove the correlation between quasi-identifier attributes (QID) and sensitive attributes by grouping and shuffling sensitive values in a group with identical QID values. They thereby address the attribute-linkage problem. (For example, they achieve l-diversity such that there are at least l distinct values for the sensitive attribute in each group with identical quasi-identifier attributes.) Perturbation distorts the data by adding noise, aggregating values, swapping values, or generating synthetic data based on some statistical properties of the original data. For

anonymization operations on graph/network and on log data, see (Hay et al. 2010; Zhou et al. 2008; Wu et al. 2010; Baeza-Yates et al. 2010).

Frankowski et al. (2006) give an example of a simple perturbation transformation that can be done by the data subjects themselves: adding noise by inserting popular items into one of the data sources (see also Berkovsky et al. 2007). However, this may amount to a suppression of "extremist" positions both in ratings and in discourse, i.e. silence certain positions and in effect push certain contents into a "private" realm in the sense of privacy-as-practice and curtailing public discourse. It should also be kept in mind that any predictability in the injected noise would make this operation vulnerable to a counteracting data-cleaning operation. A similar principle is behind services such as TrackMeNot[17] that acts as a proxy between a user and a search engine and injects noisy queries to obfuscate the information. PETs for identity management can help to keep identities separate and make integration more difficult.

However, while these technological approaches may be correct in the modelled settings, they are hampered by the impossibility of finding a general solution when attackers' background knowledge can be arbitrary (Dwork 2006) and the increasing number of findings that show the possibility of unwanted inferences with little knowledge and/or the large loss in result utility when effective anonymization is performed (e.g. Narayanan and Shmatikov 2009). Overviews of these problems for the relational and network cases, respectively, can be found in Domingo-Ferrer and Torra (2008) and Hay et al. (2010).

It may therefore be instructive to recall the more encompassing notions of legal regulations and associated tools that may at least make processes more transparent. If a data holder wants to select data, he is bound by the consent of the data subject that this data may be used for this purpose. Purpose/use limitation in this sense as a legal or quasi-legal requirement can be supported by tools like P3P that help manage the relation between data and purposes. We have proposed P3P extensions that take the inference problem into account, for Web usage data/business analytics and social network data (Berendt et al. 2008; Preibusch et al. 2007). The P3P language extension proposed there is coupled with a logic that blocks certain usages of data that would (indirectly via inferencing) violate stated purpose limitations. This is embedded in a hosted Web-based service for business analytics designed to help data holders avoid violating the purpose limitation principle. However, tools like P3P can only support, but not enforce compliance; and the database inference problem is not yet captured well in laws. Architectural principles with the same goal were put forward in the "hippocratic databases" framework (Agrawal et al. 2002).

Data distribution (explained in the next section) can also be regarded as a technical solution approach against undesired data integration: If there is no storage of personal data outside of the individual client computers (and these clients are secure), then it cannot be combined with anything else.

[17] http://cs.nyu.edu/trackmenot/

3.7.2 Solution approaches II: PPDM

The use of data transformation/construction for preserving privacy in the sense of confidentiality is one of the key techniques of PPDM: "The main objective in privacy preserving data mining is to develop algorithms for modifying the original data in some way, so that the private data and private knowledge remain private even after the mining process" (Verykios et al. 2004). PPDM techniques can be classified into data distribution, data modification, and data or rule hiding (Verykios et al. 2004).

Data distribution means a decentralized holding of the data. (It thus corresponds to a very specific meaning of data construction: the construction of decentralized data storage nodes that cooperate during mining.) It protects privacy by obviating the need, for the KD application, to collect these data from the individuals and the need to store them, thus relieving the threats outlined above under "collection and storage". Algorithms on these data usually also work in a decentralized way. This is a form of secure multi-party computation, developed in the cryptographic community also under the name of "privacy-preserving data mining" (Lindell and Pinkas 2000). An example is Canny's (2002) proposal for a distributed P2P collaborative-filtering system where all ranking data reside with the respective users.

Data modification comprises aggregation/generalization or merging into coarser categories (e.g., going from 5-digit ZIP codes to their first 3 digits), perturbation or blocking of attribute values, swapping values of individual records, and sampling. Data modification thus hides values or makes them more difficult to associate with an individual. Reconstruction-based techniques then build models (such as classifiers) from the modified data—models that contain the same aggregate information as if they had been learned from the original data. A data-modification method specifically addressing the AOL problem mentioned above was presented by Korolova et al. (2009).

Data or rule hiding hides data or patterns that are considered sensitive, by pushing their support below a threshold. Since rule hiding is tied more to modelling than to data construction/modification, we will discuss an example of this in the next section.

One general problem is that most PPDM methods have been developed for single data mining algorithms viewed in isolation, such as decision tree inducers, association rule mining algorithms, clustering algorithms, rough sets and Bayesian networks (Verykios et al. 2004; Fung et al. 2010). Thus, at present the application of a specific PPDM procedure during data preparation can either (a) not guarantee the privacy of data subjects if the full range of modelling options is tried on the data, or (b) prevent certain, potentially desirable, modelling options from being applicable. (For example, in Canny's 2002 system, users will keep their ratings private, but they may not be able to derive a clustering model of user types.)

3.8 Modelling

The object of KD is to find interesting patterns; these are identified in the modelling phase. Patterns describe regularities, i.e. they are global characterizations of the modelled set of instances (e.g., a clustering of navigation paths) or local characteriza-

tions of a—sufficiently large—subset of all instances (e.g., an association rule linking highly rated items). Modelling affects individual instances (e.g., individual people) in a way that is mediated via these groups of instances: An individual may be regarded as part of a group (e.g., a market segment), or a property may be predicted (e.g., the interest in buying some item).

Threats KD result patterns may be examples of the unwished-for social categories and norms discussed under privacy-as-practice.

This may also have implications on the public-private divide: "[A system in which individuals and groups determine which description best fits them] also addresses the second sense of privacy—that of public participation in the definition of the public/private divide. One of the most insidious aspects of market profiling is that the social models thus produced are private property [e.g., trade secrets]. …When this private social modeling is combined with the private persuasive techniques of targeted marketing, the result is an anti-democratic […] process of social shaping" (Phillips 2004, p. 703).

This problem may persist when PPDM is applied: PPDM is done under the constraint that the output of mining (e.g., a set of association rules, a classification model) should be unaffected or affected as little as possible. This results from a focus on individuals and on avoiding possibilities of re-identifying those individuals, as opposed to investigating groups and their description and the re-application of such descriptions to individuals. In addition, PPDM methods that make one data mining method privacy-preserving may fail when a different method is applied.

Yet, the modelling phase per se—in which patterns are not yet applied—affects, we believe, mainly privacy-as-practice: because of the perception of the *possibility* of knowledge discovery, because of the perception of the possibility of applying this knowledge, people may change their behaviours, for example hiding more than they normally would.

Opportunities KD result patterns may also have a liberating force, such as associating "shameful" attributes with information deemed positive or otherwise relevant. Such findings may encourage people to start talking about their previously hidden properties, shifting the public-private boundary of privacy, and engendering social change.

This most often is a gradual societal process, such as the development, since the 1970s, that males talk (more often) about emotional topics and "weaknesses" in a societal climate in which such properties are associated with more success in finding friends and partners.

Occasionally, such changes may be sparked—or furthered—by singular events. This may happen when a celebrity reports a property or is otherwise "outed". Such events may be said to invalidate the perceived pattern that this attribute is associated with weakness and other negatively connotated properties and should not be talked about in public. It can establish a new pattern that associates the attribute also with strong and positively connotated people and thus makes it easier for others to take this into the public. Examples include depression, e.g. Ärzteblatt (2011), other—in particular mental – diseases, and alcoholism, e.g. Guardian (2002).

The Web offers many medial possibilities of supporting and sustaining awareness, for example by the popular list formats.[18] In addition, its communication opportunities, such as those afforded by online support groups, appear to be very conducive to discussing sensitive topics (cf. King and Moreggi 1998). Scientific data on such developments (especially in the orders of magnitude deemed interesting by data mining) are often more difficult to extract/construct than the inputs of more standard KD applications such as Web navigation, search queries or social-networking-site friendship links. More importantly, research that draws on such online data poses hard ethical challenges (Eysenbach and Till 2001) and requires expertise beyond computer science.

An example in the scientific literature is the controversial exploratory data analysis of Donohue and Levitt (2001), who found the legalisation of abortion to be a predictor of sinking crime rates. This data analysis sparked intense debate both about its content and the statistical methods used.

Solution approaches PPDM addresses some of these issues. In addition to the modifications to the data (so that "private data remain private") as discussed in the previous section, some PPDM methods also modify the results (so that "private knowledge remains private"). An interesting development in the context of rule hiding is *discrimination-aware data mining* (Pedreschi et al. 2008). The motivation is that US (and other) laws prohibit discrimination on the basis of race, colour, religion, nationality, gender, marital status, age and pregnancy in various settings such as credit scoring or personnel decisions. The authors introduce and study the notion of discriminatory classification rules. Such rules propose a decision (e.g., whether to give a loan) based on a discriminatory attribute in a direct way (appearing in the rule premise) or in an indirect way (appearing in an associated rule). The authors propose metrics to control for such discrimination (see also Pedreschi et al. 2009; Calders and Verwer 2010; Hajian et al. 2011 for further work).

Solutions to the threats may depend on restrictions on earlier stages of KD (in particular, data selection and integration; it is unlikely that there is reason to forbid a specific model class). More openness (which attributes are defined how and inferred how, what error margin does a model have) and restrictions on deployment are likely to increase acceptance and disputability. A careful assignment of the burden of proof is often necessary: If an inference is uncertain, it is often not adequate (in the legal sense) to require that the accused customer or citizen prove that this inference is erroneous.

3.9 Evaluation

Evaluation is the step in which the analyst should ascertain that the results of the previous stages "properly achieve[…] the business objectives. A key objective is to determine if there is some important business issue that has not been sufficiently considered. At the end of this phase, a decision on the use of the data mining results should be reached" (CRISP-DM 2000, p. 14). In this phase, all the previously raised problems should be reviewed again in order to make sure that the deployment will

[18] For one example, see http://depression.about.com/od/famous/Famous_People_With_Depression.htm.

be as privacy-protecting as possible (or as desired). It is important to note that this should only concern unexpected results; many problems can be foreseen and should be dealt with in the requirements analysis, systems specification and further KD-specific phases.

One example of how current technology could support evaluation and reflection loops by analysts is given by a closer look at the idea of discrimination-aware data mining. The suppression of discriminatory mining patterns, while an important first step in criticizing and avoiding undesired profiling, rests on pre-defined discriminatory attributes and can therefore not help against the forming of new profiles. Thus, discrimination-aware data miners could evaluate their results, critically reviewing if their findings point to the necessity of adding more attributes to established catalogues such as race, religion, pregnancy status, etc. (Gao and Berendt 2011).

3.10 Deployment

In this phase, the gained insight is used, for example in real-time personalization of Web page delivery and design or other decision processes (what contract to offer or deny a customer, whether to search a traveller at the border or not, etc.). Many people regard this phase as the core danger of data mining to privacy; for example, Clifton et al. (2004, p. 192) write that "[i]t is […the] use of personal data in a way that negatively impacts someone's life, that causes concern. As long as data is not misused, most people do not feel their privacy has been violated". The considerations above as well as recent empirical findings for example in social networks show that deployment is not the only concern many people have. As an example, consider the concerns about indeterminate visibility of information (regardless of what practical consequences this may have) identified by Gürses et al. (2008). We therefore treat the deployment phase as one out of several phases relevant for privacy.

The brevity of this section should not be misunderstood as an indication of a judgment that this be less relevant than the other phases; indeed, we have made frequent references to real-life consequences throughout this paper. The reason is that deployment choices are mostly outside the discretion of the data analyst; they tend to be choices made by corporate management or governmental agents. Thus, the data miner as the primary target group of this article should be aware that she can help prepare or enable such decisions, both by the knowledge she creates (or does not create) and by the way she communicates the validity and generalisability of the reported findings. However, since the decision whether to deploy is generally a business or political decision, a more in-depth treatment of this phase is outside the scope of the present article.

Threats The operational steps of deployment may be intruding into a private sphere per se (e.g., searching someone at an airport, searching their home and/or computer). They may also contribute to the knowledge about a data subject and thus be similar to more data being collected and stored—private data may become visible generally or to new and unintended recipients (privacy-as-hiding and as control). Examples in the Web domain comprise the application of classifiers predicting political affiliation

(Lindamood et al. 2009) or social-security numbers (Acquisti and Gross 2009) from open information in social network sites. Finally, deployment operations may install social categories and norms as "facts" with all the consequences of such redefinitions of reality (less consumer choice, heightened social inequalities, more people treated as criminals, etc.). Resulting perceptions of what is "knowable" by KD may lead people to restrict their public utterances and thereby affect privacy-as-practice.

A further threat originates from the inherently statistical nature of patterns—each pattern type has certain types of errors (e.g., misclassifications for classifiers, arbitrary groupings in clustering solutions, confidence values in association rules). This means that even if data subjects would agree that the category itself exists (e.g, child pornography collectors), the inference that a particular person actually belongs to that category may simply be wrong (cf. for example Daten-speicherung.de 2010 for a survey of incidents).

Opportunities Conversely, the deployment of socially liberating insights may have beneficial consequences in the senses outlined above.

Solution approaches Data mining is a costly process—therefore there is a strong incentive to use a piece of derived knowledge that promises economic advantage. Thus, only significant expected losses of consumer goodwill and thus ultimately revenue upon using this knowledge, or effective legal provisions against using it, are likely to prevent its use. Many recent examples, especially in communities of social-network users who observe "their" network's practices closely, have illustrated how these two can collaborate (cf. Gürses et al. 2008; Gürses 2010; EPIC 2011b).

4 A major new challenge: external effects and privacy stakeholders

Communities have come to play a major role on the Web, and SNSs like Facebook, MySpace or LinkedIn offer themselves as platforms for such communities. SNSs have recently come under increased scrutiny.

First, SNSs induce detailed self-profiling. Second, first uses of SNSs relational information for marketing purposes ("viral marketing", Richardson and Domingos 2002) have backfired because they were seen as violating privacy. A much-discussed case was the Facebook Beacon that broadcast news about a Facebook user's purchases on partner sites like ebay to all of this user's contacts. Facebook decided to allow global opt-out of Beacon after heated debate especially in the blogosphere (for an overview, see EPIC 2011a). The underlying data collection uses Javascript and cookies and is a direct continuation of DoubleClick's technology (see Sect. 3.3). Regulators have begun to see these problems and are working towards developing better ways of protecting users (Article 29 Data Protection Working Party 2009) while respecting the SNS business model (which rests on utilizing profile and relational information).

However, the current discussion mostly focuses on protecting users against malpractices by site operators. This overlooks a second key aspect that has already been implicit in the discussions of privacy-as-practice above, but is today becoming highly visible in SNS: the external effects of *users* on other users. Put simply, there are cases

in which one person guarding or giving up her own privacy may breach or strengthen the privacy of others. This means that regulations may put restrictions on individual wishes to give up privacy, or at least forbid that consequences are taken that might make sense from a purely economic point of view.

For example, if a majority of health-insurance clients disclose their individual risk profiles to their insurance companies, it will invariably be concluded that the others have something to hide, present high risk, and should therefore be charged higher premiums. This may make chronically ill patients effectively uninsurable and create deep rifts through society. As a political consequence of these considerations, public health insurance companies are not allowed to act on risk information received from clients (although the clients are free to disclose this information).

Depending on how basic the right is considered to be, the need may also arise for the state to "protect people against their own wishes". For example, German law treats human dignity as an absolute human right in the sense that the state *must* protect people against being treated as objects—even if they consent to it. Recent examples include consensual cannibalism (treated in court as murder) [BVerfG(2008)] and the sport dwarf-tossing (illegal in some jurisdictions, cf. ereleases 2001). The case is less clear for privacy, as the uproar (that soon sank into oblivion) over TV shows such as "Big Brother" in various European countries has shown (Dörr 2000).

These externalities have consequences for system requirements analysis and design. Different parties may have diverging or even conflicting understandings of privacy. It may be difficult to use privacy-enhancing technologies such that these conflicting interests are articulated and resolved. Developing and implementing privacy enhancing technologies without an understanding of the larger context (= the way a system is used, the way a system is surrounded by other systems, etc.) may also fall short of attending to the privacy concerns of different users in ubiquitous environments. Such shortcomings may result in privacy breaches that may cause a loss of trust in such systems and affect technology acceptance.

In Gürses et al. (2006), we have proposed to use multilateral security requirements analysis for capturing and addressing the diverse and possibly conflicting views of privacy (and for translating them into more tangible and implementable security requirements on a system).

The way one user handles data about the relation with another user can have an impact on what data about that other individual is disclosed. As an example, consider friendship relations which are—at least in real life—symmetric. Thus, the record of person A that states that person B is a friend also contains information that is part of B's record.[19] Another example is groups of users. Group attributes may be changed by any member of the group. A user whose group membership is public thereby discloses interests, preferences, or other personal information. This means that if A discloses information about himself or groups including himself, he (whether will-

[19] It is a matter of debate to which extent friendships really are symmetric. In offline worlds as in online social networks, people may declare themselves "friends of each other" for a multitude of reasons, and this may often be asymmetric in many different senses. Some offline conventions or online platforms (such as Twitter's "follow" relationship) even allow for explicitly asymmetric social relations. We cannot investigate this question in any depth here; it suffices to point to the enforced symmetry of, e.g., Facebook's "friend" relation that requires a request and an acknowledgement in order to be created.

ingly or inadvertently) also discloses information about someone else (Preibusch et al. 2007). Expressed differently, A's treatment of his privacy has a direct effect on B's privacy.

Another example are the publication and treatment of *relational information* (such as the fact that someone has a friend-link to someone else) by the privacy settings of an SNS user. While sophisticated access-control techniques (Carminati et al. 2009) can avoid many undesired accesses, they require attentive management of settings, and they cannot avoid problems that arise intrinsically from the networked structure. For example, if A allows "friends-of-friends" to see things (such as his photos), then he is in fact giving B *transitive access control* over parts of his resources (because B can now determine, by her choice of friends, who sees A's photos). Such settings can easily lead to conflicts, such as when A stipulates that only friends may see his friend links, but B allows friends-of-friends to see her friend links (the connection A–B is then visible to more people than desired by A). For a formal analysis of such cases, see (Gürses and Berendt 2010b).

Such social-network data usually concern people who also have an ID in the same system, i.e. this privacy dependency is a problem that affects different users of the same system.

In addition, problems arise when systems support the interaction with the world outside the system. For example, Google Mail (Gmail) users consent to their emails' data being analysed by Google; however, all incoming mails of a Google Mail account (whether sent by another Gmail user or by somebody else) are also analysed. Thus, A's treatment of his privacy also has direct external effects on the privacy of C, who is a non-user of the system.

The distinction between "in the system" and "outside the system" vanishes in case of loosely coupled networks where members may engage in relationships spontaneously and without a central authority. An example are the "friend-of-a-friend" nets built by publishing FOAF files (Smarr 2001). A FOAF file describes a person's contact information, as well as his/her relationships to other people and details about them in an RDF-based standard format. As users publish their friendship details autonomously, symmetry of relationships is not enforced. However, revelation of private information is likely to occur for instance by combining real names and email addresses, and legal requirements apply (Court of Justice 2003).

Because SNSs are (by definition) built on interaction, they are typically open systems, and have certain semantic characteristics. Each privacy-related declaration has effects beyond the interaction between one individual data subject and one data collector, effects that may concern a number of stakeholders who may or may not be users of the same system.

In a quest for solutions, we identified two essential steps: First, the potential privacy conflicts that arise by social-network interaction must be identified, and methods for conflict detection, negotiation and resolution must be employed. Technological approaches that guarantee a strict sense of privacy such as anonymizers are neither always possible nor desirable. Interestingly, the need for *negotiations* of privacy has been pointed out from two very different directions. One of them is an often commercial perspective that regards privacy as the object of a "deal" between users and service providers on the Web: P3P and the wider concepts of privacy (policy) negotiations

(W3C 2006b; Preibusch 2006). The idea of users specifying their respective desired levels of privacy has also been investigated within a PPDM framework (Xiao and Tao 2006). The other is a civic-democratic perspective that criticizes the commercialization of personal data (especially if the profiles then become "private" property of the data-collecting and data-mining organizations, see Phillips 2004) and emphasises the need for negotiations on the societal (rather than individual) level. Current research on privacy in social networks shows the need to go beyond the individual level and consider also the external effects of privacy-related behaviour. These results may help focus research and practice on connecting these two positions, and may establish various forms of negotiations in the requirements engineering process.

Second, privacy preferences and requirements must be formalised sufficiently such that software can automatically detect problems, alert the user, and assist her. In data analysis routines, mechanisms need to be implemented to enforce privacy requirements. In Preibusch et al. (2007), we have extended the INFERENCE element to describe inferences based on the semantics of social relations, encoded in different privacy levels (private, group, community, public). Again, the logic on top of the INFERENCE element can ensure that data are hidden that, while themselves seemingly innocuous, would allow unwanted inferences.

Third, concerns that transcend the preferences and requirements of stakeholders must be incorporated. This is the case in particular where privacy becomes a public good, i.e. where "people need to be protected against their own wishes" as in the examples above.

Resolving differences and problems and evaluating the proposed procedures remain subjects for future work, but a language for the representation of requirements, interaction scenarios, and conflicts is a necessary first step in addressing these issues.

5 Summary and conclusions

In sum, this article shows that privacy is a multi-faceted, contextual and dynamic notion. This holds in particular in the Web as environment and vis-à-vis Web mining as a data analysis form. Technological approaches that guarantee a strict sense of privacy (hiding or confidentiality) such as anonymizers are neither always possible nor always desirable. Methods from privacy-preserving data mining and privacy-preserving data publishing address some of the problems, but a wider notion of privacy and a consideration of data mining as the whole KD cycle illustrate a much wider range of threats, opportunities and solution approaches. Major open issues are to (a) understand and map the range of privacy concepts and mining-related challenges and opportunities, (b) communicate these in usable ways, (c) develop methods and deploy technological and other solutions, and (d) create and maintain the interdisciplinary collaborations necessary for these steps.

References

Acquisti A, Gross R (2006) Imagined communities: awareness, information sharing, and privacy on the Facebook. In: Danezis G, Golle P (eds) Privacy enhancing technologies. LNCS, vol 4258. Springer, New York, pp 36–58

Acquisti A, Gross R (2009) Predicting social security numbers from public data. Proc Nat Acad Sci 106(27):10975–10980

Acquisti A, Balsa E, Berendt B, Clarke D, De Wolf R, Diaz C, Gao B, Gürses SF, Kuczerawy A, Pierson J, Piessens F, Sayaf R, Schellens T, Stutzman F, Van Alsenoy B, Vanderhoven E (2011) SPION deliverable 2.1 State of the art. COSIC Internal Technical Report, K.U. Leuven, Belgium. http://www.cosic.esat.kuleuven.be/publications/article-2077.pdf

Agrawal R, Srikant R (2000) Privacy-preserving data mining. In: SIGMOD conference. ACM, Dallas, pp 439–450

Aggarwal CC, Yu PS (2008a) A general survey of privacy-preserving data mining models and algorithms. In: Aggarwal CC, Yu PS (eds) Privacy-preserving data mining: models and algorithms. Springer, New York, pp 11–51

Aggarwal CC, Yu PS (eds) (2008b) Privacy-preserving data mining: models and algorithms. Springer, New York

Agrawal R, Kiernan J, Srikant R, Xu Y (2002) Hippocratic databases. In: VLDB. Morgan Kaufmann, San Francisco, pp 143–154

Agre PE, Rotenberg M (2001) Technology and privacy: the new landscape. MIT Press, Cambridge

Aïmeur E, Brassard G, Fernandez JM, Onana FSM (2008) Alambic: a privacy-preserving recommender system for electronic commerce. Int J Inf Secur 7(5):307–334

Anderson R (2008) Security engineering, 2nd edn. Wiley, Chichester

Article 29 Data Protection Working Party (2009) Opinion 5/2009 on online social networking. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp163_en.pdf

Ärzteblatt (2011) Initiative für psychische Gesundheit im Leistungssport (Initiative for mental health in professional sports). http://www.aerzteblatt.de/nachrichten/46692/Initiative-f_r-psychische-Gesundheit-im-Leistungssport.htm

Azevedo A, Santos MF (2008) KDD, SEMMA and CRISP-DM: a parallel overview. In: IADIS European conference data mining. IADIS, pp 182–185

Backstrom L, Dwork C, Kleinberg JM (2007) Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. In: Williamson CL, Zurko ME, Patel-Schneider PF, Shenoy PJ (eds) WWW. ACM, Ithaca , pp 181–190

Baeza-Yates R, Jones R, Poblete B, Spiliopoulou M (2010) Issues with privacy preservation in query log mining. In: Ferrari E, Bonchi F (eds) Privacy-aware knowledge discovery: novel applications and new techniques. Chapman & Hall/CRC Press, Boca Raton

Barbaro M, Zeller T (2006) A face is exposed for AOL searcher no. 4417749. New York Times, New York

BBC News (2008) Timeline: child benefits records loss. http://news.bbc.co.uk/2/hi/7104368.stm

Berendt B, Günther O, Spiekermann S (2005) Privacy in e-commerce: stated preferences vs. actual behavior. Commun ACM 48(4):101–106

Berendt B, Preibusch S, Teltzrow M (2008) A privacy-protecting business-analytics service for online transactions. Int J Electron Commer 12(3):115–150

Berkovsky S, Borisov N, Eytani Y, Kuflik T, Ricci F (2007) Examining users' attitude towards privacy preserving collaborative filtering. In: Baker R, Beck J, Berendt B, Menasalvas E, Kröner A, Weibelzahl S (eds) Proceedings of the workshop on data mining for user modelling at UM 2007. http://vasarely.wiwi.hu-berlin.de/DM.UM07/Proceedings/DM.UM07-proceedings.pdf

Bertino E, Lin D, Jiang W (2008) A survey of quantification of privacy preserving data mining algorithms. In: Aggarwal CC, Yu PS (eds) Privacy-preserving data mining: models and algorithms. Springer, New York, pp 181–200

Binder J, Howes A, Sutcliffe AG (2009) The problem of conflicting social spheres: effects of network structure on experienced tension in social network sites. In: CHI, ACM, pp 965–974

Bone RG (1998) A new look at trade secret law: doctrine in search of justification. Calif Law Rev 86(2):241–313

Bonneau J, Preibusch S (2009) The privacy jungle: on the market for data protection in social networks. In: WEIS 2009, http://preibusch.de/publications/social_networks/privacy_jungle_dataset.htm

Borra E (2007) Repurposing the Wikiscanner. https://wiki.digitalmethods.net/Dmi/WikiScanner

boyd D, Ellison N (2007) Social network sites: definition, history and scholarship. J Comput Mediat Commun 13(1), http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html

Boyens C, Fischmann M (2003) Profiting from untrusted parties in web-based applications. In: EC-Web. LNCS, Springer, vol 2738, pp 216–226

Bundesverfassungsgericht (1983) BVerfGE 65, 1—Volkszählung. Urteil des Ersten Senats vom 15. Dezember 1983 auf die mündliche Verhandlung vom 18. und 19. Oktober 1983—1 BvR 209, 269, 362, 420, 440, 484/83 in den Verfahren über die Verfassungsbeschwerden. http://www.telemedicus.info/urteile/Datenschutzrecht/88-BVerfG-Az-1-BvR-209,-269,-362,-420,-440,-48483-Volkszaehlungsurteil.html

Burton TM (2002) Florida probes Lilly's mailings of Prozac samples to patients. Wall Street Journal, 8 July 2002. http://www.mindfully.org/Industry/Eli-Lilly-Privacy-InfringementWSJ26jul02.htm

Calders T, Verwer S (2010) Three naive Bayes approaches for discrimination-free classification. Data Min Knowl Discov 21(2):277–292

Camenisch J, Kohlweiss M, Rial A, Sheedy C (2009) Blind and anonymous identity-based encryption and authorised private searches on public key encrypted data. In: Public key cryptography. LNCS, vol 5443. Springer, New York, pp 196–214

Canny JF (2002) Collaborative filtering with privacy via factor analysis. In: SIGIR. ACM, pp 238–245

Carminati B, Ferrari E, Perego A (2009) Enforcing access control in web-based social networks. ACM Trans Inf Syst Secur 13(1)

Carminati B, Ferrari E, Kantarcioglu M, Thuraisingham B (2010) Privacy issues in web-based social networks. In: Ferrari E, Bonchi F (2010) Privacy-aware knowledge discovery: novel applications and new techniques. Chapman & Hall/CRC Press, Boca Raton

Chor B, Kushilevitz E, Goldreich O, Sudan M (1998) Private information retrieval. J ACM 45(6):965–981

Ciriani V, di Vimercati SDC, Foresti S, Samarati P (2008) k-anonymous data mining: a survey. In: Aggarwal CC, Yu PS (eds) Privacy-preserving data mining: models and algorithms. Springer, New York, pp 103–134

Clifton C, Kantarcioglu M, Vaidya J (2004) Defining privacy for data mining. In: Kargupta HData mining: next generation challenges and future directions. AAAI/MIT Press, Menlo Park

Court of Justice (2003) Judgment of the court of 6 November 2003. Criminal proceedings against Bodil Lindqvist. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:62001J0101:EN:HTML

CRISP-DM (2000) CRISP-DM 1.0 Step-by-step data mining guide. http://www.spss.ch/upload/1107356429_CrispDM1.0.pdf

Dalenius T (1977) Towards a methodology for statistical disclosure control. Stat Tidskrift 15:429–444

Danezis G, Diaz C (2008) A survey of anonymous communication channels. Tech. Rep. MSR-TR-2008-35, Microsoft Research

Danezis G, Díaz C, Faust S, Käsper E, Troncoso C, Preneel B (2007) Efficient negative databases from cryptographic hash functions. In: ISC. LNCS, vol 4779, Springer, New York, pp 423–436

Daten-speicherung.de (2010) Fälle von Datenmissbrauch und -irrtümern (Cases of data abuse and errors). http://daten-speicherung.de/wiki/index.php?title=F%C3%A4lle_von_Datenmissbrauch_und_-irrt%C3%BCmern&oldid=3712

Department of Homeland Security (2007) Statement by Homeland Security Chief Privacy Officer Hugo Teufel III on the privacy act system of records notice for the automated targeting system. http://www.dhs.gov/xnews/releases/pr_1186178812301.shtm

Díaz C (2005) Anonymity and privacy in electronic services. PhD thesis. Department of Electrical Engineering, K.U. Leuven

Domingo-Ferrer J (2007) A three-dimensional conceptual framework for database privacy. In: Secure data management. LNCS, vol 4721. Springer, New York, pp 193–202

Domingo-Ferrer J, Torra V (2008) A critique of k-anonymity and some of its enhancements. In: ARES 2008. IEEE Computer Society, pp 990–993

Domscheit-Berg D (2011) Inside wikileaks: my time with Julian Assange at the worlds most dangerous website. Random House, New York

Donohue J, Levitt S (2001) The impact of legalized abortion on crime. Q J Econ 116(2):379–420

Dörr D (2000) Big Brother und die Menschenwürde: Die Menschenwürde und die Programmfreiheit am Beispiel eines neuen Sendeformats [[Big Brother and Human Dignity: Human dignity and media freedom; the case of a new television format]]. Peter Lang, Frankfurt am Main

Dwork C (2006) Differential privacy. In: Bugliesi M, Preneel B, Sassone V, Wegener I (eds) ICALP (2), LNCS. vol 4052. Springer, New York , pp 1–12

Dwork C (2008) Differential privacy: a survey of results. In: TAMC. LNCS, vol 4978. Springer, New York, pp 1–19

Electronic Privacy Information Center (2000) In the Matter of DoubleClick, Inc. Complaint and request for injunction, request for investigation and for other relief. http://epic.org/privacy/internet/ftc/DCLK_complaint.pdf
Electronic Privacy Information Center (2007) In the Matter of Google, Inc. and DoubleClick, Inc. Complaint and request for injunction, request for investigation and for other relief. http://epic.org/privacy/ftc/google/epic_complaint.pdf
Electronic Privacy Information Center (2010a) Automated targeting system. http://epic.org/privacy/travel/ats/
Electronic Privacy Information Center (2010b) Netflix cancels contest over privacy concerns. http://epic.org/2010/03/netflix-cancels-contest-over-p.html
Electronic Privacy Information Center (2011a) Facebook privacy. http://epic.org/privacy/facebook/
Electronic Privacy Information Center (2011b) In re Facebook. http://epic.org/privacy/inrefacebook/
Electronic Privacy Information Center (2011c) Privacy? Proposed Google/DoubleClick deal. http://epic.org/privacy/ftc/google/
Electronic Privacy Information Center (n.d.) EPIC online guide to practical privacy tools. http://epic.org/privacy/tools.html
ereleases (2001) Florida ban on dwarf tossing must Be upheld, announces LPA, Inc. http://www.ereleases.com/pr/florida-ban-on-dwarf-tossing-must-be-upheld-announces-lpa-inc-865
EU (1995) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML
EU (2002) Directive 2002/58/EC of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:HTML
EU (2006) Directive 2006/24/ec of the European Parliament and of the Council of 15 march 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending directive 2002/58/ec. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32006L0024:EN:HTML
European Court of Human Rights (2011) European Court of Human Rights: violation of freedom of expression in case Heinisch v. Germany. ECHR 115 (2011). http://human-rights-convention.org/2011/07/22/european-court-of-human-rights-violation-of-freedom-of-expression-in-case-heinisch-v-germany/
Eysenbach G, Till JE (2001) Ethical issues in qualitative research on internet communities. Br Med J 323:1103–1105
Fayyad UM (1996) Data mining and knowledge discovery: making sense out of data. IEEE Expert 11(5):20–25
Feiler L (2008) The data retention directive. http://www.rechtsprobleme.at/doks/feiler-DataRetentionDirective.pdf
Fleischer P (2008) Are IP addresses "personal data"? http://peterfleischer.blogspot.com/2007/02/are-ip-addresses-personal-data.html
Frankowski D, Cosley D, Sen S, Terveen LG, Riedl J (2006) You are what you say: privacy risks of public mentions. In: SIGIR. ACM, pp 565–572
FTC (2000) Privacy online: fair information practices in the electronic marketplace: a Federal trade commission report to congress. http://www.ftc.gov/reports/privacy2000/privacy2000.pdf
Fung BCM, Wang K, Chen R, Yu PS (2010) Privacy-preserving data publishing: a survey on recent developments. ACM Comput Surv 42(4)
Gao B, Berendt B (2011) Visual data mining for higher-level patterns: discrimination-aware data mining and beyond. In: Proceedings of the 20th machine learning conference of Belgium and The Netherlands. http://www.benelearn2011.org/
German Working Group on Data Retention (AK Vorrat) (n.d.) Pros and cons of data retention. http://www.vorratsdatenspeicherung.de/content/view/83/87/lang,en/
Goldberg I (2007) Privacy-enhancing technologies for the internet III: ten years later. In: Digital privacy: theory, technologies, and practices, Auerbach, pp 3–18, http://www.cypherpunks.ca/~iang/pubs/pet3.pdf
Gostin LO (2004) International infectious disease law. Revision of the World Health Organization's international health regulations. J Am Med Assoc 291(21):2623–2627

Guarda P, Zannone N (2009) Towards the development of privacy-aware systems. Inf Softw Technol 51(2):337–350

Guardian (2002) Anthony Hopkins remembers alcoholism, fondly. http://www.guardian.co.uk/film/2002/mar/01/news1

Gürses S (2010) Multilateral privacy requirements analysis in online social network services. PhD thesis, Department of Computer Science, K.U. Leuven

Gürses S, Berendt B (2010) PETs in the surveillance society: a critical review of the potentials and limitations of the privacy as confidentiality paradigm. In: Gutwirth S, Poullet Y, De Hert P (eds) Data protection in a profiled world. Springer, Dordrecht , pp 301–321

Gürses S, Berendt B (2010b) The social web and privacy: practices, reciprocity and conflict detection in social networks. In: Ferrari E, Bonchi F (eds) Privacy-aware knowledge discovery: novel applications and new techniques. Chapman & Hall/CRC Press, Boca Raton

Gürses S, Berendt B, Santen T (2006) Multilateral security requirements analysis for preserving privacy in ubiquitous environments. In: Proceedings of the workshop on ubiquitous knowledge discovery for users at ECML/PKDD 2006, Berlin, pp 51–64. http://vasarely.wiwi.hu-berlin.de/UKDU06/Proceedings/UKDU06-proceedings.pdf

Gürses S, Rizk R, Günther O (2008) Privacy design in online social networks: learning from privacy breaches and community feedback. In: ICIS. ACM

Hajian S, Domingo-Ferrer J, Martínez-Ballesté A (2011) Discrimination prevention in data mining for intrusion and crime detection. In: IEEE SSCI 2011

Hancock JT, Birnholtz JP, Bazarova NN, Guillory J, Perlin J, Amos B (2009) Butler lies: awareness, deception and design. In: CHI. ACM, pp 517–526

Hansen M (2008) Linkage control—integrating the essence of privacy protection into identity management. In: Proceedings of eChallenges. pp 1585–1592

Hay M, Miklau G, Jensen D (2010) Private analysis of network data. In: Ferrari E, Bonchi F (eds) Privacy-aware knowledge discovery: novel applications and new techniques. Chapman & Hall/CRC Press, Boca Raton

Hildebrandt M (2006) Privacy and identity. In: Claes E, Duff A, Gutwirth S (eds) Privacy and the criminal law. Intersentia, Antwerp , pp 43–58

Hildebrandt M (2008) Profiling and the identity of the European citizen. In: Hildebrandt M, Gutwirth SProfiling the European citizen: cross-disciplinary perspectives. Springer, New York

Hornick MF, Marcade E, Venkayala S (2007) Java data mining: strategy, standard, and practice. Morgan Kaufman, San Francisco

Hu J, Zeng HJ, Li H, Niu C, Chen Z (2007) Demographic prediction based on user's browsing behavior. In: WWW. ACM, pp 151–160

Infektionsschutzgesetz (2001) Gesetz zur Verhütung und Bekämpfung von Infektionskrankheiten beim Menschen (Law for the avoidance and fighting of infectious human diseases). http://www.gesetze-im-internet.de/bundesrecht/ifsg/gesamt.pdf

King SA, Moreggi D (1998) Internet therapy and self help groups—the pros and cons. In: Gackenbach JPsychology and the Internet. Academic Press, San Diego

Kohavi R, Provost F (1998) Glossary of terms. http://ai.stanford.edu/~ronnyk/glossary.html

Korolova A, Kenthapadi K, Mishra N, Ntoulas A (2009) Releasing search queries and clicks privately. In: WWW. pp 171–180

Lederer S, Hong JI, Dey AK, Landay JA (2004) Personal privacy through understanding and personal privacy through understanding and action: five pitfalls for designers. Pers Ubiq Comput 8(6):440–454

Li N, Li T, Venkatasubramanian S (2007) t-closeness: privacy beyond k-anonymity and l-diversity. In: ICDE. IEEE, pp 106–115

Lindamood J, Heatherly R, Kantarcioglu M, Thuraisingham BM (2009) Inferring private information using social network data. In: WWW. ACM, pp 1145–1146

Lindell Y, Pinkas B (2000) Privacy preserving data mining. In: CRYPTO. LNCS, vol 1880. Springer, New York, pp 36–54

Lipford HR, Besmer A, Watson J (2008) Understanding privacy settings in Facebook with an audience view. In: UPSEC'08: proceedings of the 1st conference on usability, psychology, and security. USENIX Association, pp 1–8

Liu H, Mihalcea R (2007) Of men, women, and computers: data-driven gender modeling for improved user interfaces. In: ICWSM, pp 121–128

Liu H, Maes P, Davenport G (2006) Unraveling the taste fabric of social networks. Int J Semant Web Inf Syst 2(1):42–71

Machanavajjhala A, Gehrke J, Kifer D, Venkitasubramaniam M (2006) l-diversity: privacy beyond k-anonymity. In: ICDE. IEEE Computer Society, p 24

Nakashima E (2011) Bradley manning, WikiLeaks' alleged source, faces 22 new charges. 2 March 2011. http://www.washingtonpost.com/wp-dyn/content/article/2011/03/02/AR2011030206272.html

Narayanan A, Shmatikov V (2009) De-anonymizing social networks. In: Proceedings of 30th IEEE symposium on security and privacy 2009

Nguyen DH, Mynatt E (2002) Privacy mirrors: understanding and shaping socio-technical ubiquitous computing. Technical Report GIT-GVU-02-16, Georgia Institute of Technology, USA http://smartech.gatech.edu/handle/1853/3268

OECD (1980) Guidelines on the protection of privacy and transborder flows of personal data. http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html

Owad T (2006) Data mining 101: finding subversives with amazon wishlists. http://www.applefritter.com/bannedbooks

Palen L, Dourish P (2003) Unpacking "privacy" for a networked world. In: CHI. ACM, pp 129—136

Pedreschi D, Ruggieri S, Turini F (2008) Discrimination-aware data mining. In: KDD. ACM, pp 560–568

Pedreschi D, Ruggieri S, Turini F (2009) Measuring discrimination in socially-sensitive decision records. In: SDM. pp 581–592

Pfitzmann A, Hansen M (2006–2010) Anonymity, unlinkability, unobservability, pseudonymity, and identity management—a consolidated proposal for terminology. http://dud.inf.tu-dresden.de/Anon_Terminology.shtml

Phillips D (2004) Privacy policy and PETs: the influence of policy regimes on the development and social implications of privacy enhancing technologies. New Media Soc 6(6):691–706

Pilkington E (2006) Millions assigned terror risk score on trips to the US. The Guardian, 2 December 2006 http://www.guardian.co.uk/usa/story/0,,1962299,00.html

Poblete B, Spiliopoulou M, Baeza-Yates RA (2007) Website privacy preservation for query log publishing. In: Bonchi F, Ferrari E, Malin B, Saygin Y (eds) PinKDD, LNCS. vol 4890. Springer, New York , pp 80–96

Poblete B, Spiliopoulou M, Baeza-Yates RA (2010) Privacy-preserving query log mining for business confidentiality protection. TWEB 4(3):10–11026

Preibusch S (2006) Implementing privacy negotiations in e-commerce. In: APWeb. LNCS, vol 3841. Springer, New York, pp 604–615

Preibusch S, Hoser B, Gürses S, Berendt B (2007) Ubiquitous social networks—opportunities and challenges for privacy-aware user modelling. In: Baker R, Beck J, Berendt B, Menasalvas E, Kröner A, Weibelzahl S (Eds) Proceedings of the workshop on data mining for user modelling at UM 2007. http://vasarely.wiwi.hu-berlin.de/DM.UM07/Proceedings/DM.UM07-proceedings.pdf

Privacy International (2004) Complaint: Google Inc—Gmail email service. http://www.privacyinternational.org/issues/internet/gmail-complaint.pdf

Privacy International (2007) A race to the bottom: privacy ranking of internet service companies. http://www.privacyinternational.org/article.shtml?cmd%5B347%5D=x-347-553961

Reuters (2006) AOL chief technology officer resigns: sources. http://www.kdnuggets.com/news/2006/n16/36i.html

Richardson M, Domingos P (2002) Mining knowledge-sharing sites for viral marketing. In: KDD. ACM, pp 61–70

Rowe M, Ciravegna F (2010) Disambiguating identity web references using web 2.0 data and semantics. J Web Semant 8(2)

Sagar R (2007) On combating the abuse of state secrecy. J Polit Philos 15:404–427. doi:10.1111/j.1467-9760.2007.00283.x

Sagar R (2011) Das missbrauchte Staatsgeheimnis. Wikileaks und die Demokratie (The abused state secret. Wikileaks and democracy). In: Wikileaks und die Folgen. (Wikileaks and the consequences), Suhrkamp, Berlin, pp 201–223

Shearer C (2000) The CRISP-DM model: the new blueprint for data mining. J Data Wareh 5(4):13–22

Sinus Sociovision (n.d.) Sinus-Milieus. http://www.sociovision.de/loesungen/sinus-milieus.html

Smarr J (2001) Technical and privacy challenges for integrating FOAF into existing applications. http://www.w3.org/2001/sw/Europe/events/foaf-galway/papers/fp/technical_and_privacy_challenges/

Smith-Spark L (2006) How to blog—and keep your job. BBC News. http://news.bbc.co.uk/2/hi/europe/5195714.stm
Sweeney L (2002) k-anonymity: a model for protecting privacy. Int J Uncertain Fuzziness Knowl Based Syst 10(5):557–570
Teltzrow M, Kobsa A (2003) Impacts of user privacy preferences on personalized systems—a comparative study. In: CHI-2003 workshop "designing personalized user experiences for eCommerce: theory, methods, and research. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.8180
Verykios VS, Bertino E, Fovino IN, Provenza LP, Saygin Y, Theodoridis Y (2004) State-of-the-art in privacy preserving data mining. SIGMOD Rec 33(1):50–57
W3C (2006a) Platform for privacy preferences (P3P) project. http://www.w3.org/P3P
W3C (2006b) Workshop on languages for privacy policy negotiation and semantics-driven enforcement. http://www.w3.org/2006/07/privacy-ws/
Wardlow DL (ed) (1996) Gays, lesbians, and consumer behavior. Theory, practice, and research issues in marketing. Haworth Press, New York
Warren S, Brandeis L (1890) The right to privacy. Harward Law Rev 4:193–220
Westin AF (1970) Privacy and freedom. Atheneum, New York
Wu X, Ying X, Liu K, Chen L (2010) A survey of privacy-preservation of graphs and social networks. In: Aggarwal CC, Wang H (eds) Managing and mining graph data. Kluwer, Boston , pp 421–454
Xiao X, Tao Y (2006) Personalized privacy preservation. In: SIGMOD conference. ACM, pp 229–240
Zhao Y, Du M, Le J, Luo Y (2009) A survey on privacy preserving approaches in data publishing. In: DBTA. IEEE Computer Society, pp 128–131
Zhou B, Pei J, Luk W (2008) A brief survey on anonymization techniques for privacy preserving publishing of social network data. SIGKDD Explor 10(2):12–22