PRIVACY PRINCIPLES FOR IDENTITY IN THE DIGITAL AGE

Draft for Comment - Version 1.2
Center for Democracy & Technology
March 2007

Introduction

How to create and manage individual identity is becoming a central challenge of the digital age. As identity-related initiatives are implemented in both the public and private sectors, individuals are being asked to identify themselves in some way with increasing frequency. A major goal of many of these efforts is to prevent illegal activity or enhance security, whether it be the security of our national borders, airplanes, workplaces, health records, or online transactions. Technologies – such as databases, machine-readable ID cards, and online accounts – are playing an ever more important role in these systems. Identity-related technologies can facilitate realization of the potential of the digital age, whether by making online transactions more seamless, tying together information on multiple devices, or enabling yet unimagined services.

However, the collection, storage, and disclosure of identity information involved in these systems can create risks to personal privacy and security. Poorly implemented identity systems can actually contribute to identity theft or weaken security. To mitigate these risks, it is essential that identity systems be designed with effective privacy and security measures. Incorporating such protections at the very beginning will help achieve the goals of identity systems.

In analyzing identity solutions offered in response to particular needs or problems, it is useful to understand identity as a spectrum, ranging from complete anonymity at one end to unique and full identity at the other. Some goals can be served (i.e., some transactions can be completed) without using any identity information at all, or by using minimally revealing forms of identity. Authentication operates on a similar scale – identity can be very strongly authenticated to the point where an identity claim can be verified with a high degree of reliability, or it can be weakly authenticated. The optimal strength of authentication will depend on the nature of the problem or transaction at hand.

Example

A purchase made with cash is an example of a transaction that does not use any identity information.

The choice of technology to implement an identity system carries significant privacy and security risks to both the organizations administering the system and the individuals participating in it. Thus, both public and private entities should first ask key questions: Is an identity system necessary for solving the problem at hand? If so, where should the system be located along the spectrum of identity? Does the authentication mechanism properly balance reliability with privacy protection?

This document outlines 10 privacy principles to guide government and commercial entities in developing programs or systems for the creation, authentication, and use of identity.[1] These principles deal strictly with systems for identifying individuals rather than groups or other entities.

The principles focus on privacy but also address security in certain instances. This is because privacy and security are interrelated and often must be considered together. When privacy is compromised, security of the individual, the organization or even the country is also threatened. Conversely, security breaches also lead to invasions of privacy.

This Version 1.2 is merely CDT's first iteration on the subject, and this document is open for public comment. CDT intends to convene stakeholders on all sides of this issue with the hope of achieving a comprehensive and useful set of guidelines or "best practices" that can be applied to the issues associated with identity creation and management across the public and private sectors and in many different contexts.

Principles

1 Proportionality

The uniqueness and reliability of an identity should be proportional to the purpose for which it is being created or used. An identity should only approach a person's unique and full identity as the significance of the purpose increases.

Example

An athletic club might print members' names and photos on club ID cards, but collecting fingerprints or other biometrics exceeds what may reasonably be considered necessary to achieve the goal of ensuring that only club members have access to the club.

The amount and type of information collected and stored by an identity system should be proportional to the purpose for which the identity is being created. For a transaction of lower significance, it is not appropriate to use a multitude of attributes or identifiers, or those that divulge much about a person's unique and full identity.

Not all transactions need to be tied to identity. Identity-based authentication should only be used when necessary. Identity systems relying on pseudonymous identifiers and authentication relying on anonymous attributes should be used whenever possible.

Example

The IRS may require individuals to authenticate themselves by providing their previous year's total income and a PIN number of their choice, both of which are anonymous attributes.

2 Diversity and Decentralization

Rather than attempt to serve as the perfect single solution, identity creation and authentication options should function like keys on a key ring, allowing individuals to choose the appropriate key to satisfy a specific need. Different government agencies, companies and organizations, and different types of functions within organizations, will likely need different types of identity systems.

[1] "Use" is considered to be any action besides creation and authentication. It includes storage of identity information.

Identity systems should be designed to exist in a marketplace offering multiple services that deliver varying degrees and kinds of identity creation, authentication, and use.

Example

Accessing a health record at a doctor's office will require a different authentication mechanism than accessing an email account at an Internet cafe.

As a single identity becomes more widely used and as identity information becomes more physically or logically centralized, there is increased likelihood for abuse by government, business, identity thieves, terrorists, and other criminals. Using only one or a very small handful of centralized identity solutions for multiple purposes leaves individuals with few choices and diminishes the ability of identity systems to protect privacy and security. Forcing individuals to use a single identifier or credential for multiple purposes puts their privacy and security at risk.

Creating a single identity for multiple purposes or housing identity information (or linked information) in a centralized location poses threats not only to personal privacy and security, but to national and sectoral security as well. Providing centralized access to decentralized data carries these same risks.

Example

Logical centralization occurs when data stored at multiple locations can be accessed from one location.

3 Individual Control and Choice

Individual controls are vital to building trust in identity systems. An identity system should offer individuals reasonable control and choice over the attributes, identifiers, and credentials that can be used within the system.

Individuals may choose to use a single credential or form of authentication that always discloses the same information for all interactions, but they should be able to choose to employ a variety of authentication tools for different transactions. This principle is particularly important in a system designed for both authentication and authorization, which will likely be successful only if it balances added convenience with trust in the system.

Example

Many online services allow a single individual to maintain multiple accounts. This allows individuals to use different identity information to interact with family, friends, and colleagues, for example.

Individuals should not be forced to accept the sharing of information for secondary uses as a condition of creating an identity.

4 Notice and Consent

Individuals should be provided with a clear statement about the collection and use of identifying information. Notice should be conspicuous and timely, and it should be provided in a manner appropriate to the technology being used.

Example

Providing a long, multi-page notice on a small cell phone screen would be considered inappropriate for the technology being used.

Individuals should be given the opportunity to consent to or decline the terms of the notice prior to any creation, authentication, or subsequent use of identity, identity information, and linked information. When possible, individuals should be able to consent to participation in an identity system but decline particular terms of the notice.

Individuals should be notified in situations where it may not otherwise be obvious that identities are being created for them. Prior to the creation of an identity, individuals should be notified of:

- The purposes for which the identity is being created;
- What identity information will be collected and how it will be used by the identity creator;
- How long the identity information will be stored by the identity creator;
- Whether and how the identity information or the created identity will be subsequently used by the identity creator or third parties;
- What other information will be linked to the identity and whether and how that information will be used by the identity creator or third parties;
- Whether individuals might need to authenticate themselves in the future and how to do so;
- How the individual will be able to access and correct information related to the identity; and
- How the individual may decline the creation of the identity.

Example

Information about shopping, travel, or Web browsing behavior may all be considered "linked information," as well as a record of each time a credential is used.

When identity systems make use of a technology that may be unfamiliar to participants in the system, notice should be provided about the presence of the technology and its privacy implications, in accordance with the items listed above.

Should new subsequent uses of identity information be developed after the identity is created, individuals should be notified in accordance with the items listed above and given the opportunity to consent to or decline such uses.

Example

Many individuals may be unfamiliar with RFID technology. They should be notified about how information about them can be linked to their identities through RFID.

Individuals should always be notified when other information is gathered about them and linked to their identity.

5 Limited Use

Identity information and linked information should be used only for specific, limited, and disclosed purposes.

Secondary use and sharing of identifiers or credentials can compromise privacy and security. In particular, identification numbers can become open to privacy misuses and security threats if they are used for secondary purposes. Therefore, multiple uses of such identifiers should be avoided.

Example

Bars and restaurants that swipe ID cards for age verification should not retain any identity information unless there is a legitimate, disclosed purpose for doing so.

Subsequent use of identity information and linked information, whether by the identity creator or a third party, should be minimized and disclosed. Identity information and linked information should be shared with third parties only when necessary, and should be stored by third parties only until the purpose for which it was shared has been completed.

The amount and type of data linked to an identity should be limited, and linking should only occur for specific, limited and disclosed purposes.

Government access to identity and linked information held by commercial entities should be allowed only upon service of legal process and pursuant to clear legal authority.

6 Onward Transfer

Any organization that handles identity information should include in its contracts provisions requiring that the entities with which identity information and linked information is shared will afford that shared data a level of protection consistent with or exceeding the organization's own standards, consistent with these principles.

7 Privacy and Security by Design

Privacy and security considerations should be incorporated into an identity system from the very outset of the design process. These include both safeguards for the physical system components and policies and procedures that guide the implementation of the system. Internal privacy and security practices should incorporate applicable regulatory and self-regulatory guidelines.

Example

People have difficulty remembering complicated passwords, so they choose passwords that are easy for others to guess. This human tendency should be central in deciding whether passwords are a strong enough authentication mechanism for the task at hand.

Identity systems should be designed with attention to human strengths and limitations that may impact the privacy and security of the systems. Knowledge of human behavior and how people will likely interact with an identity system should be incorporated from the first phases of a system's design.

Consistent with the principle of Limited Use, identity systems should be designed to make secondary uses difficult. Incorporating limits on the use of the system into its design will make "mission creep"[2] easier to avoid and less appealing later on.

Identity systems should be designed with consistent, robust interfaces so that individuals can learn to trust legitimate systems and distinguish them from fraudulent ones.

8 Security

Organizations that handle identity information should make reasonable and appropriate efforts to secure all technical, physical, and administrative components involved in the handling of the information. Such measures should cover credentials,

[2] Mission creep consists of authorized but initially unintended uses.

back-end systems that process and store identity information, personnel that handle the information, and physical facilities, among others. In so doing, organizations should establish and maintain an information security program in keeping with industry standards and applicable laws, appropriate to the amount and sensitivity of the information stored in their systems. Such a security program should include processes to identify reasonably foreseeable internal and external risks to the security, confidentiality, and integrity of identity information, and address those risks.

Identity systems that handle large amounts of identity information are by nature more vulnerable to tampering, loss, and unauthorized access (both internal and external). Adhering to strict security procedures should be a top priority for such systems.

The authentication mechanism used for internal access to an identity system should be at least as strong or stronger than the mechanism for external access by participants in the system.

Example

System administrators for a database of identity information may have to provide two biometric credentials for authentication while participants in the system are required to provide only one biometric credential.

9 Accountability

Organizations that handle identity information should be able to verify that they are complying with applicable privacy and security protections. Regular audits are necessary to ensure that reasonable technical, physical, and administrative privacy and security safeguards are being used. Personnel involved in handling identity information should be trained and educated about the privacy and security risks involved in dealing with identity and about applicable laws, guidelines, and procedures.

10 Access, Data Quality, and Due Process

Individuals should be provided reasonable access to the identity information and linked information that organizations maintain about them and use in the ordinary course of business. Individuals should be able to correct inaccurate identity information and linked information. This ability should be secured against unauthorized access.

The information should be easy to access, view, understand and change. Individuals should also be able to challenge conclusions drawn from identity and other information via structured and impartial administrative and judicial processes.

Access should either be provided by the identity creator or the organization interfacing with the individual, depending on the context.

Organizations should strive to ensure that the identity information they hold is timely, complete, and accurate.

Glossary

Italicized definitions are from the National Research Council's *Who Goes There? Authentication Through the Lens of Privacy*.[3]

Attribute. An attribute describes a property associated with an individual.

Authentication. Authentication is the process of establishing confidence in the truth of some claim.

Authorization. Authorization is the process of deciding what an individual ought to be allowed to do.

Credential. Credentials are objects that are verified when presented to the verifier in an authentication transaction. Credentials may be bound in some way to the individual to whom they were issued, or they may be bearer credentials. The former are necessary for identification, while the latter may be acceptable for some forms of authorization.

Identification. Identification is the process of using claimed or observed attributes of an individual to infer who the individual is.

Identifier. An identifier points to an individual. An identifier could be a name, a serial number, or some other pointer to the individual being identified.

Identity. The identity of X is the set of information about individual X, which is associated with that individual in a particular identity system Y. However, Y is not always named explicitly.

Identity Authentication. Identity authentication is the process of establishing an understood level of confidence that an identifier refers to an identity. It may or may not be possible to link the authenticated identity to an individual.

Identity Creation. Creation is the process by which an identity for individual X is established in identity system Y.

Identity Information: One or more attributes or identifiers used to create an identity.

Individual Authentication. Individual authentication is the process of establishing an understood level of confidence that an identifier refers to a specific individual.

Linked Information: Other facts about an individual, such as transactional shopping or travel behavior, tied to an identity.

Subsequent Use: Any use of identity, identity information, or linked information other than creation and authentication. Subsequent use may follow either creation or authentication of identity.

[3] National Research Council of the National Academies. *Who Goes There? Authentication Through the Lens of Privacy*. Eds. Stephen T. Kent and Lynette I. Millett. Washington: The National Academies Press, 2003.

Unique and Full Identity: An individual's true and complete identity comprised of a broad range of attributes or identifiers that can be used to distinguish the individual from every other person.