National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

*Privacy and Security in E-Learning**

El-Khatib, K., Korba, L., Xu, Y., and Yee, G.
October-December 2003

* published in International Journal of Distance Education. Volume 1, Number 4, October-December 2003. Idea Group Publishing. NRC 45786.

Canada

Privacy and Security in E-Learning

Khalil El-Khatib, Larry Korba, Yuefei Xu, and George Yee

Institute for Information Technology
National Research Council Canada
Montreal Road, Building M-50
Ottawa, Ontario K1A 0R6, Canada

{Khalil.El-Khatib, Larry.Korba, Yuefei.Xu, George.Yee}@nrc-cnrc.gc.ca

ABSTRACT

For a variety of advantages, universities and other organizations are resorting to e-learning to provide instruction on-line. While many advances have been made in the mechanics of providing online instruction, the needs for privacy and security have to-date been largely ignored. At best they have been accommodated in an ad-hoc, patchwork fashion. Privacy can be described as a learner's ability to maintain a "personal space" within which the learner can control the conditions under which personal information is shared with others. Security examines ways and means for implementing data integrity and protection policies for organizations involved with e-learning.

This paper examines privacy and security issues associated with e-learning. It presents the basic principles behind privacy practices and legislation. It investigates the more popular e-learning standards to determine their provisions and limitations for privacy and security. Privacy requirements for e-learning systems are explored with respect to the "Privacy Principles". The capabilities of a number of existing privacy enhancing technologies, including methods for network privacy, policy-based privacy/security management, and trust systems, are reviewed and assessed.

KEYWORDS

e-learning, distance education, information security, on-line privacy, privacy principles, network privacy, policy-based management, trust mechanisms

1 Introduction

One of the key characteristics of our information economy is the requirement for lifelong learning. Industrial and occupational changes, global competition, and the explosion of information technologies have all highlighted the need for skills, knowledge, and training. Focused on attracting and retaining staff, companies have placed an emphasis on training to bolster soft and hard skills to meet new corporate challenges. In many cases, career training has been placed in the hands of employees, with the understanding that employees must be able to keep ahead of technological change and perform innovative problem solving. One way of meeting the demand for these new skills (especially in information technology) is through on-line e-learning, which also offers the potential for continuous learning. Moreover, e-learning provides answers for the rising costs of tuition, the shortage of qualified training staff, the high cost of campus maintenance, and the need to reach larger learner populations.

From the corporate perspective, employee training is an approach to increase the level and variety of competencies in employees, for both hard and soft skills. On-line learning has become an important tool to implement corporate learning objectives. Indeed, specific e-learning courseware may be used to target specific corporate needs pertaining to strategic directions. Key trends for corporate e-learning, germane to privacy and e-learning include [1]:

- Learners may access courseware using many different computing devices and from different locations, via different networks.
- E-learning technology will overtake classroom training to meet the needs for "know what" and "know how" training.
- E-learning will offer more user personalization, whereas courseware will dynamically change based on learner preferences or needs. In other words, e-learning applications of the future will be intelligent and adaptive.
- Corporate training is becoming knowledge management. This is the general trend in the digital economy. With knowledge management, employee competencies are assets which increase in value through training. This trend has pushed the production of training that is more task specific than generic. Changes in corporate strategic directions are often reflected as changes in e-learning requirements prompted by the need to train staff for those new directions.
- E-learning is moving toward open standards.

Most e-learning innovations have focused on course development and delivery, with little or no consideration to privacy and security as required elements. However, it is clear from the above trends that there will be a growing need for high levels of confidentiality and privacy in e-learning applications, and that security technologies must be put in place to meet these needs. The savvy of consumers regarding their rights to privacy is increasing, and new privacy legislations have recently been introduced by diverse jurisdictions. It is also clear that confidentiality is vital for information concerning e-learning activities undertaken by corporate staff. While corporations may advertise their learning approaches to skills and knowledge development in order to attract staff, they do not want competitors to learn the details of training provided, which could compromise their strategic directions.

In this paper, we investigate the problem of privacy and security for distributed mobile e-learning systems. These kinds of e-learning systems provide service mobility, where the learner can access the learning content from anywhere, using any suitable device (e.g. desktop computer at home or work, PDA with wireless connection). We focus on the protection of personal information of a learner in an e-learning system. While it is an important issue in e-learning, we do not consider security issues related to copyright protection of course material. An overall theme of the paper is to highlight the privacy requirements for e-learning systems based on the so called "Privacy Principles" [2]. We explore the area of standards for e-learning systems and describe their deficiencies with respect to these privacy requirements. Finally, we describe several security and privacy enhancing technologies that can be applied to e-learning systems to satisfy the e-learning privacy requirements identified earlier. We do not claim that these technologies are the best fit to the requirements, only that they are *candidate* technologies to fulfill the requirements. We are currently engaged in research to identify the best fit (see Section 6).

The remainder of this paper is organized as follows: section 2 describes key "Privacy Principles" that underpin privacy practices and legislation. Section 3 investigates privacy and security issues among available e-learning standards. Section 4 examines e-learning system requirements for privacy and security using an architectural model for e-learning. Section 5 evaluates the more common Privacy Enhancement Technologies (PET), including W3C's P3P, network privacy approaches, policy-based technologies, and trust mechanisms. Section 6 offers conclusions and recommendations.

2 Privacy Principles

Incidents of privacy violation have led governments worldwide to raise privacy awareness for their citizens and to develop privacy legislation and policies to prevent exploitation of personal information. In countries where there is privacy legislation, individual control is required for the use of personal information, including the collection, use, disclosure, retention, and disposal of personal data by organizations that may handle that information. Privacy principles have been developed to expose the implications of either privacy laws or privacy policy adopted by on-line organizations. One way of assessing how well an application meets privacy requirements is to assess the application in light of the Privacy Principles. Table 1 briefly describes the ten Privacy Principles incorporated in the *Personal Information Protection and Electronic Documents Act* of Canada [2]. We will refer to these Privacy Principles in our analysis of the applicability of potential Privacy Enhancing Technologies (PET) for e-learning

applications. Generally speaking, while these principles are challenging to realize in any sector, they do offer a means for critiquing the appropriateness of different technologies [3].

Table 1: The ten Privacy Principles used in Canada.

Principle	*Description*
1. Accountability	An organization is responsible for personal information under its control and shall designate an individual or individuals accountable for the organization's compliance with the privacy principles.
2. Identifying Purposes	The purposes for which personal information is collected shall be identified by the organization at or before the time the information is collected.
3. Consent	The knowledge and consent of the individual are required for the collection, use or disclosure of personal information, except when inappropriate.
4. Limiting Collection	The collection of personal information shall be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means.
5. Limiting Use, Disclosure, and Retention	Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by the law. In addition, personal information shall be retained only as long as necessary for fulfillment of those purposes.
6. Accuracy	Personal information shall be as accurate, complete, and up-to-date as is necessary for the purposes for which it is to be used.
7. Safeguards	Security safeguards appropriate to the sensitivity of the information shall be used to protect personal information.
8. Openness	An organization shall make readily available to individuals specific information about its policies and practices relating to the management of personal information.
9. Individual Access	Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.
10. Challenging Compliance	An individual shall be able to address a challenge concerning compliance with the above principles to the designated individual or individuals accountable for the organization's compliance.

These Principles may be implemented in computer systems to varying degrees due to the nature of each principle. For example, Principle 1 is largely manual but portions of it can still be implemented to facilitate its compliance. The following suggests ways in which each principle may be "implemented":

1. Accountability: The name and contact information of the person who is accountable can be clearly advertised in the organization's online system.
2. Identifying Purpose: The purpose is clearly identified by the organization's online system and can be retrieved at will.
3. Consent: The person's consent is obtained by the organization's online system in the form of a signed certificate to guarantee authentication and non-repudiation.
4. Limiting Collection: The organization's system keeps secure logs of its data collection so that it can prove that it has complied with this principle if challenged; in addition, the organization's system identifies how it will collect the information to show that the collection will be fair and lawful.
5. Limiting Use, Disclosure, and Retention: The organization's system keeps secure logs of its uses, disclosures, or retention of the data so that it can prove that it has complied with this principle if challenged.
6. Accuracy: The system of the collecting organization can a) ask the individual providing the data to verify the data and sign-off on its accuracy and completeness, b) periodically request the individual to update his personal information, and c) run rule-based checks on the data to identify inconsistencies.
7. Safeguards: Security safeguards such as authentication and encryption can be implemented.

8. Openness: The organization's online system can advertise its policies and practices relating to the management of personal information as well as provide easily accessible links to this information.
9. Individual Access: The organization's online system can provide facilities for the individual to perform all access functions required by this principle.
10. Challenging Compliance: The organization's online system can provide a facility for the individual to address a compliance challenge to the person who has been identified as accountable by Principle 1.

3 Privacy and Security in Current E-Learning Standards

Emerging standards for distance learning and education will influence in a major way the development of on-line learning systems. Standardization and compatibility are vital for both e-learning vendors and end users to be able to sell or purchase portable content and inter-changeable components on the market. They are also very important where different e-learning systems must interact with one another.

There are currently a number of working groups seeking to develop industry-wide standards, including IEEE Learning Technology Standards Committee (IEEE LTSC) [4,5], IMS Global Learning Consortium (IMS GLC) [6], Aviation Industry CBT [Computer-Based Training] Committee (AICC) [7], Alliance of Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE) [8], and Advanced Distributed Learning- Sharable Content Object Reference Model (ADL-SCORM) [9]. Although these proposed standards mostly concern sharable components and learning objects, some of the suggested infrastructures and concepts are related to privacy and security requirements in e-learning systems. In the following subsections, we briefly review these standards for their privacy and security concerns and implications.

3.1 IEEE P1484

The IEEE P1484 is a standard for learning technology proposed by the Learning Technology Standards Committee (LTSC) of the IEEE Computer Society. The specification of Public and Private Information (PAPI) for Learners (P1484.2) outlines the syntax and semantics as well as the privacy and security of learner's information, which may be created, stored, retrieved, used, etc., by learning systems, individuals, and other entities. It defines the elements for recording descriptive information related to a learner's learning process, including personal contact information, learner relationships, learner preferences, learner performance, and portfolios. It categorizes the security and privacy concerns from the point of view of different stakeholders, such as developer, institution, regulator, and user.

The following table briefly shows the security related features of this standard.

Table 2: Security features defined in IEEE P1484

Model	Specification	Model	Specification
Session-View Security Model	D	Non-Repudiation Model	I
Security Parameter Negotiation Model	D	Repudiation Model	I
Security Extension Model	D	Privacy Model	N
Access Control Model	D	Confidentiality Model	N
Identification Model	I	Encryption Model	N
Authentication Model	O	Data Integrity Model	N
De-identification Model	O	Validation of Certificates	N
Authorization Model	I	Digital Signature Model	N
Delegation Model	I		

D - Defined: the model and/or requirements are defined or provided.
I - Implementation-dependent: the detailed methods are depended on detail implementations.
O - Outside the scope: the methods are outside the standard.
N - Non-specified: the standard doesn't specify the model and requirements.

As for privacy concerns, the P1484.2 does not specify a detailed model or technologies. It states that the implemented security techniques, including physical security, confidentiality, etc. can all be used to provide privacy. As well, it does not specify any particular privacy policy. The institutional administrators and users may act as privacy policy-makers to mandate policies, which are implemented via a variety of security techniques, technologies, processes, and procedures. A meaningful feature facilitating privacy protection is defined in the standard, which is called logical division of learner information . Using this feature, learner information may be de-identified, partitioned, and compartmentalized. Effectively, many privacy concerns for the learner may be addressed by virtue of this feature. More details can be found in the reference [5].

3.2 IMS LIP

The IMS Global Learning Consortium (IMS GLC) is another organization working on developing open specifications for distributed learning. It addresses key problems and challenges in distributed learning environments with a series of reference specifications, including Meta-data Specifications, Enterprise Specification, Content & Packaging Specification, Question & Test Specification, Simple Sequencing Specification, and Learner Information Package Specification. Among these, the IMS Learner Information Package (IMS LIP) Specification addresses the interoperability of learner information systems with other systems that support the Internet learning environment. It covers ways of organizing learner information so that learning systems can be more responsive to the specific needs of each user. Learner information is defined as the collection of information about a learner or learning producer. The typical sorts of learner information include education record, training log, professional development record, life-long learning record, and community service record (e.g. work and training experience).

The mechanisms for maintaining privacy and security of the learner information are enabled in the IMS LIP specification. A learner information server is responsible for exchanging learner's data with other information servers or other systems (e.g. a delivery system). The server will support an information owner defining what part of the information is shared with other systems. The packages that can be used to import data into and extract data from the learner information server are described in the specification.

The IMS LIP treats data privacy and integrity as essential requirements. However, the standard does not define any details of implementation mechanisms or architectures that could be employed to support learner privacy protection. The IMS LIP final specification V1.0 [10] does provide the following structures to support the implementation of "any suitable architecture" for learner privacy protection:

- The privacy and data protection meta-structure: within a learner information tree structure, each tree node and leaf has an associated set of privacy description, which defines the concerns of privacy level, access rights, and data integrity. The granularity of information is the smallest set of data where there is no further breakdown of independent privacy data.
- A "securityKey" data structure: the security keys for the learner include password, public key, and digital signatures. In this structure, the password and security codes are used for communication. The structure can allow for public key encryption, data authenticity, and password-based access control on learner information.

3.3 Other E-learning Standards

There are other standards or industry organizations working on specifications applicable for distance learning systems. These were mentioned at the beginning of this section and are: the Aviation Industry CBT [Computer-Based Training] Committee (AICC), the Alliance of Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE), and the Advanced Distributed Learning- Sharable Content Object Reference Model (ADL-SCORM). However, most of them are focusing on content management, meta-data specification, or other areas with little reference to security and privacy. For example:

- The AICC focuses on practicality and provides recommendations on e-learning platforms, peripherals, digital audio, and other implementation aspects.
- The ARIADNE focuses mainly on meta-data specification of electronic learning materials with the goal of sharing and reusing these materials.

- ADL-SCORM is mainly concerned with specifying how instructional content should be treated.

4 Privacy and Security Requirements for E-Learning

The roles of security include the following: user authentication /authorization, protection of private information from unintended access, and protection of data integrity (guarding against data corruption by attackers). We focus on requirements for privacy and data integrity. We begin by describing an architectural model for e-learning, taken from IEEE P1484.1/D9: the Learning Technology Systems Architecture (LTSA) [11]. We analyze this model as it applies to mobile, distributed e-learning with respect to the Privacy Principles and derive requirements for privacy and data integrity.

4.1 LTSA Architectural Model for E-Learning

The LTSA prescribes processes, storage areas, and information flows for E-Learning. Figure 1 shows the relationships between these elements. The solid arrows represent data flows (the thick arrows are explained below); the dashed arrows represent control flows. The overall operation is as follows: *Learning Preferences,* including the learning styles, strategies, methods, etc., are initially passed from the *learner entity* to the C*oach* process;the *Coach* reviews the set of incoming information, such as performance *history*, future *objectives*, and searches *Learning Resources*,via *Query*, for appropriate learning content; the *Coach* extracts *Locator*s for the content from the *Catalog Info* and passes them to *Delivery*, which uses them to retrieve the content for delivery to the learner as *multimedia*; *multimedia* represents learning content , to which the learner exhibits a certain *behaviour*; this behaviour is evaluated and results in an *Assessment* and/or *Learner Information* such as performance; *Learner Information* is stored in *Learner Records*; *Interaction Context* provides the context used to interpret the learner's behavior.

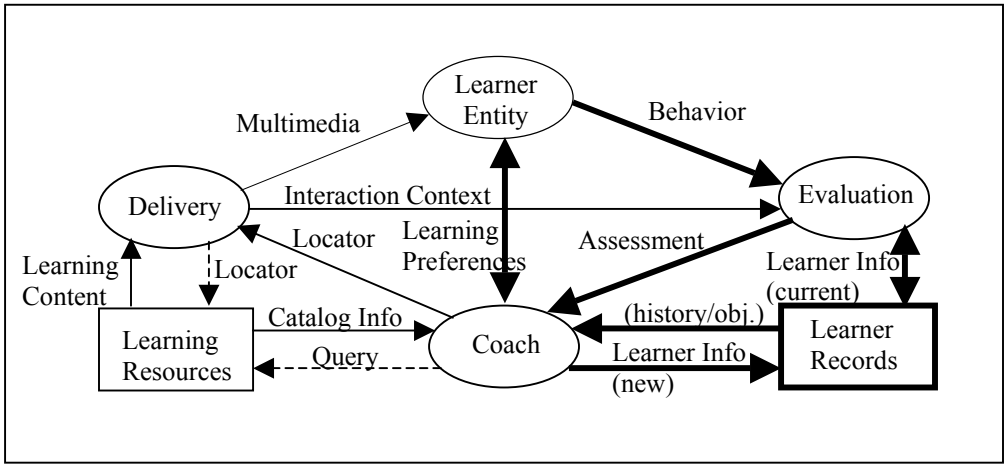


Figure 1: LSTA system components.

4.2 Fundamental Privacy Requirements

The Safeguards Principle requires security safeguards be placed on any E-Learning system component that is associated in anyway with private information. These components are highlighted with thick lines in Figure 1 and they are:

- the transmission channels between the Learner entity and both the Evaluation and Coach modules,
- the transmission channel between the Evaluation module and the Coach module,
- the transmission channel between Evaluation and Learner Records (service provider link),

- the transmission channels between Coach and Learner Records (service provider link), and
- the Learner Records themselves.

In case the learning content contains sensitive information, the transmission channels between Learner Entity, Delivery, Learning Resources, and Coach would also need to be protected leaving the Interaction Context channel as the only unprotected channel (one could argue that the Locator channels, the Catalog Info channel, and the Query channel may not need protection). Also in this case, the Learning Resources would have to be protected.

4.3 Network Privacy Requirements

With the open structure of the Internet and the readily available, easy-to-use tools for monitoring network activity, it is possible for a relative novice to extract vital information simply by analyzing the traffic patterns between the communicating entities. Some may consider that technologies such as secure sockets layer or virtual private networks would provide all of the safeguards one may require for network privacy. While these technologies may protect the data transferred between parties from network snoopers relatively well , a number of passive attack techniques can reveal sensitive information about the participating communicators [12]. Timing and communication pattern attacks, for example, extract information about the timing of communications, the locations of the communicating parties, and the amount of information being shared. By examining the pattern, timing, and origin and destination of communications, a snooper can deduce relationships between parties. For some activities in an organization, it is vital to safeguard this information. For instance, a company may have secretly chosen a new strategic initiative wherein specialized training is required for several key members of a development team. As per a recent trend, the company may have chosen to purchase a course from an online training company. In order to maintain confidentiality concerning the new strategic direction, the company would want to ensure that it would be very difficult for anyone to determine that it even has a relationship with the online training company. Indeed, the e-learning company itself may wish to distinguish its offerings from the competition by providing customers with the option of allowing students and employers to keep their network interactions confidential.

Referring again to Figure 1, all transmission channels that are used for communicating the *Learning Preferences, Behavior*, and *Multimedia* may be subject to traffic analysis and therefore counter-measures need to be in place to protect against these types of attacks. When the *Evaluation* process resides on the learner's machine, a protected transmission channel between the *Learner Entity* and the *Coach* can be used for both learning preferences and assessment information. The channel between the *Learner Entity* and the *Evaluation* process would not need protection any more.

4.4 Location Privacy Requirements

While some e-learning systems give learners the freedom to select the time and learning content according to their preferences and convenience, service mobility in e-learning offers learners additional freedom: a learner (see Learner Entity in Figure 1) can access the e-learning service anywhere using any available device. Wireless communication and device mobility compliment service mobility by delivering e-learning content to mobile computing devices, such as Personal Digital Assistants (PDA) and Internet-enabled cellular phones. Using these mobile devices, learners can receive e-learning content anywhere at any time, while traveling, commuting, or waiting in line.

Location privacy is of particular importance for mobile e-learning systems. With the convenience of delivering e-learning content to mobile devices, there is the potential of jeopardizing the location privacy of the learner. Some learners might be reluctant to reveal the location from which they are accessing e-learning content and consider this information private. Compiling this location information may provide useful information about the mobility pattern of the learner, which could be useful for a third party interested in the mobility of the learner.

5 Candidate PET for E-learning

In this section, we examine and critique a number of PET that can potentially satisfy privacy and security requirements for e-learning systems. We begin by looking at the Platform for Privacy Preference (P3P) [13], followed by approaches for network privacy. We next examine policy-based approaches for privacy/security management and go on to look at trust mechanisms. We end the section by describing the application of secure distributed logs.

5.1 Platform for Privacy Preference (P3P)

While a learner is using on-line learning services from an Internet website, he/she always has concerns about his/her privacy, such as:

- What information does the e-learning web site gather and for what purpose?
- Can the learner have access to the information related to his/her privacy?
- How long is this information kept?
- Is this information revealed to other companies and for what purpose?

The Platform for Privacy Preferences Project (P3P) [13], developed by the World Wide Web Consortium (W3C), provides a solution for answering these questions to some extent. It enables web sites to express their privacy policies in a standard format that can be automatically retrieved and interpreted by software acting on behalf of or under the control of a user (i.e. a user agent). P3P defines a machine-readable format (XML) for privacy policies. Web sites can post their privacy policies, and users can specify their privacy preferences in P3P format. . APPEL is a P3P exchange language that allows a user to express his preferences (rules) over the P3P policies. Based on these preferences, a user agent can make automated or semi-automated decisions regarding the acceptability of machine-readable privacy policies from P3P enabled Web sites. This allows P3P-enabled client software or user agents to retrieve web-site privacy policies and to compare them against the user's privacy preferences. If the user's privacy preferences are satisfied by the privacy policy of the web-site, then the user may proceed with the service; otherwise, the user might be warned that the web-site doesn't conform to his privacy preferences.

Although P3P allows web sites to express their privacy policy and notify users in a standard format, it is very limited with respect to current and emerging privacy practices and protection requirements. P3P falls short in fully supporting the Privacy Principles presented in Table 1 for the following reasons:

a) Limited coverage of privacy protection

As mentioned in Section 2, the *Personal Information Protection and Electronic Documents Act* [2] describes privacy rights with respect to personal information which are expressed as Privacy Principles. Regarding the Privacy Principles , P3P supports only the following three principles reasonably well:

- Identifying Purposes: The purposes for which personal information is collected are identified at or before the time the information is collected through the web browser.
- Consent: The individual's collection, use or disclosure of personal information are acknowledged. Consent is implicitly given when the user accepts the stated guidelines for a web site.
- Openness: Web site privacy policies on use and disclosure practices are open to public review.

The other 7 principles, including Accountability, Limiting Collection, Limiting Use/ Disclosure/ Retention, Accuracy, Safeguards, Individual Access, Challenging Compliance, are not addressed at all or are dealt with in a very weak manner in the P3P specification.

b) Lack of Privacy Policy Enforcement

P3P specification 1.0 states that it only provides a mechanism for ensuring that users can be informed about privacy policies before they release personal information. It does not provide a technical mechanism for ensuring that sites act according to their policies. The real guarantees on privacy are outside the scope of the P3P specification and depend upon specific implementations.

c) **Weak Model for Privacy and Security Protection**
Technically, P3P is a standardized set of multiple-choice questions. It is built upon the "notice and choice" privacy approach. Users are given notice of the privacy practice. If they do not like it, their choice is to leave the web site. This is a weak model for privacy and security protection.

The W3C's efforts on P3P are a positive contribution and a good beginning for privacy protection in the on-line environment. But P3P alone does not ensure strong privacy practices due to the weaknesses described above. Additional technical measures are needed to give people better control over the collection and use of personal information.

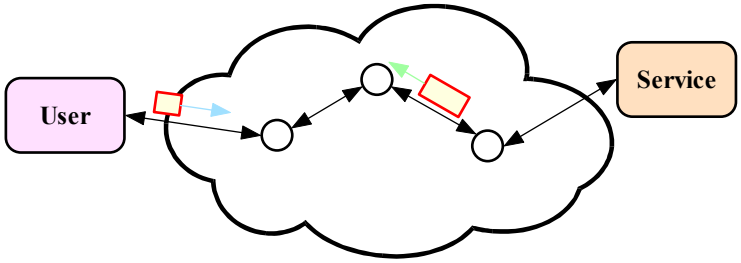
5.2 Approaches for Network Privacy

A number of approaches have been developed to provide the level of safeguarding for network privacy that is required in the company example above, in which the company needs to keep its relationship with the online training company hidden. One approach for web-based training is to use a proxy to redirect web requests [14,15]. When used in combination with secured communication channels, this approach may offer some privacy protection against casual attacks but it does have its drawbacks. It is vulnerable to timing and pattern attacks. As well, the access logs of these privacy services would provide a rich source of information concerning all users of the privacy service. Also, keeping all these logs in one location tempts hackers. Many organizations also may not want to trust a single proxy third party to protect its confidentiality and privacy. Other technologies have been developed to provide more robust privacy, such as Onion Routing [16], MIX Networks [17], DC-Net [18,19], Crowds [20] as well as commercial networks like the Freedom Network [21]. These approaches involve the deployment of a network of elements such as Chaum Mixes [17] for routing information between communicating parties. A single mix generally uses cryptographic packet tailoring techniques to hide the correlation between incoming and outgoing messages. A chain of mixes can be used to provide a more robust network privacy protection. Using a chain of mixes requires that routing at intermediate nodes be pre-determined statically by the source node, such as in the case of Onion Routing, or probabilistically by each intermediate node, as in the case of Crowds. An advantage of using multiple mixes is that these mixes are usually distributed under the control of multiple administrations in different jurisdictions so no single mix can compromise the user's privacy and collusion between mixes is not an easy task.

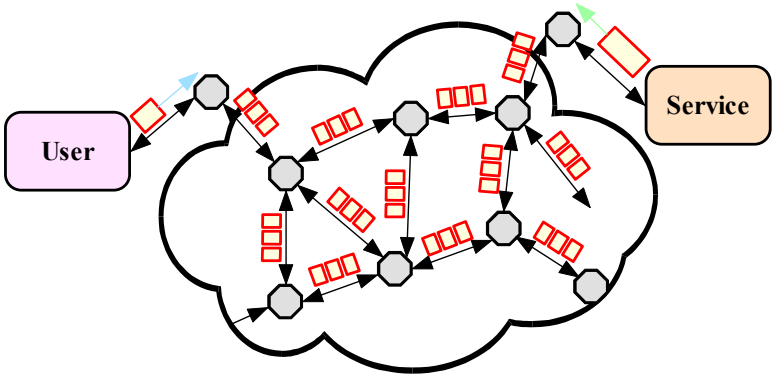
While network privacy techniques provide the required degree of anonymity, it achieves this with a certain cost. Techniques like Onion Routing incur setup overhead for each established connection. A larger delay for the data transfer is also incurred since the data is transferred along a path that may be different from the shortest path. This delay increases with the number of intermediate nodes along the path from the sender to the receiver. Cryptographic functions applied to the data in transit add more delay. This total additional delay may not affect the perceived quality of asynchronous applications such as e-mail or file transfer, but it is an issue for interactive applications such as videoconferencing. A balance needs to be established between the degree of anonymity and the perceived quality of the session.

Figure 2 illustrates two different situations. Figure 2a depicts the situation wherein a User connects over the public Internet to a Service using a conventional secured connection (VPN or SSL). Queries or data from the user are routed through various routers to a Service. It is clear that at any point along the route, various attack techniques may be used to determine the location (IP address) of the two parties and the nature of the interactions themselves.

In the case of a confidentiality network (Figure 2b), proxies at the User and Service sides modify the exchanged data so as to hide information using both cryptographic and traffic management techniques. The octagonal boxes in the network cloud represent MIX nodes that provide the cryptographic and traffic management functions. In this case, examining the traffic between any individual node pairs within the network cloud will reveal nothing about the nature and identity of the users or service.



a) In this case, only exchanges between a user and a service are encrypted. Unfortunately, communication patterns between users and a service may be attacked through traffic and timing analysis to reveal the nature and Internet addresses of the participants.



b) In the case of a MIX network, traffic and data from different users are mixed at each intermediate Mix node so as to make it difficult to determine the origin and destination of messages and the nature of the interactions.

Figure 2: The operation of a secured connection (a) as compared to a confidentiality network (b) between a User and a Service.

It is clear that not all e-learning clients will require the degree of privacy safeguarding offered by technologies such as onion routing. There will likely be varying degrees of requirements. Certainly, according to the Privacy Principles, network privacy is an important safeguard. However, offering a secure channel for exchange of information between the e-learning provider and the client may be adequate in most cases. Protecting network transmissions in a manner that would conceal location specific information and the nature of the online activities will be an important consideration for companies as they become more reliant upon third party e-learning vendors. In the near future, providing network-privacy approaches will be a differentiating factor among the offerings from different e-learning vendors.

5.3 Policy-based Approach for Privacy/Security Management

Policy-based management approaches have been used effectively to manage and control large distributed systems. In most policy-based management systems, policies are used to change the behavior of systems. Policies are usually expressed in terms of authorization and obligation imperatives over subject and object entities: authorization policies define the authorized and unauthorized actions of a subject over an object; obligation policies specify the positive and negative obligations of a subject toward an object.

As in any other distributed system, e-learning may also use a policy-based framework to manage the security and privacy aspects of operations upon objects in the system. To conform to the Privacy Principles introduced in Section 2, policies can be used to specify: Limiting Collection and Individual Access. Obligation policies can be used to

specify: Identifying Purpose, Consent (acquiring the user's consent for collecting data), supplying proof for limiting collection, Limiting Use/ Disclosure/Retention, Safeguards, and Openness.

In a policy-based e-learning system, the system administrator might specify some basic policies for the general operation of the system, and additional policies might be added based on the preferences of the entities. There would be sets of policies for each of the entities in the system (administrator, teacher, student, course material, …) as well as for the interaction between these entities. In addition, governments and other regulatory bodies may have privacy laws or regulations [3]. These may be translated into electronic policies and added to the general policies [35]. Conflicts might occur between these many policies. To streamline online activities, some sort of mechanism should be in place to detect policy conflicts and to resolve them. Thus, a facility for policy specification and negotiation would be beneficial for e-learning systems, where the e-learner and e-learning provider can identify policy conflicts and negotiate a resolution.

Interestingly, while a policy-based approach makes it possible to specify and manage privacy aspects of system operation, there is a challenge in implementing the actual controls within or around the objects themselves. Consider the principle of Limiting Collection. This principle may be readily expressed as obligation policies. Unfortunately, in implementation, limiting the extent of collection of personal information is difficult, if not impossible. For instance, an organization may specify that it will only collect names of students strictly for the purpose of managing record keeping during course execution. Yet it is difficult to imagine a system that would prevent collection of other information regarding the students' behavior during course execution, or the data mining of other information sources for further information about the user for any purpose the organization chooses. Indeed especially for the principles of Limiting Collection and Limiting Use, rather than automated means of compliance, trust and audit approaches are the most obvious recourse.

5.4 Trust Mechanisms

Like traditional face-to-face education, "trust" is an important concern in e-learning systems. In the context of networking and distributed applications, one system needs to be trusted to access another underlying system or service. Trusted interaction forms the underlying requirement between user and providers. For example, a service provider must trust that a learner truly has credentials that are not forged and is authorized to attend the course, or is limited to accessing only some services. On the other hand, the learner must trust the services. More importantly, the learner must believe the service provider will only use his/her private information, such as name, address, credit card details, preferences, and learning behavior in a manner expressed in the policy provided for the e-learning system user. The most common trust mechanisms are related to digital certificate-based approaches and are found in trust management systems.

a) Digital Certificate-based Mechanisms

These are based on the notion that "certificates represent a trusted party". The key concept behind these mechanisms is the Digital Certificate. A certification authority issues a Digital Certificate to identify whether or not a public key truly belongs to the claimed owner. Normally a certificate consists of a public key, the certificate information, and the digital signature of the certificate authority. The certificate information contains the user's name and other pertinent identification data; the digital signature authenticates the user as the owner of the public key. The most common approaches in use today are based on X.509/PKIX and PGP.

- X.509/PKIX [28] defines a framework for the provision of authentication services. This is a hierarchically structured PKI, and is spanned by a tree with a Root Certificate Authority (RCA). In this structure, the trust is centered at the root, and then transferred hierarchically to all the users in the network via Certificate Authorities (CA).
- PGP [29] Pretty Good Privacy (PGP) presents a way to digitally sign and encrypt information "objects" without the overhead of a PKI infrastructure. In PGP, anyone can decide whom he/she trusts. Unlike X.509/PKIX certificates, which come from a professional CA, PGP implements a mechanism called "Web of Trust", wherein multiple key-holders sign each certificate attesting the validity of the certificate.

The trust mechanisms based upon digital certificates, like X.509/PKIX and PGP, provide a series of systematic and comprehensive methods to define, verify, and manage trusted parties. These mechanisms have been proven to be good ways to establish one entity's credentials when doing transactions on the Internet. However, in these mechanisms, the user's confidence and trust depends on the authenticity of the public key. There are still however many uncertainties and risks that challenge certificate-based mechanisms [34]. For example, why and how can we trust a PKI vendor? There are also questions related to a vendor's authentication rules before issuing a certificate to a customer. In practice, this kind of mechanism needs to be adjusted to offer different types of security and privacy protection depending on the application, for both the user side and the service provider side. Some examples of such mature applications are PGP mail encryption and SSL-enabled connections based on PKI.

b) Trust Management Systems

Trust management systems have the goal of providing standard, general-purpose mechanisms for managing trust. Examples of trust management systems include KeyNote [30] and REFEREE [33]. Both are designed to be easily integrated into applications.

- KeyNote [30] provides a kind of unified approach to specifying and interpreting security policies, credentials, and relationships. There are 5 key concepts or components in this system.
  - 'Actions' -- the operations with security consequences that are to be controlled by the system;
  - 'Principals' -- the entities that can be authorized to perform actions;
  - 'Policies' -- the specifications of actions that principals are authorized to perform;
  - 'Credentials' -- the vehicles that allow principals to delegate authorization to other principals;
  - 'Compliance Checker' -- a service used to determine how an action requested by principals should be handled, given a policy and a set of credentials.

- REFEREE (Rule-controlled Environment for Evaluation of Rules and Everything Else) is a trust management system for making access decisions relating to Web documents, developed by Yang-Hua Chu based on PolicyMaker [31]. It uses PICS labels [32], which specifies some properties of an Internet resource, as the “prototypical credential”. It introduces the idea of "programmable credentials" to examine statements made by other credentials and fetch information from the network before making decisions.

Trust management systems provide a number of advantages for specifying and controlling authorization, especially where it is advantageous to distribute (rather than centralize) trust policy. Another advantage is that an application can simply ask the compliance checker whether a requested action should be allowed or not. However, although these trust management systems provide a more general solution to the trust management problem than public key certificate mechanisms, they mainly focus on establishing trust in resource access and possibly service provision. They still do not comprehensively cover the entire trust problem, and especially not the privacy concerns mentioned in Section 1. In e-learning, more tailored solutions or mechanisms are needed to fulfill the privacy and security requests from the learner and service provider.

5.5 Secure Distributed Logs

Secure distributed logs allow a record to be kept of transactions that have taken place between a service user and a service provider. The logs are distributed by virtue of the fact that they may be stored by different applications operating on different computers. Details of the transaction including the time of its occurrence, would be “logged” and the resulting record secured using cryptographic techniques, to provide assurance that their modification, deletion or insertion would be detectable. For e-learning, the use of secure distributed logs has important implications for privacy. In fact they support the Privacy Principles of Accountability, Limiting Use/Disclosure/Retention, and Challenging Compliance. In the case of Accountability and Limiting Use/Disclosure/Retention, the existence of a secured record of transactions allows verification that conformance to

each principle has been maintained. In the case of Challenging Compliance, the existence of a record is very useful for possibly showing where compliance has wavered.

6 Conclusions and Current Research

We have examined the Privacy Principles and investigated current e-learning standards for their privacy and security provisions. The Privacy Principles provide a basis for analyzing potential PET in terms of their capabilities to provide required privacy and security for e-learning. Current e-learning standards only treat privacy and security superficially, if at all. The LTSA architectural model for e-learning, IEEE P1484.1/D9, provides a high-level model of the components of an e-learning system. Together with the Privacy Principles, this model assists in identifying which components of an e-learning system require privacy or security safeguards. We identified such components in Section 4.2. We also looked at the requirements for network and location privacy. Existing technologies such as SSL or VPN fail to prevent traffic analysis attacks. Mobility for e-learners may lead to the need for location privacy.

We next examined a number of candidate PET for e-learning. As mentioned in the Introduction, these are only candidate PET and are not necessarily the best fit for the requirements. We are continuing our research to identify the best fit. Although P3P has some serious weaknesses with respect to privacy and security, it is a good starting point for online privacy protection. We overviewed a number of technologies for network privacy, including Onion Routing and Mixed Networks, which offer protection from traffic analysis attacks. Not all e-learning applications will require the stringent privacy offered by these privacy-enhancing networking techniques, but such levels of privacy are becoming increasingly important for more companies as they rely increasingly on third party e-learning vendors. We also looked at the policy-based approach for privacy and security management and identified how such an approach can satisfy the Privacy Principles. Finally, we examined trust mechanisms and described the use of secure distributed logs. Trust mechanisms provide for trusted interactions between a service user and a service provider. For e-learning, a trust management system can be used to set up authorizations for course access and learner privacy safeguards via policies, in conjunction with a policy-based approach to privacy and security management.

Table 3 provides a summary of our assessment of a variety of PET and indicates the degree to which they address the Privacy Principles.

We are continuing our research and development to improve privacy and security technologies for e-learning. Our focus is on the following areas:

- Network Privacy: technologies such as Onion Routing to protect from traffic analysis attacks;
- Location Privacy: technologies to ensure location privacy for mobile e-learners;
- Policy-based approach for privacy and security management: how to apply this approach to e-learning to satisfy the Privacy Principles; policy specification and negotiation mechanisms;
- Trust Mechanisms: how to apply this to e-learning to satisfy the Privacy Principles.

7 References

[1] H.W. Hodgins, “Into the Future: A Vision Paper”, commissioned by: American Society for Training and Development and the National Governor’s Association’s Commission on Technology and Adult Learning, February 2000 at: http://www.learnativity.com/download/MP7.PDF
[2] Department of Justice, Privacy provisions highlights, http://canada.justice.gc.ca/en/news/nr/1998/attback2.html
[3] Privacy Technology Review, http://www.health-canada.ca/ohih-bsi/available/tech/tech_e.html
[4] IEEE LTSC - Learning Technology Standards Committee, URL: http://ltsc.ieee.org /wg1/index.html
[5] IEEE LTSC PAPI - Public and Private Information (PAPI) for Learners, URL: http://ltsc.ieee.org/wg2/index.html; also available at http://edutool.com/papi
[6] IMS Global Learning Consortium, URL: http://imsproject.org

[7] AICC-Aviation Industry CBT [Computer-Based Training] Committee, URL: http://aicc.org
[8] ARIADNE - Alliance of Remote Instructional Authoring and Distribution Networks for Europe, URL: http://www.ariadne-eu.org/
[9] ADL-Advanced Distributed Learning, http://www.adlnet.org
[10] IMS Global Learning Consortium, Final Specification of IMS Learner Information Package Information Model, Version 1.0, 2001, available at: http://imsproject.org
[11] IEEE LTSC LTSA – Learning Technology Systems Architecture, URL: http://ltsc.ieee.org/wg1/index.html
[12] J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems", in H. Federrath, editor, Anonymity 2000, Volume 2009 of Lecture Notes in Computer Science, pages 10-29, Springer-Verlag, 2000.
[13] http://www.w3c.org/P3P
[14] Anonymizer web service at: http://www.anonymizer.com/
[15] L. P. W. Assistant, available at http://www.bell-labs.com/projects/lpwa
[16] D.Goldschlag, M.Reed and P.Syverson, "Onion Routing for Anonymous and Private Internet Connections", Communication of the ACM, vol.42, no.2, pages 39-41, 1999.
[17] D. Chaum, "Untraceable Electronic Mail, Return Address, and Digital Pseudonyms", Communications of the ACM, vol.24 no.2, pages 84-88, 1981.
[18] D. Chaum, "The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability", *Journal of Cryptology* 1/1 (1988), pp. 65-75.
[19] M. Waidner, "Unconditional Sender and Recipient Untraceability in Spite of Active Attacks", *Eurocrypt'89*, April 1989.
[20] M. K. Reiter and A. D. Rubin, "Crowds: Anonymity for Web Transactions", ACM Transactions on Information and System Security, v. 1, n. 1, pp. 66-92, 1998.
[21] P. Boucher, A. Shostack and I. Goldberg, "Freedom Systems 2.0 Architecture", Dec. 2000, available at http://www.freedom.net/info/whitepapers/Freedom_System_2_Architecture.pdf
[22] Thomas Beth, Malte Borcherding, et al, "Valuation of trust in open networks", in Proceedings of Computer Security –ESORICS '94, Brighton, UK, 2-9 Nov. 1994.
[23] Raphael Yahalom, Birgit Klein, et al, "Trust relationships in secure systems: a distributed authentication perspective", in Proc.1993 IEEE Computer Society Symposium on Research in Security and Privacy, pp.150-164, IEEE Computer Society Press, May 1993.
[24] Raphael Yahalom, Birgit Klein, et al, "Trust-based navigation in distributed systems", Computing Systems, pp.45-73, 1994.
[25] Gustavus J.Simmons and Catherine A. Meadows, "The role of trust in information integrity protocols", Journal of Computer Security, 1994.
[26] Pekka Nikander, Kristiina Karvonen, "Users and Trust in Cyberspace", Security Protocols, LNCS 2133, pp.24-35, Springer-Verlag, 2001
[27] Moti Yung, "Building Risk Management Into Commercial Electronic Transactions", http://www.certco.com/pdf/buildingriskmanagement.pdf
[28] Public-Key Infrastructure (X.509) (pkix), last modified: 11-Jan-02, http://www.ietf.org/html.charters/pkix-charter.html
[29] An Open Specification for Pretty Good Privacy (openpgp), last modified: 31-Jul-01, http://www.ietf.org/html.charters/openpgp-charter.html
[30] Matt Blaze, Joan Feigenbaum, John Ioannidis, and Angelos D. Keromytis, "The KeyNote Trust-Management System Version 2, Request For Comments (RFC) 2704", September 1999.
[31] M. Blaze, J. Feigenbaum, J. Lacy, "Decentralized Trust Management", Proceedings of the 17th IEEE Symp. on Security and Privacy, pp 164-173, IEEE Computer Society, 1996.
[32] P. Resnick and J. Miller, "PICS: Internet Access Controls without Censorship", Communications of the ACM, 39 (1996), pp. 87-93. Also available:http://www.w3.org/pub/WWW/PICS/iacwcv2.htm
[33] Y. Chu, "Trust Management for the World Wide Web", 1997, Massachusetts institute of Technology, REFEREE: Trust Management for Web Applications, http://www.w3.org/PICS/TrustMgt/presentation/97-04-08-referee-www6/
[34] C. Ellison, B. Schneier, "Ten risks of PKI: what you're not being told about Public Key Infrastructure", Computer Security Journal, V.XVI, N.1, 2000.
[35] L. Korba, "Privacy in Distributed Electronic Commerce", Proc. 35th Hawaii International Conference on System Science (HICSS), Hawaii, January 7-11, 2002.

Table 3: Privacy Principles and potential PET that may be developed for e-learning applications.

D: Direct support of a principle
I: Indirect or partial support of a principle
N: No support of a principle.

	Technology			
Principles	*Network Privacy*	*Privacy Policy Negotiation*	*Trust Mechanisms*	*Secure Distributed Logs*
1. Accountability.	*N*	*D*	*I*	*D*
2. Identifying Purposes	*N*	*D*	*N*	*I*
3. Consent	*N*	*D*	*N*	*I*
4. Limiting Collection	*I*	*I*	*N*	*I*
5. Limiting Use, Disclosure, and Retention	*N*	*I*	*N*	*D*
6. Accuracy	*N*	*N*	*I*	*I*
7. Safeguards	*I*	*I*	*I*	*I*
8. Openness	*N*	*D*	*D*	*I*
9. Individual Access	*I*	*I*	*I*	*I*
10. Challenging Compliance	*N*	*I*	*D*	*D*