ENABLING USER PARTICIPATION IN WEB-BASED INFORMATION SECURITY EDUCATION.

Ryan Goss[1], Johan van Niekerk[2]

[1]Nelson Mandela Metropolitan University
South Africa
[2]Nelson Mandela Metropolitan University
South Africa

[1]ryan@goss.co.za, [2]johan.vanniekerk@nmmu.ac.za

ABSTRACT

The greatest threat to Information Security are the employees within an organization. Many security controls rely on the user in order to be effective. It is thus vital to educate users about their role(s) in security. Many companies cannot afford, in terms of time or finances, to replace employees during training periods. The Web has long since been identified as a viable alternative to traditional training. To a certain extent, using the Web as a training platform depends on user buy-in. Web 2.0, which involves users and is largely user driven, is one way in which such buy-in could be obtained. This paper will discuss both the data acquisition and storage of Information Security principles to a centralized knowledge store, from which Web based Security Education technologies can draw inference. These web based security education applications should involve the user, thereby securing their buy-in and adding to the overall effectiveness of the training program. The use of Resource Definition Framework (RDF), SPARQL Protocol And RDF Query Language (SPARQL) and the Semantic Web will be discussed as possible solutions to the storage and transport of represented knowledge between multiple systems.

KEY WORDS

Information Security, Information Security Education, Web 2.0, Semantic Web

ENABLING USER PARTICIPATION IN WEB-BASED INFORMATION SECURITY EDUCATION.

1 INTRODUCTION

Information Systems have become a crucial tool to the success of an organization and thus need to be protected. At the same time these systems should provide adequate access to the information for the users within the organization. The protection of information resources is also known as information security, and is often described as the CIA, or Confidentiality, Integrity and Availability triangle. These three objectives fall in line with the fundamental goals of Information Communication Technology (ICT) security (Kruger, Drevin, & Steyn, 2006). Federal agencies or even Information Technology (IT) administrators cannot protect the integrity, confidentiality or availability of information in today's highly networked systems without first ensuring that each and every user of the system is aware and acting on their responsibilities within the information system (NIST 800-16, 1998). Mitnick and Simon (2002) argues that the greatest threat to Information Security are the users within an organization. In fact, this is so true that the Computer Security act of 1987 (Public law [P.L.] 100-235) required that each user within a federal agency be subjected to periodic training in both security awareness and computer best practises (NIST 800-16, 1998). The document further stipulates that these requirements include all users, from upper management to standard employees and even anyone involved within the operation of a federal computer system within the agency.

Potentially the most difficult part of security education process is ensuring user buy-in. The presentation aspect of an educational system should therefore be on-going, creative, motivational, eye-catching and intuitive, with the objective of focusing the learners attention so that the learning will be incorporated into conscious decision making (NIST 800-16, 1998). Another potential problem is that in the modern, competitive world, organizations cannot afford to take their users out of the operations of the organization for sustained periods of time without the company suffering. For sometime now, Hypermedia (Web-based) education systems have stepped up to the

plate as a solution to this problem. These systems have become known as E-Learning systems. Adaptive e-learning branched off from static, e-learning systems and provided an alternative to the "one-size fits all" e-learning scenario by allowing the system to evolve its interface and the content displayed by learning from interrogating the user and building a user model. Web-base Training (WBT), the delivery of instruction or learning content over the Internet and/or an organization's intranet is fast becoming popular amongst organizations (Lee, Chamers, & Ely, 2005). Lee et al. (2005) argues that motivation is a key element in the user acceptance of a training platform. Without user acceptance, the educational platform and the content it attempts to deliver will become nothing more than another neglected, eye-catching, yet useless lesson (Lee et al., 2005).

This paper argues that the use of recent developments in technology, such as Web 2.0 and the semantic web, make it possible to keep the user involved and motivated throughout a training program. This increased attention could enable users to better learn both simple and complex information security principles over mediums such as the World Wide Web or the organizational intranet. This paper also discusses methodologies which can be implemented to share stored knowledge, about the user profile and the information security principles being taught, amongst such systems.

2 RESEARCH PARADIGM AND RATIONALE

The purpose of this phenomenological study is to highlight the importance of the role that the user plays in information security within an organization, and to present methods the organization could employ to strengthen this "human factor". The paper is presented using argumentation theory as discussed in Van Eemeren (2001). This theory is concerned with the arts and sciences of civil debate, communication and persuasion. The paper does not necessarily cover new concepts, but rather serves to highlight various pre-existing technologies and how they are employed together to work towards the development of a successful, user-driven, Information Security Education platform.

As far as could be determined by the author, the use of Web 2.0 based

technologies coupled with the successful data sharing process for inter-educational knowledge base access (Web 3.0) for information security education systems has yet to be published. It is the author's belief that the sharing of such information by utilizing a de-centralized, generic knowledge base hosting data pertaining to information security principles and the user's profile, from which various educational tools draw inference would be a large asset in the struggle towards educating users. Creating interconnected, multi-platform compatible knowledge base access mediums will greatly aid in the strengthening of the "human factor" within information security.

The aim of this paper is therefore to show that Web 2.0, its knowledge representation storage mechanisms and transfer of this knowledge base between multiple systems is firstly possible and secondly will aid in the strengthening of the human factor within an information security environment.

3 WEB-BASED EDUCATION SYSTEMS

Hypermedia or Web-based educational systems, as mentioned previously, is by no means a technology in its infancy. In fact, ever since the establishment of the World Wide Web, scientists and scholars have been using the medium to promote information in static form to users of various web sites. Hypermedia offers a multimedia information environment, supports non-linear access to information, and provide a means of interaction with the user, all at the same time integrating the various information formats into a common display (Liaw, 2001). The rapid growth and development of the World Wide Web has been the main driving factor in the rapid migration of educational systems to hypermedia based applications (Liaw, 2001). Liaw (2001) continues to state that some of the potential benefits of hypermedia based applications would include: allowing the learner to structure their learning approach, the ability to pursue cross-references and to "remember" various aspects of the learning session.

Based on human cognition, computer assisted learning environments such as Hypermedia, are based on constructivist learning theory. Variations of this theory include social constructivism, which focuses more on the social context of learning as well as "cognitive constructivism" which which states that learners construct their own knowledge of the world through assimilation and accommodation (Liaw, 2001). Constructivism learning theory's educational

ideology is based on the learner constructing their own knowledge. This knowledge may be constructed through discovery, exploration and investigation (Cook, 2006). The teacher within a constructivist learning environment should structure the learning process so that they become a "co-constructor" of the knowledge being constructed by the learner, thus forming a partnership between both the student and the teacher (Cook, 2006).

In order for web-based systems to accommodate such learning processes, they are required to adapt to the learner, their specific needs for constructing knowledge, as well as the method of presentation of such knowledge for the learner to review. Adaptive e-learning has been around for some time now and addresses this very need. Adaptive e-learning systems can be broken into two main parts: adaptive content generation and adaptive interface design or presentation. Adaptive content generation is concerned with what content to show the learner; the learner should not necessarily be shown content that they are already familiar with. Adaptive presentation involves the user interface adapting to the preferences of the particular user, so as to avoid the heterogeneous "one size fits all" approach to education. The National Institute of Standards and Technology (NIST) IT Security Training requirements document requires that security awareness and training presentations should be designed with recognition that users practice *acclimation* or a tendency to tune-out if the stimulus or "attention-getter" is used repeatedly. Presentations should therefore be ongoing, creative and motivational with focus on the user to consciously start incorporating new knowledge into their existing behavioural pattern by way of assimilation (NIST 800-16, 1998). Adaptive hypermedia systems are perfectly aligned to allow for this constant changing presentation to occur and to assist the educational system in firstly providing the correct knowledge whilst at the same time keeping the user's attention and adapting to their individual learning style.

It is essential that adaptive e-learning systems collect and model user information so as to allow for the system to adapt to the user's characteristics and preferences (Froschl, 2005). An adaptive e-learning system should also have a strong knowledge base, from which the system draws inference. The user model is compared against this knowledge base or "domain model" and it is from this comparison that similarities are drawn and progress of the

learner is quantified. In order to build an Information Security educational adaptive e-learning suite, an extensive and accurate knowledge base is required containing various principles from within the subject domain.

New movements such as Web 2.0 have recently come to light in the struggle to keep the user involved in the training program, thereby ensuring their buy-in and allowing them to effectively participate in the information security training exercise. This participation includes both the the learning from existing information, as well as contribution of their own ontologies pertaining to particular information security principles.

4 WEB 2.0 BASED SYSTEMS

Web 2.0 is a term coined in the first O'Reilly Media Web 2.0 Conference in 2004. It is loosely defined as a business revolution within the computer industry caused by the movement to the Internet as a platform and designed to harness collective intelligence (Needleman, 2007). Web 2.0 is not a technology, but rather a way of thinking whereby users generate content which is published, used and managed through network applications in service-orientated architecture (Judicibus, 2008). Web 2.0 enabled websites also boast a host of advantages over standard "Web 1.0" websites. These include:

- *The user as a contributor*: The user is encouraged to participate in book reviews, commenting on articles, uploading multimedia such as photographs etc. Acting on what was previously discussed, this aids in the necessity to involve the user, thereby ensuring their attention whilst using the system.
- *Trust and collaboration*: Services such as wikipedia which are based on the concept that any user can add an entry and any other user can edit it (Needleman, 2007).
- *Multi-platform applications, above the level of any single device*: The World wide web provided a platform for content delivery over multiple devices. Web 2.0 takes this one step further with mobile devices

contacting remote servers, using the PC as a docking station and local cache during the transaction (Needleman, 2007).

- *Cost Reductions*: Not only are Web 2.0 applications relatively inexpensive to deploy, but in most cases Web 2.0 extensions can be added to non-Web 2.0 products to further reduce costs. For example, wikis could be deployed for users to build up knowledge bases and documentation with relatively little investment from the organization (Zambonini, 2006).

A web based system which actively involves the user as both a contributor and a casual browser could solve many of the obstacles faced by existing educational systems. Information Security Education does not always hold the interest of the users who are to take the courses and therefore whatever can be done to aid in the stimulation of the user and therefore the learning experience would be a huge asset to the training program. Many web-based information security education or awareness systems exist, however these systems operate within the confines of closed environments. One of the major downfalls of Web 2.0 technologies is in their inability to store information in a computer-readable format and therefore data-acquisition and sharing amongst various Web 2.0 websites is hindered. Whilst the majority of Web 2.0 applications typically provide some form of proprietary Application Program Interfaces (API) access to their underlying knowledge store, in order for a remote application to access this knowledge, the accessing application should have extensive parsing ability for the remote API set, with programs often traversing large eXtensible Markup Language (XML) trees to recover the required data (Heath & Motta, 2007). A storage and transport medium needs to be identified which will solve the problem of data storage, facilitating the interoperability of many of these potential Web 2.0 learning environments, thereby allowing the user access to a wealth of information and training material, all from a single website. One such technology, proposed by the World Wide Web Consortium (W3C) is already gaining wide acceptance - namely the Semantic Web.

5 SEMANTIC WEB AS A KNOWLEDGE TRANSPORT AND STORAGE SYSTEM

Breners-Lee (1998) described the Web as being an information space, whos goal is to be useful not only for human to human communication, but also

that machines would be able to participate and help. Breners-Lee (1998) further discusses that one of the major obstacles has been that information on the web has in the past been designed for human consumption and even if the data was represented in a technically sound manner, the structure of such representation would not be evident to a robot browsing the web. One of the core goals of the semantic web is to bring progressively more meaning to the information published on the web (Java et al., 2007). The semantic web encapsulates information with a collection of metadata which describes this information. Using standardized query languages such as RDF and Web Ontology Language (OWL) allows machines and human readers alike access to the information. The machine readers have access to the underlying metadata, whilst for the human readers, this information is masked so as to hide the underlying architecture and merely provide the information requested. Machines being exposed to the metadata will benefit from the deep semantic annotations in their application-orientated task processing (Java et al., 2007).

The semantic web therefore provides a near perfect platform for the development of shareable knowledge models on particular problem domains for the construction of knowledge base systems on an open environment such as the Internet (Chan, 2007). Such knowledge base systems enable semantic web agents to draw inference in common formats, thereby allowing for the ease of distribution and querying of remote knowledge stores without having to locally store the data. The various engines require a common protocol for data acquisition and transfer. Some examples of such protocols are RDF and SPARQL. The exact operation of these protocols is beyond the scope of this paper.

In order for the semantic web to facilitate the process of knowledge storage and retrieval for Information Security Educational applications, a suitable front end environment needs to be created in order for the user to be able to contribute to the knowledge store. One of the greatest downfalls of the Semantic Web is the lack of intuitive interface design for creating, modifying and querying data within the grid.

This system should be able to translate the information from the user to a semantic web based format (such as RDF), thereby enabling remote user applications to share in the accumulated ontology. One such front end has already been discussed : Web 2.0. The problem therefore becomes whether Web 2.0 and Semantic Web technologies can co-exist in order to promote user involvement and support within an Information Security Education System, thereby attending to both the presentation and data retrieval aspects of a successful adaptive e-learning system.

6 WEB 2.0 AND THE SEMANTIC WEB

Web 2.0 has aided in the contribution of an unprecedented volume of knowledge to the World Wide Web, through simple yet engaging interfaces, allowing the user to contribute to a vast number of subject domains (Heath & Motta, 2007). Heath and Motta (2007) continues to describe these heterogeneous knowledge stores as using techniques that do not facilitate the scaling beyond a handful of sources. The semantic web on the other hand provides the key to large-scale data integration, yet lacks the interactive user interfaces necessary to allow for contributions by non-specialists (Heath & Motta, 2007). The perfect hypermedia educational system should therefore provide an interface using Web 2.0 technologies, yet store the knowledge acquired in RDF data sets, ready to be shared via an underlying semantic web. This provision for contribution of knowledge by users would aid in the development of Information Security Education systems whereby experts contribute knowledge which would span world wide for various other educational systems to access. The following two sub-issues deserves special attention:

1. *Contributor Credibility*
   All users contributing to this wealth of knowledge should be rated to ensure the validity of such data. As the proposed educational system will be based on the semantic web, an RDF or Friend Of A Friend (FOAF) object would be built for each user and other users could rate this user, increasing their credibility or score on the system. A user with a high score could be said to be credible, conversely one with a low score would be deemed an amateur whos contributions should be questioned or confirmed by a higher ranking contributor. Harnessing the power of the semantic web, a particular user may already maintain

an existing RDF describing themselves on a remote site also supporting semantic web standard query languages. In this case, a user may link their profile to the remote FOAF object Unique Resource Identifier (URI), thereby allowing the Information Security Education system access to additional information about the user, such as qualifications, employment details, location or whatever the user has decided to publish in their RDF object. This remote access ability allows the base system to capture only minimal information about the user onto its local data store, encouraging the user to rather link to an alternate URI for the enhancement of their profile.

2. *Tagging, not classifying*
   Heath and Motta (2007) describes a method of tagging Web 2.0 data, now encapsulated in RDF format, instead of requiring the user to link the new knowledge under a particular heading or category. This ensures ease of contribution by the user, since the information supplied no longer needs to be fixed within the confines of a particular category. Knowledge which may not easily be classified is now easily tagged and stored in the underlying RDF database (Heath & Motta, 2007). Data about tags associated with particular Information Security principles would be described using the Tag Ontology and published on the website in Hyper Text Markup Language (HTML) (for human readers) and via the website's SPARQL endpoint, enabling machines access to the knowledge store. Each Information Security Principle is able to be tagged numerous times, thereby allowing web searches more accuracy whilst querying the knowledge store. Having tags also allows for related principles to be displayed to the user whilst they browse the site or provides a semantic pathway to machines traversing the data.

   As the website learning environment would be Web 2.0 powered, each Information Security Principle would have a section where users could post their views on the principle, argue for or against its validity and generate a discussion around the subject area. These discussions would too be published in RDF format to the underlying semantic web.

   As each user on the website has a FOAF object, stored on the central server, any Information Security application drawing inference from this knowledge store is able to keep track of the progress of the user, using user modeling techniques. These user models can then be con-

verted to RDF and linked against the FOAF object for a particular user. The ability to track the progress of the education of the user ensures feedback to organizations pushing for user training in the field of Information Security, thereby strengthening the human factor. Organizations could flag certain tags within the system and ensure that their users complete all training related to these tags. The system will keep track and quantify the results of the training and give detailed reports back to the organization as to the understanding of the employee.

Future information security education systems, which incorporates concepts from both Web 2.0 and the Semantic Web should thus, for best results, exhibit the following characteristics:

- An intuitive, user-involved and morphing interface
- A common knowledge storage format
- A common interface for querying stored knowledge
- Knowledge and data contributions by users of the site
- A simplistic classification of submitted user information

From the above it should be clear that the Web 2.0 philosophy is well suited to use for interface and interaction design in information security educational systems. Similarly, the Semantic Web would be an appropriate methodology for common data (knowledge) storage and querying, using the Tag Ontology for classification of the data. The way forward is therefore quite clear - toward a third generation, Web 3.0, educational system where various *Mashups* are able to interconnect and share ontologies and knowledge stores, enabling better access to the knowledge and a more customized system for all users. Web 3.0 based educational systems should focus on the backend transports and storage layers, rather than primarily the front end as has been the case on the Web as of late. Nova Spivak of Radar Networks describes Web 3.0 as the next big step in Internet development which is still in its infancy and should be mainstream as of 2010. Spivak is ambitious in his discussion as to what follows the Web 3.0 era - Web 4.0. Spivak finishes by stating that in the world of Web 4.0, users will benefit from distributed searches, intelligent personal agents, semantic databases etc truely working towards 'The WebOS'. Future research regarding Information Security

Education using the latest Web technologies could include an investigation into the movement toward distributed searches, a discussion of the technical storage mechanisms for converting Web 2.0 input into RDF format for use within the semantic web and the underlying technical implementation for communication on the network.

7 CONCLUSION

This paper introduced the idea of implementing a hypermedia Information Security Education platform, powered by Web 2.0 and Semantic web technologies to get the best of user interaction and knowledge base sharing across multiple systems - giving rise to a Web 3.0 training platform.

It was argued that by using Web 2.0, non-specialists could generate semantic annotations suitable for use within a semantic web. The use of Web 2.0 ensures user buy-in to the training experience, thereby keeping their attention and educating them in the various Information Security Principles.

It was further argued that these principles could be captured by any user, however the credibility of such input would be based on a scoring facility, indicated by the credibility and acceptability of each contributor. The comments facility was also discussed to initiate inter-user communication and arguing, ensuring a better understanding of each principle, rather than a blind acceptance of it. This aids in the embedding of such knowledge in the day to day actions of the user, greatly adding to the effectiveness of security awareness campaigns and overall organizational security practises.

Whilst this paper does not, at a technical level, provide a solution on how to use Web 2.0 to enable user-participation in information education, it does show that such an approach is definitely possible. Future efforts in line with this research will be aimed at delivering the more technical hands-on parts of this solution. It should also be clear that the idea of ensuring user buy-in into security education programmes by implementing a Web 2.0 interface still needs to be tested empirically.

References

Breners-Lee, T. (1998). Semantic web roadmap. *[WWW document]. URL http://www.w3.org/DesignIssues/Semantic.html, Sited 2 June 2008.*.

Chan, C. (2007). Development of an ontology for an industrial domain. *Intl Journal of Cognitive Informatics and Natural Intelligence*, *1*(3).

Cook, P. (2006). The project approach: An appreciation for the constructivist theory. *Published by the Forum on Public Policy*.

Froschl, C. (2005). *User modeling and user profiling in adaptive e-learning systems*. Unpublished master's thesis, Graz University, Austria.

Heath, T., & Motta, E. (2007). Ease of interaction plus ease of integration: Combining web 2.0 and the semantic web in a reviewing site. *Web Semantics: Science, Services and Agents on the World Wide Web*.

Java, A., Nirneburg, S., McShane, M., Finin, T., English, J., & Joshi, A. (2007). Using a natural language understanding system to generate semantic web content. *Intl Journal on Semantic Web and Information Systems*, *3*(4).

Judicibus, D. de. (2008). World 2.0. *[WWW document]. URL http://lindipendente.splinder.com/post/15354690/World+2.0., Sited 23 April 2008.*.

Kruger, H., Drevin, L., & Steyn, T. (2006). A framework for evaluating ict security. *Information Security South Africa Conference*.

Lee, D., Chamers, T., & Ely, T. (2005). Web-based training in corporations: design issues. *Intl Journal of Instructional Media*, *32*(1).

Liaw, S. (2001). Designing the hypermedia-based learning environment. *Intl Journal of Instructional Media*, *28*(1).

Mitnick, K., & Simon, W. (2002). *The art of deception: Controlling the human element of security*. Wiley Publishing.

Needleman, M. (2007). Web 2.0 and lib 2.0 - what is it? (if its anything at all). *Serials Review*.

NIST 800-16: Information Technology Security Training Requirements: A Role- and Performance-Based Model. NIST Special Publication 800-16, National Institute of Standards and Technology. (1998).

Van Eemeren, F. (2001). Crucial concepts in argumentation theory. *Amsterdam University Press*.

Zambonini, D. (2006). Why you should let web 2.0 into your hearts. .

8 ACKNOWLEDGEMENTS

The financial assistance of National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the National Research Foundation.