Secure Knowledge Management

S. Upadhyaya
University at Buffalo, USA

H. Raghav Rao
University at Buffalo, USA

G. Padmanabhan
GE Transportation Systems, USA

INTRODUCTION

As the world is getting more and more technology savvy, the collection and distribution of information and knowledge need special attention. Progress has been made on the languages and tools needed for effective knowledge management1 and on the legal issues concerning the consumption and dissemination of critical knowledge. From a business perspective, a knowledge-management system (KMS) within a firm generally strives to maximize the human-capital utilization and profitability of the firm. However, security is becoming a major issue revolving around KMS; for instance, the KMS must incorporate adequate security features to prevent any unauthorized access or unauthorized dissemination of information. Acquiring the information that one needs to remain competitive while safeguarding the information one already has is a complicated task. Firms must balance the advantages of openness against its inevitable risks, and maximize the efficiency of electronic communication without making it a magnet for intruders. One must integrate offense and defense into a comprehensive strategy, and scholars have suggested that it is time to integrate intelligence and security imperatives with other knowledge-management strategies and processes (Barth, 2001).

Since the widely reported attacks on knowledge repositories in 2001 (e.g., Amazon was hit by denial-of-service attacks and the NIMDA virus hit financial markets), many organizations, especially the U.S. government, have increased their concern about KMSs. With the advent of intranets and Web access, it is even more crucial to protect critical corporate knowledge as numerous individuals now have access to the assets of a corporation. Therefore, we need effective mechanisms for securing data, information, and knowledge as well as the applications (Thuraisingham, 2003, 2004).

Security methods for knowledge-management systems may include authentication or passwords, cryptography programs, intrusion-detection systems, or access-control systems. Issues include insider threat (protecting from malicious insiders), infrastructure protection (securing against subversion attacks), and establishing correct policies, refinement, and enforcement. KMS content is much more sensitive than raw data stored in databases, and issues of privacy also become important (Thuraisingham, Chadwick, Olivier, Samarati, & Sharpston, 2002).

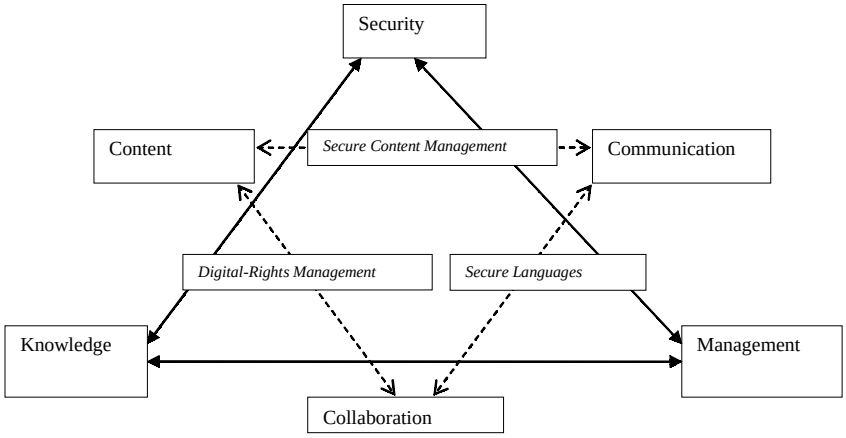
Asllani et al. (2003) surveyed over 300 knowledge managers about their job roles and found little or no evidence of security issues being considered in their jobs; their primary role was focused on communication within the organization. This article about secure knowledge management raises a number of issues in this critical area of research that need to be tackled by knowledge-management practitioners. The following sections focus on three important aspects of secure knowledge management: secure languages, digital-rights management (DRM), and secure content management (SCM).

BACKGROUND

A firm exists as a repository of knowledge over time (Zander & Kogut, 1995). Knowledge management is the methodology for systematically gathering, organizing, and disseminating information (Morey, Maybury, & Thuraisingham, 2003) in a firm. It essentially consists of processes and tools to effectively capture and share data, as well as use the knowledge of individuals within a firm. Knowledge management is about sharing information more freely such that firms derive benefit from such openness.

Secure knowledge-management (SKM) systems can be described in terms of the three *C*s: communication, collaboration, and content. SKM systems act as a gateway to the repository of intellectual content that resides within an organization. SKM systems need to source and/or provide access to knowledge that resides in multiple machines across an organization or multiple organizations for collaborative efforts. Secure languages are uti-

Figure 1. A framework for secure knowledge-management systems



lized to transfer information safely. At the same time, digital-rights management becomes critical in cross-organizational transfers of knowledge, while access control and identity management play an important role in securing the knowledge-management system. A framework for secure knowledge management is shown in Figure 1 as two interlinked, triangular chains: The larger chain focuses on security, knowledge, and management, while the smaller triangular chain (with dotted links) focuses on content, communication, and collaboration. Different aspects within the smaller chain include secure content management, digital-rights management, and secure languages. This article focuses on the interarticulation of the different concepts in the triangles.

SECURE LANGUAGES

In order to communicate securely and collaborate with one another, organizations need to use secure languages. These languages can be implemented to enhance the security of knowledge-management systems. Some of these are detailed in the following sections.

Security-Assertion Markup Language

The security-assertion markup language (SAML) can secure the KMS from insider or outsider threat by managing access control and identity. SAML is an extensible-markup-language- (XML) based framework (Cohen, 2000) for exchanging security information. In SAML, the expression of security is in the form of assertions about subjects. Most other security approaches use a central authority to authenticate the identity or the data. However, SAML does not use a central authority that authenticates the identity; it is up to the receiving application to accept if it trusts the assertion. The security-assertion markup domain model is depicted in Figure 2.

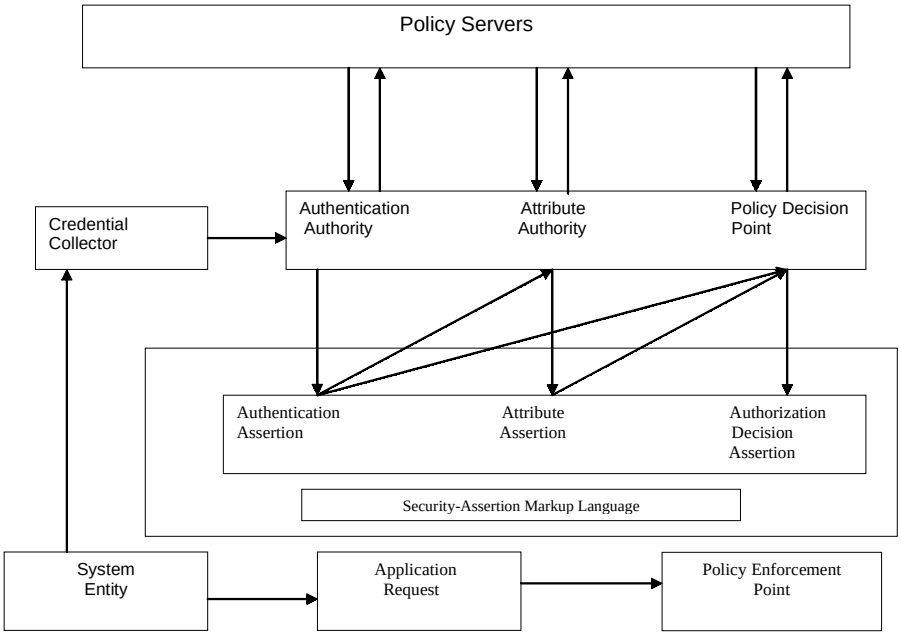
SAML shows how to represent users, identifies what data need to be transferred, and defines the process for sending and receiving authorization data (Cohen, 2000). SAML also has extensive applications in automated business-to-business (B2B) transactions that require secure transactions between the two parties. The increased collaboration among the various businesses has necessitated the need for such a technology (Patrizio, 2003). A case in point is that of Southwest Airlines (Wagner & Witty, 2003)—one of the first to use SAML-enabled identity management on a large scale to perform cross-domain trust. This implementation also marks an early step in the movement toward federated identity management.

SAML does not provide a complete security solution, but it does provide the identity-management functionality. In addition, it provides password management and access control, and a framework for implementing the "single sign-on" mechanism where authentication needs to be shared across multiple systems. Single sign-on becomes an absolute necessity when implementing complex KMSs that need to source or access data from multiple machines.

A number of commercial and open-source products provide SAML, including the following:

- Entegrity Solutions AssureAccess (http://www.entegrity.com/products/aa/aa.shtml)
- Internet2 OpenSAML (http://www.opensaml.org/)
- Netegrity SiteMinder (http://h71028.www7.hp.com/enterprise/cache/8258-0-0-225-121.aspx)

Figure 2. Security-assertion markup model



(Adapted from http://www.fawcette.com/xmlmag/2002_03/magazine/departments/marketscan/SAML/)

- RSA Security ClearTrust (http://www.rsasecurity.com/node.asp?id=1186)
- VeriSign Trust Integration Toolkit (http://www.xmltrustcenter.org/developer/verisign/tsik/download.htm)

Secure Knowledge-Query and Manipulation Language

KQML or the knowledge-query and -manipulation language is a language for exchanging information and knowledge. KQML focuses on an extensible set of performatives that defines the permissible operations that agents may attempt on each other's knowledge and goal stores. The performatives comprise a layer on which to develop higher level models of interagent interaction such as contract nets and negotiation. In addition, KQML provides a basic architecture for knowledge sharing through a special class of agents called communication facilitators, which coordinate the interactions of other agents. The ideas that underlie the evolving design of KQML are currently being explored through experimental prototype systems that are being used to support several test beds in such areas as concurrent engineering, intelligent design, and intelligent planning and scheduling (Lebrau, Finnin, Sherman, & Rabi, 1997).

An extension of KQML is secure KQML, which is being developed to take into account security and privacy concerns that agents could encounter whenever they cross multiple administrative domains. Since traditional agent communication-language standards lack the necessary constructs that enable secure cooperation among software agents, SKQML enables KQML-speaking agents to authenticate one another, implement specific security policies based on authorization schemes, and, whenever needed, ensure the privacy of the messages exchanged. SKQML employs public-key cryptographic standards and it provides security mechanisms as an integral part of the communication language. In summary, SKQML incorporates a synthesis of public-key certificate standards and agent communication languages to achieve an infrastructure that meets the security needs of cooperating agents.

B2B Circles of Trust

As can be seen from the discussion above, while the secure languages do allow secure communication to an extent, they are not complete solutions. An alternate mechanism for enhancing secure communication and collaboration across organizations in the knowledge-management environment has been termed "circles of trust."

Circles of trust involve two or more organizations sharing supplier or customer authentication information among themselves via a common interface or single sign-on capability. XML provides the basis for operating circles of trust (Varney, 2003).

One of the premier organizations espousing the concept of circles of trust is the Liberty Alliance—a consortium of more than 150 organizations working worldwide to create open, technical specifications for federated network identity. The alliance outlines the specifications for simplified sign-on capabilities using federated network-identity architecture. Permission-based attribute sharing is utilized to enable organizations to provide users with choice and control over the use and disclosure of personal information. A commonly accepted platform and mechanism for building and managing identity-based Web services is based on open industry standards. The Liberty Alliance specification addresses privacy and security concerns, and enables the participating organization to build more secure, privacy–friendly identity-based services that can comply with local regulations and create a trusted relationship with customers and partners (Varney, 2003).

DIGITAL-RIGHTS MANAGEMENT

The confluence of content and collaboration across organizations has brought up the concept of digital-rights management. DRM has traditionally focused on security and encryption to alleviate copyright-infringement and unauthorized-use problems. In order to do so, DRM techniques have implemented a mechanism to lock content and limit distribution to subscribed customers. Current DRM solutions include the description, identification, trading, protection, monitoring, and tracking of all forms of rights usages over both tangible and intangible assets including the management of rights holders' relationships (Ianello, 2001).

DRM systems are supposed to serve markets in which the participants have conflicting goals and cannot be fully trusted, yet need to collaborate and share knowledge content with each other. This adversarial situation introduces interesting new twists on classical problems studied in cryptology and security research, such as key management and access control (Feigenbaum, Freedman, Sander, & Shostack, 2002). Furthermore, novel business models and applications often require novel security mechanisms. Recent research has also proposed new primitives for DRM that make it possible to identify content in an adversarial setting.

Functional Architecture

The overall DRM framework suited to building digital-rights-enabled systems is illustrated in Figure 3. The functional architecture stipulates the roles and behavior of a number of cooperating and interoperating modules under the three areas of intellectual property (IP): asset creation, management, and usage (Figure 3).

The concept of intellectual-property asset creation and capture refers to the key question of how to manage the creation of content so it can be easily traded. This includes asserting rights when content is first created (or reused and extended with appropriate rights to do so) by various content creators or providers. The IP asset-creation and -capture module supports (a) rights validation to ensure that content being created from existing content includes the rights to do so, (b) rights creation to allow rights to be assigned to new content, such as specifying the rights owners and allowable usage permissions, and (c) a rights work flow to allow for content to be processed through a series of work-flow steps for review and/or approval of rights (and content).

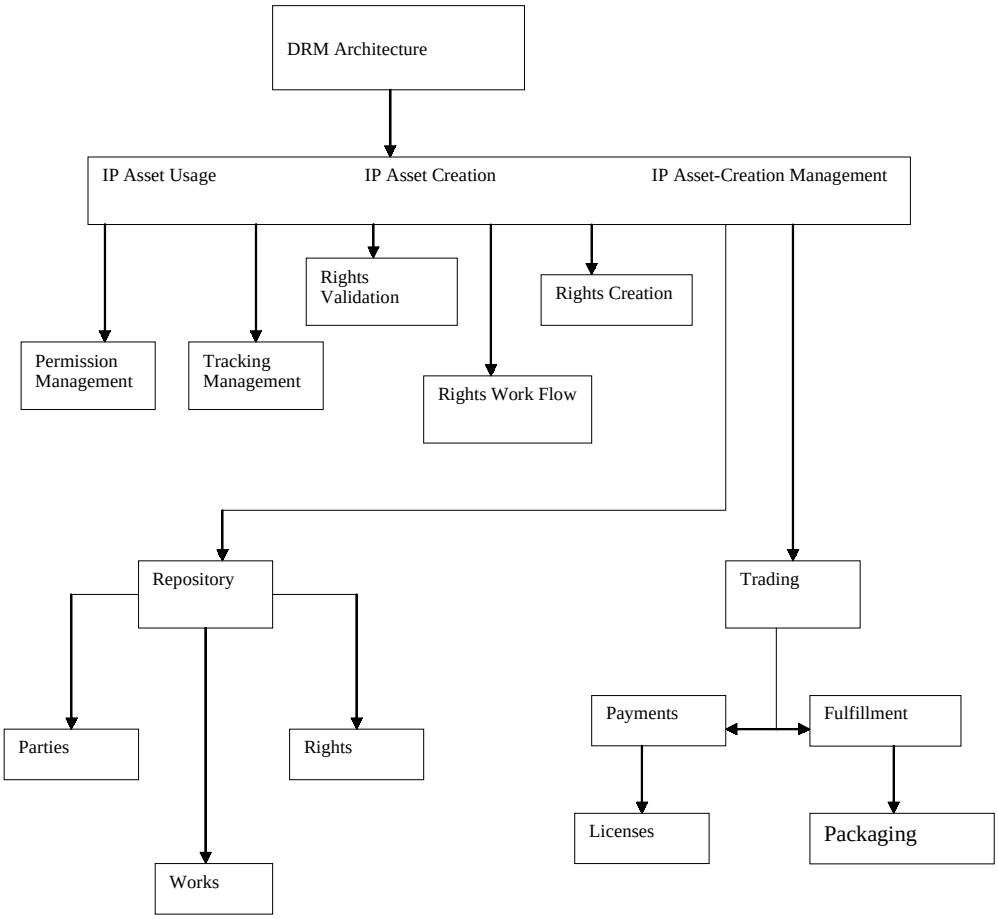
IP asset management involves the management and enabling of the trade of content. This includes accepting content from creators into an asset-management system. The trading systems need to manage the descriptive metadata and rights metadata (e.g., parties, usages, payments, etc.).

The IP asset-management module supports repository functions to enable the access or retrieval of content in potentially distributed databases and the access or retrieval of metadata. The metadata cover information regarding parties, rights, and descriptions of the work. The module also supports trading functions that enable the assignment of licenses to parties who have traded agreements for rights over content, including payments from licensees to rights holders (e.g., royalty payments). In some cases, the content may need to go through fulfillment operations to satisfy the license agreement. For example, the content may be encrypted, protected, or packaged for a particular type of desktop usage environment.

Once the IP asset has been traded, this module focuses on how to manage the usage of content. This includes supporting constraints over traded content in specific desktop systems or software.

The IP asset-usage module supports permissions management to enable the usage environment to honor the rights associated with the content. For example, if the user only has the right to view the document, then printing will not be allowed. It also allows tracking management to enable the monitoring of the usage of content where such

Figure 3. Digital-rights-management architecture



tracking is part of the agreed-to license conditions (e.g., the user has a license to play a video 10 times; Iannella, 2001)

SECURE CONTENT MANAGEMENT

The final link in the secure knowledge-management chain is the one that links content and communication, that is, secure content management. The Internet is a tremendous tool for enterprises to share intellectual property with customers, partners, and suppliers. It is an instant distribution network any corporation can use to improve communications while lowering operating costs (Ogren, 2003). The Yankee Group estimates that the market for secure content-delivery products and services amounted to $302 million in 2002 and will grow to $580 million by 2007. It is widely believed that more destructive and harder-to-detect threats, spam, legal liability, employee productivity, and compliance with privacy regulations will continue to fuel the growth of the secure content-management market over the next several years (http://www.csoonline.com/analyst/report1490.html).

The Internet, instant messaging, and the availability of Web content have transformed everyday business activities (Robb, 2003). As a result, CIOs (chief information officers) and IT management are increasingly looking for solutions to help enforce corporate policy, comply with privacy regulations, limit legal liability, increase employee productivity, and reduce network bandwidth consumption. All this is made possible by secure content-management solutions.

Secure content-management tools help to correctly label business-related content. The first generation of SCM products is now beginning to appear on the market. Generally, they consist of the following features: antivirus capabilities, proactive identification to block only malicious code, smart filtering of spam and URLs (uniform resource locators), keyword identification to safeguard against the transmission of proprietary and confidential information via e-mail, and centralized management of all facets to bring simplicity to the task of security administration (Robb, 2003).

Secure Content Delivery

With the advent of the Internet, content that enterprises once closely guarded in private databases is now being placed on the Internet to save distribution costs throughout the supply chain and to increase customer satisfaction. A Web initiative can take multiple forms: for example, an employee portal or Web-enabled self-service partner extranet. Each such initiative involves delivering business value. The Web has been instrumental in expanding communication channels and providing endless opportunities. Globalization has led to increased collaboration among trading partners that require the sharing of confidential information. The quest for cost-effective solutions for secure content delivery is intense since it must not only ensure the privacy of the electronic customers, but also reliably deliver important information only to designated recipients.

The trend has been to centralize identity management and documents in secure server repositories and portals accessed by browsers, and to avoid the complexities of client-side software installations. Content in transit has traditionally been protected by secure-sockets-layer (SSL) communications for browsers, and virtual private networks (VPNs) for application access, encrypted e-mail, and proprietary application solutions.

CONCLUSION

We are moving into a knowledge-based economy in the 21st century. Knowledge-based assets are gaining in importance, and it is becoming extremely important to protect these assets. In the area of national security, the knowledge that must be shared comes from many fields including homeland-defense activities, tactical intelligence missions, diplomatic channels, and direct military support. A range of KMS approaches and technologies and their security features need to be examined to enable critical intelligence gathering. Critical issues in secure knowledge management include content, communication, and collaboration. In this context, SAML, SKQML, circles of trust, DRM, secure content management, and secure content-delivery mechanisms would ensure the security and privacy of knowledge repositories.

REFERENCES

Barth, S. (2001). Open yet guarded: Protecting the knowledge enterprise. *Knowledge Management Magazine*.

Cohen, F. (2003). Debunking SAML myths and misunderstandings. *IBM Developer Works*.

Dingledine, R., Freedman, J. M., & Molnar, D. (2001). The free haven project: Distributed anonymous storage service. *International Workshop on Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability*, (pp. 67-95).

Doukidis, G., Mylonopoulos, N., & Pouloudi, N. (in press). In C. Dellarocas (Ed.), *Information society or information economy? A combined perspective on the digital era*. Hershey, PA: Idea Book Publishing.

Feigenbaum, J., Freedman, M., Sander, T., & Shostack, A. (2002). Privacy engineering in digital rights management systems. In *Lecture notes in computer science: Vol. 2320. Proceedings of the 2001 ACM Workshop on Security and Privacy in Digital Rights Management.* Berlin, Germany: Springer.

Feigenbaum, J., & Miller, E. (2002). Taking the copy out of copyright. In *Lecture notes in computer science: Vol. 2320. Proceedings of the 2001 ACM Workshop on Security and Privacy in Digital Rights Management.* Berlin, Germany: Springer.

Iannella, R. (1997). Digital rights management (DRM) architectures. *D-Lib Magazine, 7*(6).

Labrou, Y., Finnin, T., Sherman, A., & Rabi, M. (1997). *A proposal for a new KQML specification.* Baltimore, MD: Computer Science and Electrical Engineering Department, University of Maryland.

Morey, D., Maybury, M., & Thuraisingham, B. (2003). *Knowledge management: Classic and contemporary works.* MIT Press.

Ogren, E. (2003). Secure content delivery protects shared, transmitted and post-delivery digital assets. *CSO*.

Patrizio, A. (2003). *SAML advances single sign-on prospects.* Fawcette Technical Publications.

Robb, D. (2003). The emergence of secure content management. *IT Management*.

Schirmer, A. L. (2003). Privacy and knowledge management: Challenges in the design of the Lotus Discovery Server. *IBM Systems Journal, 42*(3).

Thuraisingham, B. (2003). Data mining and cyber security. *Proceedings of the Third International Conference on Quality Software (QSIC'03)*.

Thuraisingham, B. (2004). Cybertrust, data and applications security. *Proceedings of the Secure Knowledge Management Workshop,* Buffalo, NY.

Thuraisingham, B., Chadwick, D., Olivier, S. M., Samarati, P., & Sharpston, E. (2002). Privacy and civil liberties. In *IFIP Conference Proceedings* (256). Kluwer.

Van den Heuvel, S. A. F. A., Jonker, W., Kamperman, F. L. A. J., & Lenoir, P. J. (2002). *Secure content management in authorized domains*. The World's Electronic Media Event IBC.

Varney, C. (Ed.). (2003). *Privacy and security best practices*. Liberty Alliance Project.

Wagner, R., & Witty, R. (2003). *Southwest Airlines shows SAML's promise*. Gartner Research.

Zander, U., & Kogut, B. (1995). Knowledge and the speed of the transfer and limitation of organizational capabilities: An empirical test. *Organization Science, 6*.

KEY TERMS

Digital-Rights Management: DRM is a platform to protect and securely deliver content on a computer.

IP Asset Management: This involves management and enabling the trade of content, and includes accepting content from creators into an asset-management system.

Knowledge Management: Knowledge management is the methodology for systematically gathering, organizing, and disseminating information. It essentially consists of processes and tools to effectively capture and share data as well as use the knowledge of individuals within a firm.

Secure Content-Delivery Space: Content that enterprises once closely guarded in private databases is now being placed on the Internet to save distribution costs. Hence, content has to be delivered securely. The mechanisms that allow this form the secure content-delivery space.

Secure Knowledge Management: The management of knowledge while adhering to principles of security and privacy. Enterprises must find cost-effective solutions to ensure the privacy of electronic customers, reliably deliver important information only to designated recipients, and offer revenue-generating services based on access profiles.

Secure Knowledge-Management Trends: The trend has been to centralize identity management and documents in secure server repositories and portals accessed by browsers.

Security-Assertion Markup Language: SAML is an XML-based framework for exchanging security information. This security information is expressed in the form of assertions about subjects (either human or computer) that have an identity in some security domain. Assertions can convey information about authentication acts and authorization decisions about whether subjects are allowed to access certain resources.