Privacy-Enhancing Technologies

Yang Wang
Bren School of Information and Computer Sciences
University of California, Irvine
Donald Bren Hall 5091
Irvine, CA 92697-3440
yangwang@uci.edu
(949)-351-9336 (phone)
(949) 824-4056 (fax)

Alfred Kobsa
Bren School of Information and Computer Sciences
University of California, Irvine
Donald Bren Hall 5092
Irvine, CA 92697-3440
kobsa@uci.edu
(949) -202-5704 (phone)
(484)-762-6644 (fax)

Accepted for publication in: M. Gupta and R. Sharman, eds.: Handbook of Research on Social and Organizational Liabilities in Information Security. Hershey, PA: IGI Global, to appear.

Privacy-Enhancing Technologies

ABSTRACT

Privacy-enhancing technologies (PETs), which constitute a wide array of technical means for protecting users' privacy, have gained considerable momentum in both academia and industry. However, existing surveys of PETs fail to delineate what sorts of privacy the described technologies enhance, which makes it difficult to differentiate between the various PETs. Moreover, those surveys could not consider very recent important developments with regard to PET solutions. The goal of this chapter is two-fold. First, we provide an analytical framework to differentiate various PETs. This analytical framework consists of high-level privacy principles and concrete privacy concerns. Secondly, we use this framework to evaluate representative up-to-date PETs, specifically with regard to the privacy concerns they address, and how they address them (i.e., what privacy principles they follow). Based on findings of the evaluation, we outline several future research directions.

INTRODUCTION

Privacy has been recognized as a fundamental human right at least since the seminal treatise of Warren and Brandeis (Warren & Brandeis, 1890). However, it is only in recent decades that privacy issues have attracted substantive attention in society, due to the proliferation and advancement of innovative information technologies such as computers, the Internet, and recently mobile and ubiquitous computing applications. Despite its importance, the concept of privacy is difficult to grasp. Privacy is a truly multi-dimensional notion. It involves, but is not limited to, cultural, social, legal, political, economic and technical aspects.

Privacy-enhancing technologies (PETs), which constitute a wide array of technical means for protecting users' privacy, have gained considerable momentum in both academia and industry. A number of overviews of the PET landscape have already been published (Blarkom, Borking, & Verhaar, 2003; Burkert, 1997; Camp & Osorio, 2003; Goldberg, 2002; Senicar, Jerman-Blazic, & Klobucar, 2003; Tavani & Moor, 2001). However, most of these studies fail to delineate what sorts of privacy the described technologies enhance, which makes it difficult to differentiate between the various PETs. Moreover, those surveys could not consider very recent important developments with regard to PET solutions. We will therefore focus on these newer solutions here (specifically on privacy policy languages and systems aimed at empowering users in their privacy decisions), and conduct an in-depth examination of the privacy landscape in which these PETs are supposed to make meaningful contributions. More classical PETs such as authentication and identity management systems as well as systems that provide authorization and access control will only be briefly mentioned in the passing. So does another class of very specialized PETs, namely privacy-preserving personalization methods, which have been described in (Y. Wang & Kobsa, Forthcoming).

The goal of this chapter is to provide an analytical framework upon which to chart past, present and future research on PETs. It is our belief that a deeper understanding of their underpinnings will enable us to identify gaps that may still exist, and research directions in developing next-generation PETs.

The remainder of the chapter is organized as follows. Firstly, we provide a review of current privacy-related regulatory requirements and users' privacy concerns and preferences. Secondly, we introduce our analytical framework consisting of privacy principles and privacy concerns. Thirdly, we use this framework to evaluate representative PETs, specifically with regard to the privacy concerns they address, and how they address them (i.e., what privacy principles they follow). Fourthly, we further discuss the findings. Finally, we conclude with an outline of promising future research directions.

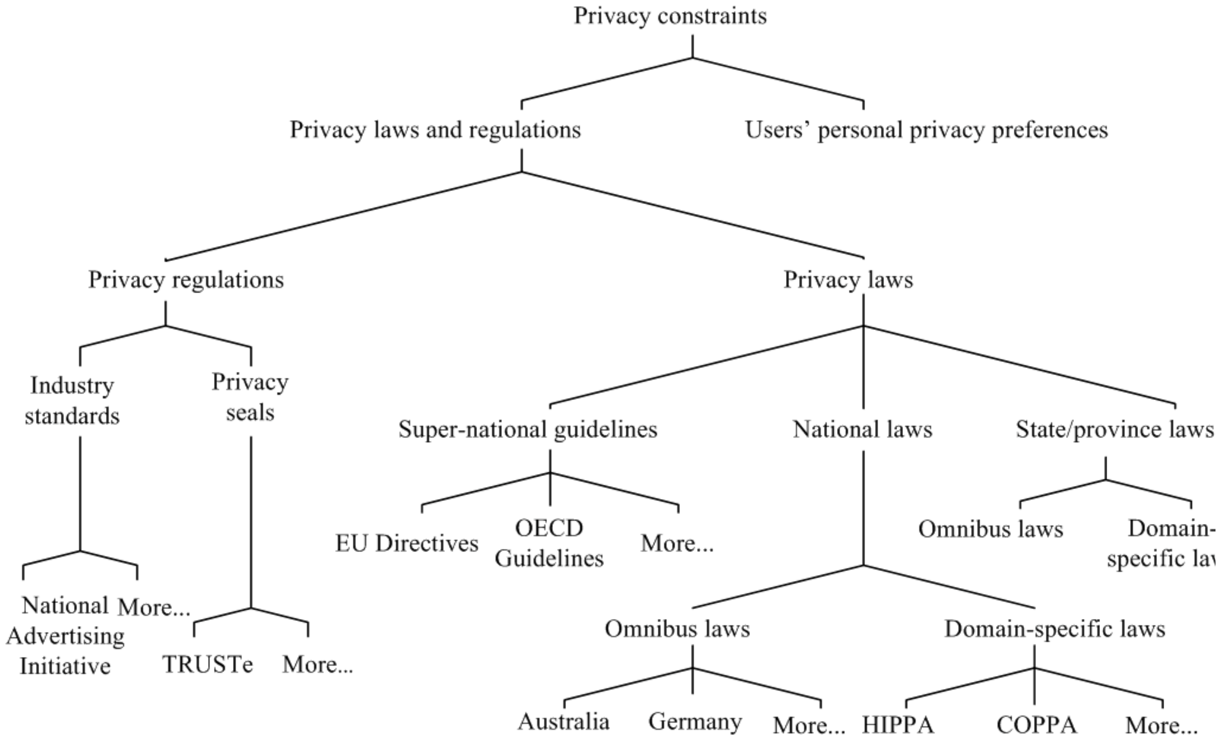

Fig. 1. The hierarchy of potential privacy constraints

THE PRIVACY LANDSCAPE

Privacy has been studied for decades, and many different definitions of privacy have been proposed. This is largely due to the fact that privacy is "an overwhelmingly large and nebulous concept" (Boyle & Greenberg, 2005). Young (Young, 1978) wittedly commented that "privacy, like an elephant, is … more readily recognized than described". In essence, privacy is personal, nuanced, dynamic, situated and contingent (Dourish & Anderson, 2006; Palen & Dourish, 2002).

If privacy considerations are taken into account in the design of computer systems, they restrain the possible design space for such systems. Solutions that violate privacy constraints cannot be considered any more. Privacy constraints for computer systems stem primarily from two sources, namely from privacy laws and regulations and from personal privacy expectations of the computer users. Figure 1 shows the hierarchy of these constraints with a focus on privacy laws and regulations.

Privacy laws and regulations

We have witnessed a proliferation of privacy laws and regulations during the past 30 years. More than 40 countries currently have national privacy laws enacted. In addition, numerous other types of privacy regulations, industry seal programs and company self-governing policies have been introduced as well.

Privacy laws and regulations generally apply when personal data of people are being processed who can be identified with reasonable means. These laws and regulations usually lay out organizational and technical requirements for ensuring the protection of personal data that is stored and/or processed in information systems. These requirements include, but are not limited to, proper data acquisition, notification about the purpose of use, permissible data transfer (e.g., to third parties and/or across national borders) and permissible data processing (e.g., organization, modification and destruction). Other requirements prescribe user opt-ins (e.g., asking for their consent before collecting their data), opt-out (e.g., of data collection and/or data processing) and user inquiries (e.g., regarding what personal information was collected and how it was processed and used). Other stipulations mandate the establishment of adequate security mechanisms (e.g., access control for personal data), and the supervision and audit of personal data processing.

Historically, Europe and the U.S. launched parallel initiatives in privacy and data protection (Westin & Gelder, 2003). In the 1970s, the U.S. inaugurated legislation for the protection of citizen and consumer information databases, with the Fair Credit Reporting Act of 1970 and the Privacy Act of 1974. The Fair Information Practice Principles that were first formulated by the U. S. Department of Health, Education and Welfare in 1973 became the basis for many privacy laws and regulations worldwide. A number of Western European countries, such as Sweden, Germany and France followed the move in the early 1970s and early 1980s.

In 1980, the Organisation for Economic Co-operation and Development (OECD) drafted *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (OECD, 1980) These guidelines are however not binding for its currently 30 member countries, which include the U.S.. The European Union issued two privacy-related directives (EU, 1995, 2002) that set out the minimum standards for its member states to implement in their respective national privacy laws. The Asia-Pacific Economic Cooperation (APEC) recently also drafted a privacy framework (APEC-FIP, 2004), serving as recommendations for its currently 21 member countries including the U.S.

In the U.S., several sector-specific laws have come into effect such as the Health Insurance Portability and Accountability Act of 1996 (HIPAA) for medical privacy, the Children's Online Privacy Protection Act of 1999 (COPPA) for protecting children under the age of 13, the Gramm-Leach-Bliley Act of 1999 for financial privacy, and the Sarbanes–Oxley Act of 2002 for accounting and financial reporting. In 2000, the Federal Trade Commission (FTC) published a widely known report to Congress on Fair Information Practice Principles (FTC, 2000b). In the same year, the FTC also issued the so-called Safe Harbor Principles (FTC, 2000c) to meet the adequacy standard imposed by the EU. In 2006, the Association for Computing Machinery (ACM), the world's largest computer science association, announced recommendations of privacy principles drafted by its U.S. Public Policy Committee (USACM, 2006).

Users' personal privacy preferences

While privacy laws and regulations are on the normative side of privacy, personal privacy attitudes represent its subjective aspect. Numerous opinion polls and empirical studies have revealed that Internet users harbor considerable privacy concerns regarding the disclosure of their personal data to websites, and the monitoring of their Internet activities. These studies were primarily conducted between 1998 and 2003, mostly in the United States but also in other countries. In the following, we summarize a number of important findings (the percentage figures indicate the ratio of respondents from multiple studies who endorsed the respective views). For more detailed discussions we refer to (Kobsa, 2007; Teltzrow & Kobsa, 2004).

Personal data.

1. Internet users who are concerned about the privacy or security of their personal information online: 70% - 89.5%;
2. People who have refused to give personal information to a web site at one time or another: 82% - 95%;
3. Internet users who would never provide personal information to a web site: 27%;
4. Internet users who supplied false or fictitious information to a web site when asked to register: 6% - 40% always, 7% often, 17% sometimes;
5. People who are concerned if a business shares their data for a purpose that is different from the original purpose: 89% - 90%.

User tracking and cookies.

1. People concerned about being tracked on the Internet: 54% - 63%;
2. People concerned that someone might know their browsing history: 31%;

3. Users who feel uncomfortable being tracked across multiple web sites: 91%.
4. Internet users who generally accept cookies: 62%;
5. Internet users who set their computers to reject cookies: 10% - 25%;
6. Internet users who delete cookies periodically: 53%.

Behavioral experiments moreover show that Internet users also follow up on their privacy concerns to some extent, and supply more data about themselves, make purchases more frequently, and are willing to pay a small premium when interacting with e-commerce sites that noticeably have good privacy practices ((Gideon, Cranor, Egelman, & Acquisti, 2006; Kobsa & Teltzrow, 2005; Metzger, 2006; Tsai, Egelman, Cranor, & Acquisti, 2007)) .

ANALYTICAL FRAMEWORK FOR EVALUATING PRIVACY-ENHANCING TECHNOLOGIES

To evaluate the effectiveness of various PETs, we propose an evaluation framework that analyzes solutions along two dimensions:

(1) *What high-level principles the privacy solution follows*
We identify a set of fundamental privacy principles from various privacy laws and regulations, and treat them as high-level guidelines for enhancing privacy. Other principles that are desirable for privacy enhancement (e.g., usability) are also included.

(2) *What privacy concerns the privacy solution addresses*
While privacy principles are high-level guidelines to enhance privacy, privacy concerns are more concrete and mundane. Ideally one would need user studies to examine how effectively solutions address users' changing and contingent privacy needs and preferences. Since running such studies for every evaluated solution is barely realistic, we instead chose to investigate privacy concerns that are somewhere in between high-level privacy principles and low-level contingent privacy needs of users.

Furthermore, to be better able to assess the privacy protections that privacy-enabling technologies afford, we propose to group them into the following three categories:

- Protection of identity: this type of privacy protection aims to prevent users' true identities from being revealed (i.e., who they are).
- Seclusion: this type of privacy protection attempts to prevent users from being bothered by unwanted contact or solicitation (e.g., spam emails).
- Control over data: this type of privacy protection allows users to have control over their data, e.g. regarding what data can be collected or disclosed for what purpose, how the data will be used, and with whom the data may be shared or to whom it may be transferred.

Principles

Privacy legislation and regulations are usually instantiations of more fundamental privacy principles. We select a core set of privacy principles that are frequently addressed in privacy laws and regulations, and add other principles/properties that are also desirable for privacy enhancement. This list of principles is by no means exhaustive, but meant to initiate a discussion on what principles are desirable for enhancing privacy effectively. The principles are grouped by their origin in the listing below.

Privacy principles from privacy laws and regulations

1. Notice/Awareness
   - Privacy policy: *Make […] privacy policy statements clear, concise, and conspicuous to those responsible for deciding whether and how to provide the data* (USACM, 2006);

- Notice upon collection: *Whenever any personal information is collected, explicitly state:*
    - *the precise purpose for the collection,*
    - *all the ways in which the information might be used,*
    - *all the potential recipients of the personal data,*
    - *how long the data will be stored and used;* (USACM, 2006)

2. Data minimization
   Before deployment of new activities and technologies that might impact personal privacy, carefully evaluate them for their necessity, effectiveness, and proportionality: the least privacy-invasive alternatives should always be sought (USACM, 2006).

3. Purpose specification
   The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose (OECD, 1980).

4. Collection limitation
   There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means […] (OECD, 1980).

5. Use limitation
   Personal data should not be disclosed, made available or otherwise used for purposes other than those specified (OECD, 1980).

6. Onward transfer
   Personal data should not be transferred to a third country/party if it does not ensure an adequate level of protection (EU, 1995; FTC, 2000c).

7. Choice/Consent
   Where appropriate, individuals should be provided with clear, prominent, easily understandable, accessible and affordable mechanisms to exercise choice in relation to the collection, use and disclosure of their personal information (APEC-FIP, 2004). The two widely adopted mechanisms are (FTC, 2000a):
   - Opt-in: *requires affirmative steps by the consumer to allow the collection and/or use of information*;
   - Opt-out: *requires affirmative steps to prevent the collection and/or use of such information*.

8. Access/Participation
   An individual should have right to:
   - *know whether a data controller has data relating to her (OECD, 1980),*
   - *inspect and make corrections to her stored data (USACM, 2006).*

9. Integrity/accuracy
   A data controller should ensure the collected personal data is sufficiently accurate and up-to-date for the intended purposes and all corrections are propagated in a timely manner to all parties that have received or supplied the inaccurate data (USACM, 2006).

10. Security
    Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data (OECD, 1980).

11. Enforcement/Redress
    - Effective privacy protection must include mechanisms for enforcing the core privacy principles. At the minimum, the mechanisms must include (FTC, 2000c):

- Recourse mechanisms for customers: *readily available and affordable independent recourse mechanisms by which an individual's complaints and disputes can be investigated and resolved and damages awarded where the applicable law or private sector initiatives so provide;*
- Verification mechanisms for data controllers: *follow-up procedures for verifying that the attestations and assertions businesses make about their privacy practices are true and that privacy practices have been implemented as presented*;
- Remedy mechanisms: *obligations to remedy problems arising out of failure to comply with these principles by organizations announcing their adherence to them and consequences for such organizations.*

Anonymity-related principles

12. Anonymity
    Anonymity means that users cannot be identified nor be tracked online.

13. Pseudonymity
    Pseudonymity also means that users cannot be identified, but they can still be tracked using a so-called alias or persona. The German Telemedia Law (DE-TML, 2007) mandates that profiling necessitates the use of pseudonyms or the prior consent of the user.

14. Unobservability
    A data controller cannot recognize that a system/website is being used or visited by a given user.

15. Unlinkability
    A data controller cannot link two interaction steps of the same user.

16. Deniability
    Deniability means that users are able to deny some of their characteristics or actions (e.g., having visited a particular website), and that others cannot verify the veracity of this claim.

Other desirable principles for privacy enhancement

17. User preference
    Different users can potentially have different privacy preferences. A data controller should tailor its privacy practices to each individual user's preferences.

18. Negotiation
    This principle calls for the support of negotiation between a user and a website, during which they can reach an agreement on privacy practices that the website may employ for the respective user.

19. Seclusion
    Seclusion means that users have the right to be left alone. Violations of this principle in the electronic world are popup ads and junk emails.

20. Ease of adoption
    Oftentimes privacy protection mechanisms rely on the presence of other infrastructures or technologies, and this fact may pose significant barriers for adoption. This principle relates to the readiness of *organizations* to adopt the examined privacy protection (e.g., whether the solution relies on special protocols or technologies that are proprietary or not readily available).

21. Ease of compliance
    This principle is concerned with the ease of fulfilling legal requirements by adopting a specific privacy protection solution.

Table 1. Privacy guidelines/frameworks and privacy principles

Specification **Principle**	OECD Guidelines (OECD, 1980)	EU Directive on Data Protection (EU, 1995)	German Telemedia Law (DE-TML, 2007)	APEC Privacy Framework (APEC-FIP, 2004)	FTC Safe Harbor Principles (FTC, 2000c)	FTC Fair Info Practice (FTC, 2000b)	ACM Principles (USACM, 2006)
Notice/Awareness	X	X	X	X	X	X	X
Minimization							X
Purpose specification	X	X	X	X	X		X
Collection limitation		X	X	X	X		
Use limitation	X	X	X	X	X		X
Onward transfer		X	X		X		
Choice/Consent	X	X	X	X	X	X	X
Access/Participation	X	X	X	X	X	X	X
Integrity/accuracy	X	X	X	X	X	X	X
Security	X	X	X	X	X	X	X
Enforcement/Redress		X	X		X	X	
Anonymity-related principles							
Anonymity							
Pseudonymity		X	X				
Unobservability							
Unlinkability							
Deniability							
Other desirable principles for privacy enhancement							
User preference							
Negotiation							
Seclusion							
Ease of adoption							
Ease of compliance							
Usability							
Responsiveness							

22. Usability
    The privacy protection solution should be easy for *users* to adopt (e.g., the efforts required from users to utilize the solution should be reasonable).

23. Responsiveness
    The privacy protection solution should respond promptly to changes in users' privacy decisions.

Privacy laws and regulations typically only include subsets of the above principles. For a comparison, Table 1 shows a group of representative privacy laws and regulations in its columns, and the privacy principles discussed above in its rows. An "X" in a cell means that the framework includes the respective principle.

Applying our privacy protection taxonomy to the principles

We now categorize the 23 identified principles based on the type of privacy protection they relate to. Note that the general category contains principles that pertain to all three types of privacy protection. Table 2 represents which category each privacy principle falls into.

Table 2. Categorization of principles based on the type of privacy protection

Protection **Principle**	General	Protection of Identity	Seclusion	Control over data
Notice/Openness	X			
Minimization				X
Purpose specification				X
Collection limitation				X
Use limitation				X
Onward transfer				X
Choice/Consent	X			
Access/Participation				X
Integrity/accuracy				X
Security				X
Anonymity		X		
Pseudonymity		X		
Unobservability		X		
Unlinkability		X		
Deniability		X		
Enforcement/Redress	X			
User preference	X			
Negotiation	X			
Seclusion			X	
Ease of adoption	X			
Ease of compliance	X			
Usability	X			
Responsiveness	X			

Privacy concerns

Whereas privacy principles are high-level guidelines for enhancing privacy, users' privacy concerns are more concrete and down to the earth. We discuss and analyze them here in order to also be able to evaluate the effectiveness of privacy enhancements from a subjective stance. Privacy concerns usually arise from characteristics of a specific application domain. To illustrate this, we will focus on the potential privacy concerns that may arise in web personalization (Brusilovsky, Kobsa, & Nejdl, 2007), such as Amazon's personalized book recommendations.

Wang et al. (H. Wang, Lee, & Wang, 1998) present a taxonomy of privacy concerns in Internet marketing that includes improper access, improper collection, improper monitoring, improper analysis, improper transfer, unwanted solicitation and improper storage. These high-level concerns as well as concerns about improper merging of data also apply in web personalization. Table 3 shows what privacy concerns (columns) can arise from typical web personalization activities (rows).

Table 3. Potential privacy concerns in typical web personalization activities

	Control over data							Seclusion	Protection of identity
	Improper acquisition			Improper use			Improper storage	Unwanted solicitation	Identity fraud/theft
	Improper access	Improper collection	Improper monitoring	Improper analysis	Improper merge	Improper transfer			
Tracking		XX	XX						
Profiling		X	X	X	X	X			X
Cross-website recommendation		X	X	X	XX	XX	X	X	
Single-website recommendation		X	X	X	X	X	X	X	
Third-party data sharing				XX	X	XX	X	X	X
Direct mailing				X				XX	

XX: very likely X: likely

EVALUATING PETS

In this section, we will review major privacy-enhancing technologies, namely privacy policy languages, anonymity techniques, authentication and identity management, authorization and access control, usable security and privacy mechanisms. We will contrast them against the two evaluative elements of our framework, namely what principles they follow and what privacy concerns they address. This close examination of existing PETs will allow for a more comprehensive view of their pros and cons as well as their current gaps, and thus point out future research avenues.

Privacy policy languages

The U.S. Federal Trade Commission (FTC) defines a privacy policy as a comprehensive description of a company's information practices, accessible by clicking at a hyperlink on the company's website (FTC, 1998). Its aim is to enhance users' awareness of the privacy practices of the website. Privacy policies thus are directed at human readers.

Privacy policy languages, in contrast, are intended to be machine-readable. They can be roughly divided into two types: external policy languages to describe websites' public privacy policies or users' privacy preferences, and internal ones to specify companies' or websites' internal rules for privacy practices. In general, external privacy policy languages are declarative without enforcement mechanism, while internal privacy policy languages are normative with support for enforcement.

External privacy policy language

P3P: The Platform for Privacy Preferences

Developed by the World Wide Web Consortium (W3C), the Platform for Privacy Preference (P3P 1.1) (L. Cranor et al., 2006) aims at increasing the transparency of websites' privacy practices in such a way that users can easily decide whether or not these websites meet their privacy expectations. Technically, P3P consists of two parts: (1) a standard machine-readable (XML) language/syntax that allows websites to describe their privacy practices regarding the collection, use, and distribution of personal information, and (2) a "handshake" protocol built on top of the HTTP protocol that enables P3P-enabled user agents (e.g., web browsers) to retrieve websites' P3P privacy policies automatically (Garfinkel & Cranor, 2002). Agents can also be configured to inform users about the sites' privacy policies, to notify them when those change, to warn them when those deviate from their pre-specified privacy preferences (expressed in languages like APPEL (L. Cranor, Langheinrich, & Marchiori, 2002) or XPref (Agrawal, Kiernan, Srikant, & Xu, 2003)), and to semi-automate or automate the decision whether or not to disclose the requested information on users' behalf.

A P3P policy file can be applied to a whole website or certain parts of it such as web pages, images, cookies, forms and even a single form field[1]. Every P3P policy contains a description of the legal entity responsible for the privacy policy, whether the site allows users to have access to the information collected about them, (optional) information regarding dispute resolution and remedy, and at least one statement. Each statement describes the data being collected (physical contact information, online contact information, purchase information, click stream data, etc.), the purpose(s) for collection (web site administration, research and development, profiling, etc.), whether the site supports user opt-in or opt-out for those purposes, what organizations will have access to the collected data (primary service provider only, delivery services, unrelated third parties, etc.), the retention of the collected data (single session, stated purpose, indefinitely, etc). Personalization can be considered as one type of purposes dubbed as "individual decision"[2] and similarly anonymous personalization as "pseudo decision".

P3P was designed as part of a broader privacy protection framework (including privacy legislation and enforcement) and is applicable to any web-based systems. P3P implementations include

- P3P user agents (such as Internet Explorer 6 (Microsoft, 2000) that supports cookie management as well as websites' privacy policies disclosure,
- AT&T Privacy Bird (L. Cranor, 2002), an add-on to the Internet Explorer, that utilizes differently colored bird icons in the corner of the browser window to indicate whether or not a site's P3P policy matches the user's preferences,
- Privacy Bird Search Engine (Byers, Cranor, Kormann, & McDaniel, 2004) that annotates regular search results with an indication to what extent the P3P policy of each site matches the user's requirements,
- P3P policy generators/editors/checkers (e.g., P3PEdit), and
- server-side P3P support (e.g., IBM Tivoli Privacy Manager For E-business that can enforce privacy policies internally in a system).

P3P's official website currently lists about 2900 websites worldwide that have adopted P3P 1.0[3]. The latest P3P adoption study conducted in the summer of 2005 (Egelman, Cranor, & Chowdhury, 2006) estimates the overall P3P adoption rate at about 10% using a list of "typical" search terms taken from AOL users' queries, and the government adoption rate roughly at 36% (this is by far the largest sector to adopt P3P, which is probably due to the P3P adoption mandate of the E-Government Act[4] (USA, 2002)). The usage of P3P has also been proposed in the context of ubiquitous computing (Langheinrich, 2002).

Despite its relative popularity, P3P has a number of limitations:

First, P3P does not include any technical mechanism for enforcing privacy policies. It is totally up to the websites to follow their stated privacy policies, and users cannot verify whether a site acts as promised.

Second, P3P (even the latest version 1.1) does not support different policies for different users, albeit offering users a choice of P3P policies is mentioned in P3P's future plan. Nevertheless, several proposals for individual negotiation of P3P policies have been made (Buffett, Jia, Liu, Spencer, & Wang, 2004; Preibusch, 2006).

Third, P3P might not be expressive enough to be able to fully encode the nuances of websites' privacy practices. For example, P3P cannot handle cases where privacy concerns crosscut more than one statement (e.g., that personal data that were obtained for different purposes may not be grouped (CZ, 2000)). Because of this lack in expressiveness and enforcement, it is difficult for websites to keep their P3P privacy policies, human-readable privacy policies and actual practices all consistent. P3P

[1] The P3P 1.1 specification provides a new mechanism that binds a P3P policy to an XML element that does not have to be associated with a URI.

[2] Information may be used to determine the habits, interests, or other characteristics of individuals and combine it with identified data to make a decision that directly affects that individual.

[3] http://www.w3.org/P3P/compliant_sites.php3

[4] The act mandates that government agencies post machine-readable privacy policies on their web sites.

allows sites to further explain those nuances in the human-readable fields, which are not technically analyzable though.

Furthermore, P3P is not able to accurately capture the subtleties of privacy laws and regulations, nor does it develop a minimum set of privacy or security standards that web sites should follow. Therefore, websites cannot rely on P3P as a technical means to comply with relevant privacy provisions, and discrepancies might exist between the P3P policies and the applicable privacy laws which in turn can expose the websites to legal jeopardy (L. F. Cranor & Reidenberg, 2002).

Last but not least, P3P has been criticized for facilitating websites' data collection, rather than protecting users' privacy (Coyle, 1999). A convincing example is that the default value for data retention is "indefinitely" instead of "no retention" or "stated purpose". Besides, P3P makes it difficult for users to protect their privacy (for instance, users find changing defaults for cookie settings to be burdensome and confusing (EPIC & Junkbusters, 2000). P3P could also effectively excludes non-compliant sites (e.g., websites without P3P compact policies would be blocked by IE6) and P3P-compliant sites do not make themselves more trustable (a study has shown that among the top 500 companies, only nine out of the 65 sites that adopted P3P have P3P satisfactory to a pragmatic user who wants some privacy protection (Ashrafi & Kuilboer, 2005)).

In order to gauge P3P's expressiveness, we attempted to describe six representative privacy provisions that have the biggest impact on the internal operation of web personalization (Y. Wang & Kobsa, 2006). The results show that P3P can express most of the provisions, with the following deficiencies: First, it needs more fine-grained expressiveness (e.g., retention time cannot be set in a continuous time scale; this can be easily solved by introducing an "expiry"-like sub-element for the retention element). Secondly, perhaps the biggest issue with the P3P language, it cannot express interactions across different statements; a potential solution is to introduce logic operators in the statement-group element so that different relationships between statements can be captured. Thirdly, the overall P3P framework is short of an interface with systems that enforce P3P privacy policies.

APPEL: A P3P Preference Exchange Language

APPEL (L. Cranor, Langheinrich, & Marchiori, 2002) was designed to complement P3P by allowing users to express their privacy preferences in terms of rules that specify certain conditions under which user information may be collected and used, so that P3P-enabled agents are able to check users' preferences against a website's P3P policy to make automated or semi-automated decisions whether or not users' data may be released to the website. A rule includes a behavior, an optional persona, optional explanation and prompt messages, and a number of expressions (L. Cranor, Langheinrich, & Marchiori, 2002). An expression is used to match a full XML element or a single attribute and its value in an XML element in the evaluated P3P policy.

APPEL only allows logical operations at nodes corresponding to P3P elements. The matching scheme of APPEL is problematic: a P3P policy can contain multiple statements, and a rule will fire and then stop being further evaluated against the policy if any of the statements satisfies the rule. Therefore, the remaining statements will be ignored. Because of this deficiency, APPEL only works correctly when rules express what is unacceptable rather than what is acceptable (Agrawal, Kiernan, Srikant, & Xu, 2003).

XPref

XPref (Agrawal, Kiernan, Srikant, & Xu, 2003) is another preference language for P3P and is based on the XPath language (Clark & DeRose, 1999), a W3C Recommendation for navigating and matching the hierarchical structure of an XML document. The biggest difference between XPref and APPEL is that APPEL uses the sub-elements of a rule to specify acceptable and unacceptable combinations of P3P elements, while XPref utilizes XPath expressions for the same purpose. XPref outweighs APPEL in that it can specify what is acceptable as well as what is unacceptable, and combinations of both. Agrawal et al (Agrawal, Kiernan, Srikant, & Xu, 2003) show that XPref subsumes APPEL, and that APPEL can be programmatically translated into XPref.

Individual privacy policy negotiation

Buffet et al (Buffett, Jia, Liu, Spencer, & Wang, 2004) present a framework for the negotiation between users and websites about the disclosure of user information for compensation. It relies on utility theory to allow users to express the value associated with each piece and combination of personal data. The proposed PrivacyPact protocol enables the transmission of messages for negotiation. Preibusch (Preibusch, 2006) identified relevant privacy dimensions (recipient, purpose, retention, and data) for negotiation and proposed a simple extension of P3P to allows for the expression and implementation of such negotiation processes.

Internal privacy policy language

EPAL: Enterprise Privacy Authorization Language

EPAL (Ashley, Hada, Karjoth, Powers, & Schunter, 2003) is a formal language developed by IBM that allows enterprises to write their internal privacy policies, so that those can be enforced across IT applications and systems in an automated manner. To take advantage of EPAL, an enterprise first defines an EPAL vocabulary (i.e., concrete element types) that caters to its own needs, and then specifies its EPAL privacy policies. Applications that aim to share an EPAL policy obviously must agree on the vocabulary and interpret it in the same way.

An EPAL policy contains a set of privacy authorization rules that allow or deny requests. Rules have the following constituents (Ashley, Hada, Karjoth, Powers, & Schunter, 2003):

- the action for which authorization is requested (e.g., disclose or read)
- the data categories upon which these actions are going to be performed (e.g. medical records or contact info),
- the user categories that are affected (e.g., a department or a particular employee),
- the usage purposes (e.g., direct marketing or auditing),
- associated conditions (e.g., the purpose category must be billing purpose), and
- associated obligations (e.g., delete data after 7 days or obtain consent).

An authorization request contains a user category, an action, a data category, and a purpose.

An authorization result is a statement that includes a ruling (allow or deny), a user category, an action, a data category, and a purpose. A rule may also contain conditions and obligations. Authorization results are used to determine if a request is allowed or denied. At first sight, EPAL policies seem quite similar to P3P policies since both describe privacy practices. However, they differ in the following ways:

- P3P is data-centric (i.e., a P3P statement covers different aspects of a specific type of data), while EPAL is access-centric (i.e., a EPAL rule refers to an instance of information access) (Stufflebeam, Anton, He, & Jain, 2004);
- EPAL aims to describe the internal privacy practices of an enterprise, which are probably not to be shared with the public, while P3P is used to describe public privacy policies;
- unlike P3P that predefines the elements of a privacy policy, EPAL elements (such as actions or user categories) are abstract types which are mapped to actual instances of elements during the implementation;
- the design of EPAL takes privacy legislation and regulations into account, by including an obligation element;
- EPAL policies are machine enforceable: they are akin to access control policies in the security domain. An authorization engine parses the EPAL policies to generate a ruling given a request, and subsequently an enforcing environment/software will execute the ruling;
- Conflicting EPAL rules are allowed and solved by prioritizing rules, to allow for general rules and exceptions;
- P3P policies are always formulated in a positive manner (i.e., what is acceptable, not what is unacceptable), whereas EPAL can express both via the ruling element (allow or deny).

Although EPAL was designed to be expressive and flexible so as to capture evolving privacy legislation and customized privacy policy, we observe that it fails to express some privacy provisions, for example, that personal data that were obtained for different purposes may not be grouped (CZ, 2000). Essentially, as with P3P, EPAL cannot express the interactions between different rules. EPAL's current abstraction of actions is not sufficient for privacy authorization purposes. Actions need to be modeled with finer granularity (e.g., modeled as hierarchyType like data categories). For example, in the case of the action “personalization/inference”, we believe that it should be further categorized into different specific personalization techniques (such as collaborative filtering and incremental machine learning). Our justification is based on the observation that if privacy laws apply to a personalized website, they may affect the methods (i.e., actions) that may be used for processing them (Kobsa, 2002). Or in other words, different specific actions may lead to opposite privacy authorization results. For instance, one-time machine learning methods that rely on a record of raw data from several user sessions are not permitted without the user’s consent under the German Telemedia Law (DE-TML, 2007). In contrast, incremental machine learning methods that discard the raw data of each user session and only retain the learning results may be employed. However, in the current language both types of machine learning methods/actions are modeled as an abstract “inference” action and thus the critical distinction in authorization results is lost.

XACML: eXtensible Access Control Markup Language

XACML[5] (OASIS, 2005) is a general-purpose access control policy language. XACML can be used to describe not only general access control policies/rules, but also access control decision requests and responses. The root element of an XACML policy document is a Policy or PolicySet element (a container for a set of policies). A Policy contains a set of access control Rules. A Rule includes a Condition (that can be nested) and the Rule’s effect (Permit or Deny). If the Condition evaluates to true, the Rule’s effect (Permit or Deny) is returned. If the Condition evaluates to false, it means the Rule is not applicable (NA is returned). If an error occurs during the evaluation, “Indeterminate” is returned.

When a user/subject wants to perform some action on a resource, she will make a request to the so-called “Policy Enforcement Point” (PEP) that protects the requested resource (e.g., a file system or web server). The PEP will generate a request that includes Subject, Resource, Action and optionally Environment attributes (values). The PEP will then submit the request to a Policy Decision Point (PDP), which will identify all policies that apply to the request by evaluating their Targets. A Target is a set of conditions associated with a PolicySet or Policy or Rule. When the Target evaluates to true, the corresponding PolicySet/Policy/Rule applies to the request. The PDP will then evaluate the request against the applicable policies, yielding a response that consists of an access control decision regarding whether or not the request should be allowed, and optionally a list of obligations (i.e., actions that the PEP is obligated to perform before granting or denying access). Since each Rule and Policy may evaluate to yield different access control decisions, the XACML utilizes a collection of Combining Algorithms (e.g., Deny Overrides Algorithms) to derive a single final access control decision. Finally, the PDP returns the final decision back to the PEP, which can then enforce the decision (namely either allow or deny the access).

XACML also comes with an approved OASIS Standard profile for privacy policies, for modeling how personally identifiable information is collected and used. An attribute of “resource:purpose” defines the purpose for which the data resource was collected. Another attribute of “action:purpose” indicates the purpose for which access to the data resource is requested. A Rule element mandates that access shall be denied unless the purpose for which access is requested matches, by regular-expression match, the purpose for which the data resource was collected.

A thorough comparison of the latest versions (EPAL 1.2 and XACML 2.0) shows that XACML is a functional superset of EPAL 1.2 and outweighs EPAL in expressing not only access control policies but also privacy policies. Specifically, XACML provides the following important features that EPAL lacks (Anderson, 2005):

- the ability to combine results of multiple policies developed by different policy issuers;
- the ability to reference other policies in a given policy;
- the ability to specify conditions on multiple subjects that may be involved in making a request;
- the ability to return separate results for each node when access to a hierarchical resource is requested;
- support for subjects who must simultaneously be in multiple independent hierarchical roles or groups;
- policy-directed handling of error conditions and missing attributes;

[5] At the time of writing, XACML 3.0 is being drafted.

- support for attribute values that are instances of XML schema elements; and
- support for additional primitive data types (including X.500 Distinguished Names, RFC822 names, and IP addresses).

Summary of privacy policy languages

Table 4 shows a summary of various privacy policy languages and their characteristics. Two observations can be made in this table. First, negotiation and enforcement still call for wider support. Second, XACML seems to surpass all other existing privacy policy languages, although it lacks support for negotiation.

Table 4. Privacy policy languages and their characteristics

	External privacy policy	Internal privacy policy	User preference	Expressiveness	Negotiation	Enforcement
P3P	+			+		
APPEL			+	+		
XPref			+	++		
PrivacyPact	+		+		+	
EPAL		+		++		+
XACML	+	+	+	+++		++

+++: Very strong support ++: Strong support +: Support

An integrated privacy management system based on privacy policy languages

The IBM Tivoli Privacy Manager is a comprehensive enterprise privacy management system that aims at supporting a variety of privacy enhancement functionalities (IBM, 2006):

- centralized authorship and management of an enterprise's privacy rules,
- a natural language interface to author and manage privacy policies,
- translation of privacy policy from prose to P3P,
- enforcement of privacy policies across the enterprise's IT infrastructure,
- monitoring access to personal information and generating detailed audit logs,
- notification and consent preferences for information sharing across the enterprise, and
- automatically generation of reports detailing compliance to corporate policies.

This solution focuses on privacy protection in the sense of "control over data". It marginally addresses "seclusion and barely touches the protection of identity. In other words, the Tivoli Privacy Manager cannot protect end users' identities, which is understandable since it is geared towards enterprise privacy management.

Anonymity techniques

Anonymity of a user means that she cannot be identified nor tracked online. One way to improve anonymity is what Goldberg et al. (Goldberg, Wagner, & Brewer, 1997) called "strip identifying headers and resend" approach. This approach has been used in anonymous email remailers (Gülcü & Tsudik, 1996) and anonymous web browsing tools like Anonymizer (ConneXion, 1996), a web proxy that strips off identifying headers and source addresses from the web browser.

Another approach is "onion routing" which is built upon the notion of "mix network" (Chaum, 1982). A mix network is essentially a chain of proxy servers (called mixes). In onion routing, a message or packet is encrypted to each mix node using public key cryptography. The resulting encryption is like a layered "onion" with the original message in the innermost layer. As the message traverse over the network, each mix node strips off its own layer of encryption to reveal where to send the message next. Untraceability can be achieved unless all mix nodes are compromised. For example, Tor (Tor), a concrete onion routing system can provide anonymous communication such as web browsing, remote login sessions, instant messaging and other applications that rely on the TCP protocol.

The third major approach is centered on the concept of "k-anonymity" (Sweeney, 2002). It is concerned with a practical problem of releasing data about individuals without revealing identifying information about them. In a k-anonymized release, each individual's record is indistinguishable from at least k-1 others' records. A myriad of policies and techniques (e.g., clustering (Aggarwal et al., 2006)) have been proposed to achieve k-anonymity.

Authentication and identity management

Authentication seeks to ensure that a user is actually the person who she claims to be. This is usually achieved by employing a username in combination with a password, where the username is considered as a digital identity of the bearer and the password as her authentication. A more sophisticated and thus more secure scheme is the so-called two-factor authentication, which involves two independent ways for verifying identity. It may include a user having something (e.g., a bank ATM card or a time-dependent token card) and the user knowing something (e.g., a PIN).

One of the goals of the emerging identity management systems is to allow users to have more than one digital identity and be able to freely choose which identity to use. For example, Google allows its users to use different identities/accounts in its various applications (so that, for instance, one's interactions with Google Calendar will not be combined and used in Google's personalized search).

Another recent industry example is Microsoft's CardSpace (Microsoft), an "identity metasystem" that allows users to create multiple virtual ID cards. Each virtual card created by the user would only contain the minimum amount of information (retrieved from an identity provider) that individuals will need to divulge to carry out the transaction to which the card applies. CardSpace thereby uses the metaphor of the various cards that we use to identify ourselves in the physical world, such as business cards, driver's licenses and credit cards. With these virtual cards, users no longer have to hassle with daunting passwords. CardSpace has been integrated into Microsoft's operating system Vista.

OpenID (OpenID) is an open specification of a truly distributed identity system. OpenID providers are essentially authentication brokers between users and OpenID-enabled websites. They allow users to log into an OpenID-suported website without registration, using a URI as a username that belongs to the user (e.g., the URL of her homepage or blog). Users' passwords and other credentials are safely stored by OpenID (which can be run by the user or by a third-party identity provider). Because of its open and distributed nature, ease of use, and easy adoption for websites (free libraries are available in most web programming languages), OpenID is gaining more and more momentum and emerges as the de-facto industry standard.

Authorization and access control

Authorization involves granting or denying specific access rights. In a classic access control model, an access matrix specifies what permissions each subject has on the resources the system retains. In a role-based access control model, permissions are assigned to roles instead of subjects directly (subjects can take on multiple roles, and multiple subjects can take on the same role). In a directory-based access control model, subjects are managed and organized in directories (e.g., in an LDAP server), and permissions are granted based on these different directories (Cannon, 2005).

Privacy policy languages such as P3P and XACML have an access control aspect since they prescribe who can access what information under what condition for what purpose.

Systems for empowering users in their privacy decisions

Security has long been primarily regarded as a technical and theoretical problem. It is well known though that many established security mechanisms are barely used in practice since they pose usability problems. A growing number of security researchers have therefore shifted towards so-called "usable security and privacy", which studies the usability of security and privacy mechanisms. This emerging field aims at uncovering the reasons behind the mismatch between technical security mechanisms and their practical usage by end users, and on ways of bridging the gap to better meet users' security needs.

Whitten and Tygar's conducted a seminal usability analysis of PGP 5.0 (Whitten & Tygar, 1999), to find out why users failed to achieve their security goals (encrypting and decrypting email messages in this case). They found that this is largely due to interface design problems, causing a mismatch between users' needs and the structure of the encryption technology. Bellotti and Sellen (Bellotti &

Sellen, 1993) identified two primary sources for a number of potential security and privacy problems, from their experiences in ubiquitous computing: disembodiment (the actors are invisible in actions) and dissociation (actions are invisible to actors), both of which are visibility issues.

In the light of rendering the invisible visible (privacy threats in this case), Ackerman and Cranor (Ackerman & Cranor, 1999) proposed privacy critics that are semi-autonomous agents and can monitor users' online actions, warn users about potential privacy threats and suggest available countermeasures. Gideon et al. (Gideon, Cranor, Egelman, & Acquisti, 2006) and Tsai et al. (Tsai, Egelman, Cranor, & Acquisti, 2007) confirmed the effectiveness of such awareness mechanisms empirically.

de Paula et al. (Paula et al., 2005) moved one step further. Instead of simply examining the usability of secure mechanisms, they framed security as an interaction problem (a practical, situated and contingent problem of decision making) and looked at a broader concern: "how security can manifest itself as part of people's interactions with and through information systems". In other words, security cannot be confined within components of a system specifically designed to attain security, but is an intrinsic and pervasive aspect of a broader context that includes end users, work practices and information systems. They argued that in practice the key issue is not how theoretically secure the underlying security mechanisms are, but rather to what extent end users can understand and make effective use of the secure mechanisms. They deliberately turned their "attention away from traditional considerations of expression and enforcement and towards explication and engagement". They designed Impromptu, a peer-to-peer file-sharing application based on supporting informed decision-making via two design principles: (1) the dynamic real-time visualization of system state, and (2) the integration of configuration and action. The former principle aims at helping users comprehend and assess the consequences of their actions when making privacy decisions. The later is based on the observation that "the separation of configuration and action may result in either overly rigid or ineffective control over security".

In short, these solutions underlie the strategy dubbed as "user empowerment" – helping users make informed privacy decisions.

DISCUSSION

Table 5 below presents how a set of representative PETs address the privacy concerns, and Table 6 shows in what ways these solutions follow the privacy principles. We now discuss some observations from these tables, and then propose implications for future research in the next section.

Two observations can be made in these two tables. First, privacy protection solutions form clusters. Solutions of the same type tend to address almost identical privacy concerns by following similar privacy principles, while different types of solutions address different but not necessarily disjoint concerns, and follow different but not necessarily disjoint principles. For example, P3P-enabled user agents such as Privacy Bird, PrivacyPact and Ackerman and Cranor's (Ackerman & Cranor, 1999) privacy critics address all listed privacy concerns except improper access. They do this by applying general principles (e.g., Notice/Openness) and data principles (e.g., purpose specification) but not identity principles. In contrast, identity management tools (such as CardSpace and OpenID) and anonymizers (e.g., Anonymizer) attend to concerns such as improper monitoring and improper use (e.g., improper merge) by observing identity principles (e.g., pseudonymity) but not data principles. This phenomenon indicates that research in data protection and identity management is still somewhat fragmented, albeit some overlap exists. Since both form integral parts of privacy enhancement, collaborations between the two research communities to integrate the two types of PETs would be desirable.

Second, no current PET effectively addresses all privacy concerns, nor follows all privacy principles. The IBM Tivoli privacy manager is the most comprehensive solution among those examined in this section. However, since it is designed and implemented as a server-side enterprise privacy management system, principles like usability (i.e., a PET solution should be easy for *end users* to adopt) are inevitably hard to achieve. Rather than mulling over whether one can develop a technical solution that effectively addresses all privacy concerns, our pragmatic strategy is to merely highlight directions that deserve more attention. For example, principles such as responsiveness, enforcement

Table 5. How PETs address privacy concerns

	Control over data							Seclusion	Protection of identity
	Improper acquisition			Improper use			Improper storage	Unwanted solicitation	Identity fraud/theft
	Improper access	Improper collection	Improper monitoring	Improper analysis	Improper merge	Improper transfer			
Privacy Bird		+	+	+	+	+	+	+	
Privacy Pact		+	+	+	+	+	+	+	
IBM Tivoli privacy manager		+	+	++	++	++	++	+	++
PGP	++								
CardSpace			+	+	++	+			+
OpenID			+	+	++	+			+
Anony-mizer			++	+	+	+			+
History/ cookie manager	+	++	++	+	+		+		+
Popup blocker/ Antispam								++	
Privacy critics		+	+	+	+	+	+	+	+

++: Effective +: Partially effective

and ease of compliance are barely supported by the discussed solutions, except for the IBM Tivoli privacy manager.

CONCLUSION AND FUTURE DIRECTIONS

We proposed and applied a multi-faceted approach to the investigation and evaluation of privacy-enhancing technologies, which considers both privacy principles and individuals' privacy concerns. Privacy principles thereby serve as high-level guidelines for the conceptual evaluation of technical solutions, while privacy concerns constitute user needs that privacy mechanisms need to address. Our analysis reveals trends, identifies deficiencies, and suggests future directions in this research area.

Based on our investigation of existing privacy enhancing technologies, we suggest the following directions for future research:

- Privacy needs to be treated as a first-class requirement from the early onset in the design of an information system since, like for security and usability, it is extremely difficult if not impossible to "retrofit" a completed system to make it more privacy-friendly.
- While compliance has long been technically framed and treated as a server-side problem, we believe that the "user empowerment" strategy has a great potential for compliance since the "expression and enforcement" paradigm seems too rigid to accommodate users' changing and context-dependent privacy desires.
- Since users' privacy needs and preferences are inherently dynamic and contingent, solutions need to cater to users' individual privacy needs. We start to see solutions like negotiable privacy policies that follow this promising direction.

Table 6. What privacy principles PETs follow

Solution Principle	Privacy Bird	Privacy Pact	IBM Tivoli privacy manager	PGP	Card Space	Open ID	Anony-mizer	History/ cookie manager	Popup blocker/ Antispam	Privacy critics
GENERAL										
Notice/Openness	+	+	+							+
Choice/Consent	+	+	+					+		
Accountability										
Enforcement/Redress			+							
User preference	+	+	+					+	+	
Negotiation		+								
Ease of adoption	+	-	-		-	+				+
Ease of compliance			+							
Usability	+					+				+
Responsiveness	+	+	+							
IDENTITY										
Anonymity							+			
Pseudonymity					+	+	+			
Unobservability							+			
Unlinkability					+		+	+		
Deniability										
SECLUSION										
Seclusion			+						+	
DATA										
Minimization		+								
Purpose specification	+	+	+							+
Collection limitation										
Use limitation	+	+	+							+
Onward transfer	+	+	+							+
Access/Participation										
Integrity/accuracy										
Security			+	+			+			+

++: Strong support +: Support -: negative impact

REFERENCES

Ackerman, M., & Cranor, L. (1999). *Privacy Critics: UI Components to Safeguard Users' Privacy.* In proceedings of the CHI '99 Extended Abstracts on Human Factors in Computing Systems (pp. 258-259).

Aggarwal, G., Feder, T., Kenthapadi, K., Khuller, S., Panigrahy, R., Thomas, D., et al. (2006). *Achieving anonymity via clustering.* In proceedings of the Symposium on Principles of Database Systems (pp. 153 - 162).

Agrawal, R., Kiernan, J., Srikant, R., & Xu, Y. (2003). *An XPath-based Preference Language for P3P.* In proceedings of the 12th Int'l World Wide Web Conference (pp. 629 - 639) Budapest, Hungary.

Anderson, A. (2005). *A Comparison of Two Privacy Policy Languages: EPAL and XACML*: Sun Microsystems Laboratories. from http://research.sun.com/techrep/2005/smli_tr-2005-147/ TRCompareEPALandXACML.html

APEC-FIP. (2004). *APEC Privacy Framework* (No. 2004/AMM/014rev1): Asia-Pacific Economic Cooperation.

Ashley, P., Hada, S., Karjoth, G., Powers, C., & Schunter, M. (2003). Enterprise Privacy Authorization Language (EPAL 1.2). W3C Member Submission 10 November 2003.

Ashrafi, N., & Kuilboer, J.-P. (2005). Privacy Protection via Technology: Platform for Privacy Preferences (P3P). *International Journal of E-Business Research, 1*(2), 56-69.

Bellotti, V., & Sellen, A. (1993). *Design for Privacy in Ubiquitous Environments.* In proceedings of the Third European Conference on Computer-Supported Cooperative Work (ECSCW'93) (pp. 77-92) Milan, Italy.

Blarkom, G. W. v., Borking, J. J., & Verhaar, P. (2003). PET. In G. W. v. Blarkom, J. J. Borking & J. G. E. Olk (Eds.), *Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents*: College bescherming persoonsgegevens.

Boyle, M., & Greenberg, S. (2005). The Language of Privacy: Learning from Video Media Space Analysis and Design. *ACM Transactions on Computer-Human Interaction (TOCHI), 12*(2), 328 - 370.

Brusilovsky, P., Kobsa, A., & Nejdl, W. (2007). *The Adaptive Web: Methods and Strategies of Web Personalization*. Heidelberg, Germany: Springer Verlag.

Buffett, S., Jia, K., Liu, S., Spencer, B., & Wang, F. (2004). Negotiating Exchanges of P3P-Labeled Information for Compensation. *Computational Intelligence, 20*(4), 663-677.

Burkert, H. (1997). Privacy-Enhancing Technologies: Typology, Critique, Vision. In P. E. Agre & M. Rotenberg (Eds.), *Technology and Privacy: The New Landscape* (pp. 126-143). Boston, MA: MIT Press.

Byers, S., Cranor, L., Kormann, D., & McDaniel, P. (2004, May, 2004). *Searching for Privacy: Design and Implementation of a P3P-Enabled Search Engine.* In proceedings of the 2004 Workshop on Privacy Enhancing Technologies (PET2004) (pp. 26-28) Toronto, Canada.

Camp, J., & Osorio, C. (2003). Privacy-Enhancing Technologies for Internet Commerce. In *Trust in the Network Economy*: Springer-Verlag (Berlin).

Cannon, J. C. (2005). *Privacy: What Developers and IT Professionals Should Know*: Addison-Wesley.

Chaum, D. (1982). *Blind Signatures for Untraceable Payments.* In proceedings of the CRYPTO 82 (pp. 199-203).

Clark, J., & DeRose, S. (1999). XML Path Language (XPath) Version 1.0, W3C Recommendation 16 November 1999. from http://www.w3.org/TR/1999/REC-xpath-19991116

ConneXion, C. (1996). Anonymous Surfing. from http://anonymizer.com

Coyle, K. (1999). P3P: Pretty Poor Privacy? A Social Analysis of the Platform for Privacy Preferences (P3P) [Electronic Version]. from http://www.kcoyle.net/p3p.html

Cranor, L. (2002). *Web Privacy with P3P*: O'Reilly.

Cranor, L., Dobbs, B., Egelman, S., Hogben, G., Humphrey, J., Langheinrich, M., et al. (2006). The Platform for Privacy Preferences 1.1 (P3P1.1) Specification. W3C Working Draft 10 February 2006. from http://www.w3.org/TR/2006/WD-P3P11-20060210/

Cranor, L., Langheinrich, M., & Marchiori, M. (2002). A P3P Preference Exchange Language 1.0 (APPEL1.0): W3C Working Draft 15 April 2002.

Cranor, L. F., & Reidenberg, J. R. (2002). *Can User Agents Accurately Represent Privacy Notices?* In proceedings of the 30th Research Conference on Communication, Information and Internet Policy, Alexandria, VA. from http://articles.ssrn.com/sol3/articles.cfm?abstractid=328860

CZ. (2000). Czech Republic Act of 4 April 2000 on the Protection of Personal Data and on Amendment to Some Related Acts (Vol. 101).

DE-TML. (2007). *German Telemedia Law.* from http://www.gesetze-im-internet.de/tmg/13.html

Dourish, P., & Anderson, K. (2006). Collective Information Practice: Exploring Privacy and Security as Social and Cultural Phenomena. *Human-Computer Interaction, 21*(3), 319-342.

Egelman, S., Cranor, L., & Chowdhury, A. (2006). An Analysis of P3P-Enabled Web Sites among Top-20 Search Results. In proceedings of the 8th International Conference on Electronic Commerce (pp. 197 - 207).

EPIC, & Junkbusters, Pretty Poor Privacy: An Assessment of P3P and Internet Privacy. (June 2000). from http://www.epic.org/reports/prettypoorprivacy.html

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. *Official Journal of the European Communities* (23 November 1995 No L. 281), 31ff.

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.

eXtensible Access Control Markup Language (XACML), Version 2.0; OASIS Standard, February 1, 2005. (2005). from http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml

FTC. (1998 June 1998). Federal Trade Commission. Privacy online: A Report to Congress.

FTC. (2000a). Online Profiling: A Report to Congress.

FTC. (2000b, May 2000). Privacy Online: Fair Information Practices in the Electronic Marketplace. A Report to Congress.

FTC. (2000c). The Seven Safe Harbor Principles.

Garfinkel, S., & Cranor, L. (2002). P3P: Privacy Primer.

Gideon, J., Cranor, L., Egelman, S., & Acquisti, A. (2006). *Power Strips, Prophylactics, and Privacy, Oh My!* In proceedings of the Second Symposium on Usable Privacy and Security (pp. 133-144) Pittsburgh, Pennsylvania.

Goldberg, I. (2002, April 2002). *Privacy-Enhancing Technologies for the Internet, II: Five Years Later.* In proceedings of the Workshop on Privacy Enhancing Technologies 2002.

Goldberg, I., Wagner, D., & Brewer, E. (1997). *Privacy-Enhancing Technologies for the Internet.* In proceedings of the 42nd IEEE Spring COMPCON San Jose, CA.

Gülcü, C., & Tsudik, G. (1996). *Mixing Email with BABEL.* In proceedings of the 1996 Symposium on Network and Distributed System Security (SNDSS '96) (pp. 2).

IBM. (2006). IBM Tivoli Privacy Manager for E-Business. 2006.

Kobsa, A. (2002). Personalization and International Privacy. *Communications of the ACM* (5), 64-67.

Kobsa, A. (2007). Privacy-Enhanced Web Personalization. In P. Brusilovsky, A. Kobsa & W. Nejdl (Eds.), *The Adaptive Web: Methods and Strategies of Web Personalization*: Springer-Verlag.

Kobsa, A., & Teltzrow, M. (2005). Contextualized Communication of Privacy Practices and Personalization Benefits: Impacts on Users' Data Sharing Behavior. In D. Martin & A. Serjantov (Eds.), *Privacy Enhancing Technologies: Fourth International Workshop, PET 2004, Toronto, Canada* (Vol. LNCS 3424, pp. 329-343). Heidelberg, Germany: Springer Verlag.

Langheinrich, M. (2002). *A Privacy Awareness System for Ubiquitous Computing Environments* In proceedings of the 4th International Conference on Ubiquitous Computing (pp. 237-245) Göteborg, Sweden.

Metzger, M. J. (2006). Effects of Site, Vendor, and Consumer Characteristics on Web Site Trust and Disclosure. *Communication Research, 33*(3), 155-179.

Microsoft. CardSpace [Electronic Version]. Retrieved 2006, from http://cardspace.netfx3.com/

Microsoft. (2000). Microsoft Announces Privacy Enhancements for Windows, Internet Explorer. Retrieved 21 June 2000, from http://www.microsoft.com/PressPass/press/2000/jun00/p3ppr.asp

OASIS. (2005). eXtensible Access Control Markup Language (XACML), Version 2.0; OASIS Standard, February 1, 2005. from http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml

OECD. (1980). Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data.

OpenID. OpenID: An Actually Distributed Identity System [Electronic Version]. Retrieved 2006, from http://openid.net

Palen, L., & Dourish, P. (2002). *Unpacking "Privacy" for a Networked World.* In proceedings of the CHI-02 (pp. 129-136) Fort Lauderdale, FL.

Paula, R. r. d., Ding, X., Dourish, P., Nies, K., Pillet, B., Redmiles, D. F., et al. (2005). In the Eye of the Beholder: A Visualization-based Approach to Information System Security. *International Journal of Human-Computer Studies, 63*(1-2), 5-24.

Preibusch, S. (2006). *Personalized Services with Negotiable Privacy Policies.* In proceedings of the PEP06, CHI 2006 Workshop on Privacy-Enhanced Personalization (pp. 29-38) Montreal, Canada.

Senicar, V., Jerman-Blazic, B., & Klobucar, T. (2003). Privacy-Enhancing Technologies: Approaches and Development. *Computer Standards & Interfaces, 25*(2), 147-158.

Stufflebeam, W., Anton, A. I., He, Q., & Jain, N. (2004). *Specifying Privacy Policies with P3P and EPAL: Lessons Learned.* In proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society (pp. 35-36).

Sweeney, L. (2002). K-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10*(5), 557-570.

Tavani, H., & Moor, J. (2001). Privacy Protection, Control of Information, and Privacy-Enhancing Technologies. *ACM SIGCAS Computers and Society, 31*(1), 6-11.

Teltzrow, M., & Kobsa, A. (2004). Impacts of User Privacy Preferences on Personalized Systems: a Comparative Study. In C.-M. Karat, J. Blom & J. Karat (Eds.), *Designing Personalized User Experiences for eCommerce* (pp. 315-332). Dordrecht, Netherlands: Kluwer Academic Publishers.

Tor. (2004). from http://tor.eff.org/

Tsai, J., Egelman, S., Cranor, L., & Acquisti, A. (2007). *The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study.* In proceedings of the Sixth Workshop on the Economics of Information Security, Pittsburgh, PA.

USA. (2002). The E-Government Act. from http://thomas.loc.gov/cgi-bin/bdquery/z?d107:HR02458:|TOM:/bss/d107query.html

USACM. (2006). USACM Policy Recommendations on Privacy. from http://usacm.acm.org/usacm/Issues/Privacy.htm

Wang, H., Lee, M. K. O., & Wang, C. (1998). Consumer Privacy Concerns about Internet Marketing. *Communications of the ACM, 41*(3), 63-70.

Wang, Y., & Kobsa, A. (2006). *Impacts of Privacy Laws and Regulations on Personalized Systems.* In proceedings of the PEP06, CHI06 Workshop on Privacy-Enhanced Personalization (pp. 44-46) Montréal, Canada.

Wang, Y., & Kobsa, A. (Forthcoming). Technical Solutions for Privacy-Enhanced Personalization. In Mourlas, C. & Germanakos, P. (Eds.), *Intelligent User Interfaces: Adaptation and Personalization Systems and Technologies*: IGI Global.

Warren, S., & Brandeis, L. (1890). The Right to Privacy. *Harvard Law Review, 4*(5), 193-220.

Westin, A., & Gelder, V. v. (2003). Privacy & American Business: Special Issue on Consumer Privacy in Japan and the New National Privacy Law. from http://www.privacyexchange.org/pab_japanissue.pdf

Whitten, A., & Tygar, D. (1999). *Why Johnny Can't Encrypt. A Usability Evaluation of PGP 5.0.* In proceedings of the Ninth USENIX Security Symposium. from http://www.cs.berkeley.edu/~tygar/papers/Why_Johnny_Cant_Encrypt/OReilly.pdf

Young, J. (1978). Introduction: A Look at Privacy. In J. Young (Ed.), *Privacy*. New York: John Wiley & Sons.

Key terms:

- Authentication: a process for verifying the digital identity of users or processes.
- Identity management: the management and provisioning of information about users across different applications (sometimes users may entertain different identities with partially different characteristics).
- Authorization: a process for verifying whether an identified user or role enjoys specific access rights to certain resources.
- Privacy policy language: a machine-readable language for expressing the privacy policies of organizations.
- Anonymity: the property that a user cannot be identified within the total user population, nor her interactions be tracked.
- Pseudonymity: the property that a user cannot be identified within the total user population, but her interactions nevertheless be tracked.

- P3P: a machine-readable (XML) language that allows websites to describe their privacy practices and P3P-enabled user agents (e.g., web browsers) to retrieve these privacy policies automatically and potentially analyze them.
- APPEL: a P3P preference language that allows users to express their privacy preferences.
- XACML: a general-purpose access control language that can be used to describe access control decision requests and responses as well as access control rules and policies.