

By JinKyu Lee, Shambhu J. Upadhyaya, H. Raghav Rao,
and Raj Sharman

SECURE KNOWLEDGE MANAGEMENT AND THE SEMANTIC WEB

Strengthening security within the domain of shared knowledge is a critical issue, and great challenge, to businesses today. A number of different protocols currently available offer an array of benefits and limitations.

Knowledge has long been recognized as an essential prerequisite for quality decision making [9]. An individual's knowledge can be transferred to others through different modes, such as socialization, externalization, combination, internalization—mechanisms that create and expand organizational knowledge [7]. Knowledge management (KM) can be viewed as a class of managerial processes that

creates strategic value through knowledge creation, storage/retrieval, transfer, and application processes [7]. Knowledge management systems (KMS) are IT-based systems that can help organizations manage their knowledge by supporting those KM processes [1].

Is it possible to capture tacit and implicit knowledge that resides within individuals into an organizational knowledge base, and transfer

Illustration by Gianpaolo Pagni

the organizational knowledge to other members who may be separated by time and space? If so, how? While some deep tacit knowledge may be extremely difficult to articulate, codify, and transfer, most knowledge exists on a continuum of tacitness, and transformation of implicit knowledge to explicit knowledge is, in many cases, a matter of effort to verbalize its terminology and rules [4].

Berners-Lee's vision of the Semantic Web, where information flow is significantly enhanced by machine-processable metadata [2], envisages a new generation of KMS that can foster knowledge transfer, both implicitly and explicitly [8]. Semantics-enabled KMS (hereafter Semantic KMS) can allow multiple groups of knowledge engineers and users, within or across organizational boundaries, to build and share organizational knowledge [3].

SECURITY AND SEMANTIC KMS

As KM has become a more central part of organizational activities and dependent upon technologies, securing organizational knowledge has become one of the most important issues in the KM area [6]. When groups of individuals who must share knowledge are distributed across different places and times, it is expected their need to transform implicit knowledge into explicit knowledge and to share the articulated knowledge with other group members will increase [4].

In conjunction with heavy reliance on information and communication technologies in today's distributed environment, such knowledge externalization efforts will result in digitalized taxonomies and related rules that can easily be stored and transferred by KMS [10]. Therefore, Semantic KMS can capture more articulated organizational knowledge that would otherwise have remained as tacit knowledge within an individual, and the externalized once-tacit explicit knowledge can now easily be transferred to collaborators or be amenable to theft by competitors. This implies that strategic competence from organizational knowledge is dependent largely upon knowledge security. To safeguard knowledge against theft, secure knowledge management is a necessity. When an organization fails to protect its externalized organizational

knowledge from theft, the organization will lose its competitive advantage.

Knowledge is now resident in metadata models and connections between different pieces of informa-

| KMS User Group/Role | Description | Annotated Data |
|--|---|--|
| Internal Knowledge Engineer <ul style="list-style-type: none"> Design and manage knowledge classifications. Capture and organize knowledge | Design and Integrate versioning, confidentiality classification, and rules.* | Encrypt/sign highly confidential information in annotate data repository (XMLEnc, XMLDSig) |
| Internal User <ul style="list-style-type: none"> Create and share the source of organizational knowledge. Use organizational knowledge. | Have partial and indirect access through KMS services on the intranet (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) | Encrypt/sign annotated data she creates, revises, and allows to share (XMLEnc, XMLDSig) Access annotated data through KMS services on the intranet (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) |
| External User <ul style="list-style-type: none"> Group of knowledge users outside of a principal organization Collaborate with internal users to develop and share organization knowledge. | Have no access to knowledge partners' conceptual schemas. Indirectly access mapping services (within or outside of the organization) to access knowledge partners' annotated data (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) | Access knowledge partners' annotated data through inter-organizational KMS services (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) Encrypt/sign annotated data she provides to knowledge partners (XMLEnc, XMLDSig). |
| External Knowledge Engineer <ul style="list-style-type: none"> Collaborate with Internal knowledge engineer to develop interoperable KMS. | Have no direct access to knowledge partners' live ontologies Securely communicate with partners' knowledge engineers in creating ontologies.** | Same as external users |

* Readers are referred to Knowledge Control System (KCS) in Ontology Middleware Module developed in On-To-Knowledge Project; www.ontoknowledge.org/and [3].

** Traditional security technologies (S/MIME, PGP, SSH, VPN, among others) can be used for message/file transfer.

Table 1. Examples of security technologies applied to the case scenario.

tion made by an organization to gain strategic advantage in a competitive world. Although all data and information systems in organizations must be protected by using authentication/authorization, cryptography, intrusion detection/prevention, and access control mechanisms, particular attention must be paid to protecting strategic knowledge resources [11]. Here, we describe some technologies that can help organizations protect their knowledge in Semantic KMS from knowledge theft.

TECHNOLOGIES FOR SECURE KM ON THE SEMANTIC WEB

We focus here on those technologies likely to be considered industrial standards¹ in the Semantic Web area.² To help readers contextualize the security technologies within Semantic KMS, we use the following case scenario:³

Many institutional lenders (for example, mortgage or auto loan lenders) need an insurance-tracking service in

¹We introduce only those technologies recommended or being evaluated by W3C, OASIS, IETF, and IBM/MS consortium.

²For more detail, see www.ninebynine.org/SWAD-E/Security-formats.html#toc#toc or msdn.microsoft.com/library/default.asp?url=/library/en-us/dnwssecur/html/securitywhitepaper.asp.

³This scenario is built on an insurance-tracking Web service case; www.microsoft.com/biztalk/evaluation/casestudies/casestudy.asp?CaseStudyID=15096.

order to make sure their loan customers have enough insurance to protect the lender's interest in the collateral and, if not, purchase additional insurance on behalf of the borrower. For insurance-tracking service providers, the capability to handle a large volume of data from insurance carriers, adequately analyze the risk of collaterals, and provide accurate assessment results to the lenders is critical organizational knowledge that must be securely protected.

Let's assume insurance-tracking company (A) and mortgage company (B) are sharing knowledge through two interconnected Semantic KMS.

- S1. A risk analyst (knowledge engineer) in company A develops a better way to calculate the risk associated with a type of collateral. She modifies the business rules for risk calculation.
- S2. A home insurance company informs company A that a home insurance contract was terminated. Company A identifies the property is in the collateral list of company B's mortgage loan and triggers a risk analysis process.
- S3. The risk analysis process issues queries for all the relevant information, some from internal KMS and some from company B.
- S4. Company B as well as company A's internal KMS responds to queries from the risk analysis process.
- S5. The risk analysis process concludes the insured value of the house is less than the collateral value and informs company B of the risk. Company B purchases additional insurance for the house on behalf of the home owner.

Table 1 presents some examples of the security technologies applied to this scenario.

ENCRYPTION AND DIGITAL SIGNATURE FOR CONFIDENTIALITY AND INTEGRITY

As emphasized earlier, strategic organizational knowledge should be secret and kept away from competitors. Encryption technologies protect the confidentiality of the knowledge stored in a knowledge base or distributed over the network. Because the Semantic Web uses XML syntax as the primary building block, most data stored in or distributed over the network must be in XML format, and this requires an encryption mechanism that can work with XML messages. Although Secure Sockets Layer (SSL) and Transport Layer Security (TLS) are widely used in the HTTP protocol, which also provides the primary communication channel for Web services (such as SOAP), these point-to-point protocols are

not suitable for the Semantic Web where a Web service request can trigger multiple additional requests to some other Web services. In this scenario, each connection between a pair of intermediaries requires a new SSL/TLS session, and the proof of the original requestor's identity and authorization credentials established with the initial Web service cannot be propagated to the end points.

XML Encryption (XMLEnc) is a W3C standard that specifies how to encrypt/decrypt a XML-formatted data object. XMLEnc supports end-to-end (as opposed to point-to-point protocols) encryption of a XML object (whole or a part of a XML document), which can be transmitted in XML or non-XML format. This technique can be used in various stages in Semantic KMS, including knowledge storage, internal/external knowledge transfer, and authentication.

When organizational knowledge is captured, stored, and reused, the users of the knowledge must know who provided the knowledge and if the knowledge has been modified by a third party. A digital signature provides a mechanism by which the user of knowledge can verify its authenticity (such as, it is indeed from the purported author) and integrity (it has not been tampered with) of the knowledge. As in the XMLEnc case, the Semantic Web requires a digital signature mechanism for XML objects.

XML Digital Signatures (XMLDSig), also called XML Signatures, is an IETF/W3C joint standard that specifies how to digitally sign and verify a signature of a XML data object. XMLDSig enables digital signatures on arbitrary digital content (XML or non-XML) within a particular view to XML content. Like XMLEnc, XMLDSig can be used in many phases in Semantic KMS (for example, authenticity verification for retrieved/updated knowledge and involved intermediaries, among others). XMLEnc and XMLDSig, however, should be used with care as these techniques significantly increase the volume and process overhead. This necessitates that only a limited amount of highly confidential knowledge be encrypted permanently in a knowledgebase, while other recipient-specific confidential knowledge should be encrypted and signed upon transmission or replaced by a N/A flag. The accompanying figure depicts two organizations sharing knowledge through interoperable Semantic KMS. The arrows represent interactions between various Semantic KMS components, including automated agents, knowledge engineers, domain ontology, knowledge users, and so on.

PUBLIC KEY MANAGEMENT, AUTHENTICATION, AND AUTHORIZATION FOR ACCESS CONTROL

When a Semantic KMS user accesses an organiza-

tional knowledge base that user should be authenticated and/or her digital signatures should be validated by a Web service. These processes often utilize a public key infrastructure (PKI), which requires a complex and non-XML syntax communication. Key management services can alleviate this error-prone procedure by making PKI-using Web service applications a client of key management services.

XML Key Management Specification (XKMS) is a W3C protocol specification that describes how to distribute and register public keys. XKMS consists of two components: XML Key Information Service Specification (X-KISS) and XML Key Registration Service Specification (X-KRSS). X-KISS describes how to verify public key information contained in a XML message, and X-KRSS describes how a Web service registers public key information. These specifications encapsulate key information/registration processes within XML syntax, thus making applications using PKI free from the complex, non-XML syntax trust establishment processes. XKMS, in conjunction with XMLDSig and XMLEnc, can work with various PKI specifications, including X.509/PKIX, SPKI or PGP (www.w3.org/TR/xkms).

IBM and Microsoft also offer a similar technology. Web Services Trust Language (WS-Trust)⁴ is a model they are developing and specifies how to establish trust relationships directly or indirectly (via intelligent agents and Web services intermediaries) by using security token issuance services. WS-Trust also describes how to allow delegation and impersonation.

When two or more agents/Web services communicate, they must make sure the other party has the right to see what they request. Extensible Access Control Markup Language (XACML) is an OASIS⁵ specification that describes how to impose control over access policies and authorization mechanisms. XACML determines appropriate response to user requests, using specified rules, policies, and/or policy sets to evaluate the requester's attributes, the protocol

used in the request, the type of requested activities, and the range of possible input. Web Services Policy (WS-Policy),⁶ also being developed by IBM and Microsoft, is similar to XACML in that it defines the requirements and capabilities in communication with intermediaries and endpoints. WS-Policy also specifies how to associate service policies with SOAP mes-

| KMS User Group/Role | Description | Annotated Data |
|--|---|--|
| Internal Knowledge Engineer <ul style="list-style-type: none"> Design and manage knowledge classifications. Capture and organize knowledge | Design and Integrate versioning, confidentiality classification, and rules.* | Encrypt/sign highly confidential information in annotate data repository (XMLEnc, XMLDSig) |
| Internal User <ul style="list-style-type: none"> Create and share the source of organizational knowledge. Use organizational knowledge. | Have partial and indirect access through KMS services on the intranet (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) | Encrypt/sign annotated data she creates, revises, and allows to share (XMLEnc, XMLDSig) Access annotated data through KMS services on the intranet (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) |
| External User <ul style="list-style-type: none"> Group of knowledge users outside of a principal organization Collaborate with internal users to develop and share organization knowledge. | Have no access to knowledge partners' conceptual schemas. Indirectly access mapping services (within or outside of the organization) to access knowledge partners' annotated data (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) | Access knowledge partners' annotated data through inter-organizational KMS services (SAML, XACML, WS-Policy/Privacy, XKMS, WSS) Encrypt/sign annotated data she provides to knowledge partners (XMLEnc, XMLDSig). |
| External Knowledge Engineer <ul style="list-style-type: none"> Collaborate with Internal knowledge engineer to develop interoperable KMS. | Have no direct access to knowledge partners' live ontologies Securely communicate with partners' knowledge engineers in creating ontologies.** | Same as external users |

* Readers are referred to Knowledge Control System (KCS) in Ontology Middleware Module developed in On-To-Knowledge Project: www.ontoknowledge.org/and [3].

** Traditional security technologies (S/MIME, PGP, SSH, VPN, among others) can be used for message/file transfer.

Table 2. Security needs in Semantic KMS.

sages. Some basic requirements and capabilities include privacy attributes, encoding formats, security token requirements, and supported algorithms.

Security Assertions Markup Language (SAML) is an OASIS standard language that specifies how to describe authentication, authorization, and other information, and how to bind the transportation protocol. The contents of a SAML message are determined by the policy it communicates (for example, XACML), and the values of the contents influence the policy-based decision. This framework allows Web service components to exchange security information without a predefined authorization message. This information, expressed in the form of assertion, usually includes information about subjects (human or machine), authentication acts, and authorization decisions (whether allowed to access certain resources). SAML is also a key enabler for single sign-on (SSO) (for example, Microsoft Passport), where the initial sign-on Web service must endorse the validity of the user's authentication, authorization, and attribute information to other Web service intermediaries.

⁴msdn.microsoft.com/library/en-us/dnglobspec/html/ws-trust.pdf.

⁵OASIS is a not-for-profit, international consortium dedicated to the development, convergence, and adoption of e-business standards; www.oasis-open.org.

⁶msdn.microsoft.com/Webservices/default.aspx?pull=/library/en-us/dnglobspec/html/ws-policy.asp.

INTEGRATION AND COMMUNICATION

Originally submitted by IBM and Microsoft, Web Services Security (WSS) is an OASIS specification that offers message integrity, confidentiality, and single-message authentication mechanisms. WSS is basically an extension of SOAP, an envelope framework for XML messages. By attaching one or more security header blocks to SOAP, WSS passes the security mechanism-related information to multiple receivers. However, those security mechanisms need other Web service extensions and higher-level application-specific protocols in order to implement various security models and security technologies. The security-related information that can be included in the security header blocks are security tokens, endorsement of claims, and verifiable proof of possession of the tokens. Security tokens can be a username/password, Binary tokens (for example, X.509 certificates or Kerberos tickets), XML tokens (for example, XML signatures or SAML assertions), and so on.

Endorsement of claims is left open to various XML signature specifications. Designed to work with XML Signature, XML Encryption, and many other security mechanisms (for example, XKMS), WSS provides a means to integrate authentication, privacy, and authorization mechanisms into a SOAP-based application framework. Security needs and relevant security technologies for knowledge repositories are summarized in relation to various Semantic KMS users groups in Table 2.

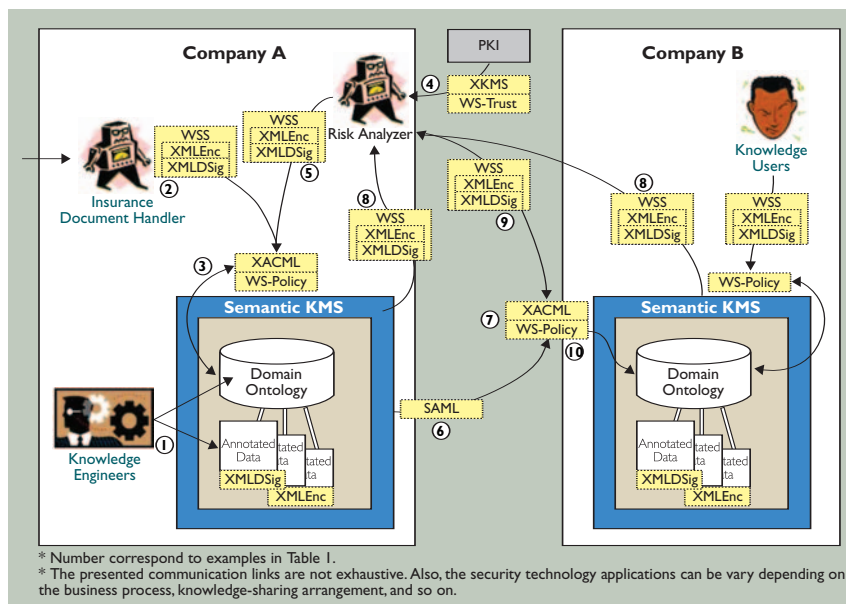
CONCLUSION

We have presented some prominent security technologies for Semantic KMS. Although applicable to non-KMS Web services [5] as well, these technologies are particularly important to knowledge-intensive businesses. It is also important to note that interoperability is a key for Semantic KMS, and thus standardization is a critical necessity.

Although some of the security technologies presented here are proprietary protocols (for example, WS-Trust and WS-Policy), we expect they will be submitted to a standardization body (OASIS, W3C) or further developed as a compatible extension of a

standard protocol. Thus, rather than wondering about which brand of technology to use, businesses may want to focus more on designing security rules and developing knowledge-sharing processes.

There are several other open issues worthy of further investigation in the area of more secure KMS, particularly the area of knowledge security. Knowl-



Security technologies for interoperable KMS on the Semantic Web.

edge classification and description logic often include very valuable organizational knowledge (for example, S1 in our case scenario) and should be protected.

A related issue is the role allocation between IT staff and non-IT staff. Although research on role-based access control deals with dynamic and static separation of roles in assigning security permissions, the permission assignment rules themselves can be integrated into knowledge systems, which would then be managed by non-IT oriented knowledge engineers. An important factor is the inferred/reconstructed knowledge problem. When more than two parties share a knowledge base, nuggets of knowledge—each of which do not undermine confidentiality—prove more revealing when combined. This is an emergent property in sharing knowledge. For example, let's assume a consulting firm disclosed three prices for successful project cases (x, y, z) to three different potential customers, where each of the previous projects was manned by two consultants A&B, B&C, and A&C. When the three customers decide to share their knowledge, they can calculate each of the consultant's prices. In this case, the consulting firm stands to lose. Research to prevent such situations has started to appear [12] and would be of great practical significance to the industry. Such studies may include the

investigation of properties in knowledge-sharing rules, rules to block or restrain emergent knowledge, or zero-knowledge proofs. These will lead to significant contributions in the area of Semantic Web-enabled KMS. **C**

REFERENCES

1. Alavi, M. and Leidner, D.E. Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Q.* 25, 1, (2001), 107.
2. Berners-Lee, T., Hendler, J., and Lassila, O. The Semantic Web. *Scientific American* 284, 5 (May 2001), 34–43.
3. Davies, J., Fensel, D., and van Harmelen, F. (Eds.). *Towards the Semantic Web: Ontology-Driven Knowledge Management*. John Wiley, Chichester, West Sussex, U.K., 2003.
4. Griffith, T.L., Sawyer, J.E., and Neale, M.A. Virtualness and knowledge in teams: Managing the love triangle of organizations, individuals, and information technology. *MIS Q.* 27, 2 (2003), 265–287.
5. Kim, D.J., Agrawal, M., Jayaraman, B., and Rao, H.R. A comparison of B2B e-service solutions. *Commun. ACM* 46, 12 (Dec. 2003ve) 317–324.
6. King, W.R., Marks, P.V., and McCoy, S. The most important issues in knowledge management. *Commun. ACM* 45, 9 (Sept. 2002), 93–97.
7. Nonaka, I. A dynamic theory of organizational knowledge creation. *Organization Science* 5, 1 (1994), 14–37.
8. Singh, R., Iyer, L., and Salam, A.F. Semantic ebusiness. *Intern. J. on Semantic Web IS* 1, 1 (2005) 19–35.
9. Thuraisingham, B. Keynote: Secure knowledge management. Presented at the the Workshop on Secure Knowledge Management, Buffalo, NY, 2004.
10. Tidwell, L.C. and Walther, J.B. Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations.

Human Commun. Research, 28, 3 (2002), 317–348.

11. Upadhyaya, S., Rao, H.R., and Padmanabhan, G. Secure knowledge management. *Encyclopedia of Knowledge Management*. D. Swartz, Ed. ISWorld, 2005.
12. Xu, S. And Zhang, W. PBKM: A secure knowledge management framework. Presented at the Workshop on Secure Knowledge Management, Buffalo, NY, 2004.

This research has been supported by NSF under grant 0423014.

JINKYU LEE (jklee2@buffalo.edu) is a Ph.D. candidate in the School of Management at the State University of New York, Buffalo, NY.

SHAMBHU J. UPADHYAYA (shambhu@cse.buffalo.edu) is an associate professor in the department of computer science and engineering at the State University of New York, Buffalo, NY.

H. RAGHAV RAO (mgmtrao@buffalo.edu) is a professor in the School of Management and adjunct professor in the Department of Computer Science and Engineering at the State University of New York, Buffalo, NY.

RAJ SHARMAN (rsharman@buffalo.edu) is an assistant professor in the School of Management at the State University of New York, Buffalo, NY

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

© 2005 ACM 0001-0782/05/1200 \$5.00

STAY ON TOP OF ACM NEWS WITH **MEMBERNET**

NOW IN THE CURRENT ISSUE:

- Reports on ACM's OOPSLA and Supercomputing conferences.
- New books and courses come to the Professional Development Centre.
- A new awards program from ACM's Committee on Women and Computing honors women in technology.

And much more!

All online, in MemberNet: www.acm.org/membernet.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.