Persistent Tagging
Enforced through Encryption

Introduction

The management of information flow can be achieved through tagging techniques and tagging protocols. Over time, tagging has taken on different means such as labeling, asserting and credentialing. The intent is to define information flow and control. Different approaches have surfaced to address labeling such as identity tags in communications routing and access distribution. More recently, shifts in computer protocols such as from HTML to XML have expanded the possibilities of using tags.

Data that is affected by these tagging techniques and protocols can be more persistent in that data management can be viewed from the creation of the data, the processing of the data, the storage of the data, and finally the expulsion of the data. The tag persists with the data through its life cycle. Data in this context is more knowledge based and takes on the properties of content with its expanded usage of tags. (See section The Significance of Information)

A significant tagging protocol is XML. The XML protocol is a markup language that is designed to allow an XML designer to describe stored data via custom-defined tags. XML generally includes one or more data elements. Each of these data elements is provided with at least one respective tag that specifically describes the particular data element or group of elements. For example, a data element can be "12341234156", and its respective tag can be "<credit card data>." When stored as plaintext, the data element alone might not be readily identifiable or usable by an unauthorized party. However, when viewed with its descriptive tag, the data element's defined nature is known, and the risk of an undesirable disclosure or use is significantly increased compared with that of untagged data elements.

From a security perspective, encryption can be used to reduce the risk of undesirable disclosure. Encryption can be leveraged to define access control limits that enhance tagging. Tagging can effectively separate data into predefined security information boundaries that further can be enforced through encryption. Also, tagging with encryption can be used to extend the boundaries of the concepts of 'data in transit' (in motion, for example, over the Internet, Intranet, LAN, WAN, Satellite, etc.) and 'data at rest' (stored, for example, on servers, PCs, laptops, PDAs, in databases, on USBs, memory sticks or CDs, etc.).

The strategy of implementing encryption is also shifting to accommodate the advances in tagging. The perimeter defense concepts that have focused on primarily protecting the physical or transport layer (Open System Interconnection (OSI) model) can be broadened to protect the data itself, at the object layer. Protecting the object or content allows for (1) the movement of the data via whatever transport mechanism is available, and (2) the storage of data wherever it is convenient, as well as (3) the access to the information controlled and accessible only to selected parties. Access to the data objects can be enforced with encryption in a

role-based access control scheme.   Only users who have the access rights to the object can view and/or manipulate the data.   Similarly, data can only be re-purposed by authorized users.   The definition of content protection can also be extended to protecting virtual folders.   In this context, the folder is a mechanism to visualize and manipulate information.

To have a high assurance that the content is continuously being protected from creation through transmission and storage, as well as while being manipulated and finally destroyed, the content along with its tagging scheme needs to be bound to the encryption methodology.   From a security system perspective, access to the content needs to be defined through the concepts of Authentication, Authorization, and a Trusted Platform.   Encryption can be a thread that further binds the security process to the content.   Encryption can have a generalized role in each of the concepts: PKI for Authentication, Constructive Key Management® (CKM) for authorization and CKM® for enforced data separation within a Trusted Platform.

PKI can enforce access to information as a single entity and necessitates a user or entity validation process to complete a high assurance process. CKM® can enforce access to information through roles or one-to-many distribution as can be illustrated in a publish-and-read architecture.   The enforcement tools of CKM are *Credentials* that bind the content and an encryption process.   Both PKI and CKM® include management tools to address the key management cycle associated with managing keys.   Both encryption technologies can complement one another, or be used in an independent context.   The encryption processes of PKI and CKM® have matured sufficiently to now focus on object or content protection.

TecSec Incorporated
www.tecsec.com
Copyright 2014

Page 2
Proprietary

June 16, 2010
SR_15_001_Persistent Tagging
All Rights Reserved

The Significance of Information

Information Theory is the application of mathematical principles to the problems of transmitting and storing information. In 1948, Bell Labs scientist Claude Shannon developed Information Theory, and the world of communications technology has never been the same. Information Science is the interdisciplinary academic field that deals with the generation, collection, organization, storage, retrieval, and dissemination of information. Without going into an in-depth discourse on Information Science and Theory, it is helpful to establish a few definitions before getting started.

In the computing world, information is defined as computer data that has been organized and presented in a systematic fashion to clarify the underlying meaning, while data is information for computer processing. For example, numbers, text, images, and sounds, in a form that is suitable for storage in or processing by a computer[1]. A simple example of the difference between data and information follows:

1234567.89 is data.

"Your bank balance has jumped 8087% to $1,234,567.89" is information.

And finally, content is information that is made available by an electronic medium or product.

Consider information; whether an engineering or accounting description of carrier service, production input, intermediate product, or processed output. This point is easily seen through the conceptual conflicts between information and data. From one standpoint, organizing, re-organizing, and processing information is a production process that uses inputs of data (or raw information) to yield an information output. Under this view, information is an economic resource that can be an input, output, or both.

From an engineering standpoint, the term used to describe communication content (whether it contains data or information) is not necessarily relevant. Instead, quantitative concepts such as the timely, speedy, and error free transmission of digitally coded communications treat the transport of information as a commodity.

One source of such a distinction is whether an organization creates and sells information or simply carries data. To the information producer, there is an important difference between data and information, a distinction that is absent from the communications carrier's worldview. The terms data and information refer to different concepts. Once data is organized, the result is an output of information, a value-added product.

According to LaFrance[2] the economist's role is itself based on the distinction between data and information. In an information age, it is increasingly important that we do not confuse data with information. Information is data which is placed within a particular context. It is the context and underlying conceptual framework that makes the data useful in decision-making. Without the context and framework, the value of data is indeterminate.

Agricultural economists are often the vital link in producing useful information out of data and defining what data are needed to produce information. Constructing a framework and establishing a context for data are what we do as economists. Another conceptualization of

---

[1] Source: http://encarta.msn.com/dictionary
[2] Vincent LaFrance is a member of the faculty of the Department of Management and Business Administration at Messiah College in Grantham, Pennsylvania.

information is as a "perception of pattern". According to Braman,[3] "Information from this perspective has a past and a future, is affected by motive and other environmental and causative factors, and itself has effects." Instead of counting bits or treating information as a homogenous commodity or input, the richness and meaning of content is considered. Information diffusion is a process where information is exchanged within an organization or among consumers in a target market. Information literacy has to do with the human ability to find and evaluate useful information. Information processing refers to the psychological tasks people use to remember information and act on it. The combination of information processing, the process of information diffusion and information literacy allow an organization to perceive information patterns so they can be used in problem solving.

A fundamental error in discussions of the information economy is the belief that it, like all economies, is based on some form of scarcity, but there is no scarcity of information in the world - we live in a world of information overload. Instead, there is a scarcity of time to "spend" and attention to "pay" to the information available. It is the scarce commodity that defines an economy of attention governing the relationship between information produced and information consumed.

Enter the Internet. Its low cost, ease of use, high speed, and reliability make the Internet almost perfectly suited to those of us who spend so much of our time producing and consuming information. Therein lies the fundamental dilemma of an "attentional" economy. Because the Internet is such a good way to distribute and exchange information, we are increasingly using it for these purposes. Information thus proliferates at an increasing rate. Yet, our time remains constant. As a result, the limits of our time force us to pay less attention to more information. By distributing more information more widely, more quickly, and more economically, the Internet intensifies an already fierce competition for our limited resource.

Processed information is typically more valuable than mere data or unprocessed information. Information becomes valuable through organization, categorization, analysis, and other information processing activities. In this sense, Information Technology (IT) processes information for problem solving.

It is from the conceptualization of information as a heterogeneous, behavioral input to be processed into a recognizable pattern and diffused throughout an organization that the relevance of information literacy skills arises. The American Library Association[4] gives information literacy this definition:

> "To be information literate an individual must recognize when information is needed and have the ability to locate, evaluate, and use effectively the information needed. Ultimately information literate people are those who have learned how to learn. They know how to learn because they know how information is organized, how to find information and how to use information in such a way that others can learn from them. An important implication of information literacy is that differences in the ability to deal with information are important explanations of variation in human capital."

---

[3] Change of State: An Introduction to Information Policy. Braman, 1989, p.238
[4] Source: http://www. ala.org/ala/acrllacrlissues/acrlinfolitlinformationliteracy.htm

Information's value must be balanced against its cost.   There is a balance to information value between information producers and consumers.

Putting barriers between potential users and the creator of the information is to limit the degree to which economic value will be derived from it.   The other side of this dilemma is that if no mechanism is in place for the inventors, or producers, or publishers, or disseminators of the new information to recapture their costs, then people will not produce as much information as is socially desirable.

Information is different from other economic goods; unlike them it cannot be displayed to a potential buyer, otherwise he will possess without having to buy.   Therefore, those who buy information are always uncertain of precisely what it is they are buying.   So imperfect is the market for information that price alone may determine demand, with information unwanted when it is sold cheaply and the same information in much demand when it is expensive. Consultants are more likely to sell a report whereas academics may give away the same report.

Exclusivity generally increases the value of information to the buyer, but as information may be reproduced at little cost and always remain with the seller anyway, exclusivity can be hard to guarantee.   Information tends to be of little use in isolation; the buyer seeks to purchase only that information which is compatible with what he already has.   Information itself may be a non-perishable good, but its value to many customers tends to be extremely time sensitive.

TecSec Incorporated

www.tecsec.com

Copyright 2014

Page 5

Proprietary

June 16, 2010

SR_15_001_Persistent Tagging

All Rights Reserved