View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Stellenbosch University SUNScholar Repository



# Web 3.0: Governance, Risks And Safeguards

Mr. Rikus (Hendrik Jacobus) Bruwer, Stellenbosch University, South Africa
Mr. Riaan Rudman, Stellenbosch University, South Africa

**ABSTRACT**

*Many organisations consider technology as a significant asset to generate income and control cost. The Web is recognised as the fastest growing publication medium of all time. This mass of unstructured information presents many new opportunities for organisations. The Web acts as an enabler for technological advancement, and has matured in its own unique way. From the static informative characteristics of Web 1.0, it progressed into the interactive experience Web 2.0 provides. The next phase of Web evolution, Web 3.0, is already in progress. Web 3.0 entails an integrated Web experience where the machine will be able to understand and catalogue data in a manner similar to humans. This will facilitate a world wide data warehouse where any format of data can be shared and understood by any device over any network. Organisations need to be ready, and acquire knowledge about the opportunities and risks arising from Web 3.0 technologies. The objective of this study is to investigate the risks an organisation will be exposed to when interacting with Web 3.0 technologies. The study proposes to provide insight into the risks arising from the use of Web 3.0, and to recommend possible safeguards to mitigate these risks to an acceptable level.*

*Identified opportunities can mainly be characterised as the autonomous integration of data and services which increases the pre-existing capabilities of Web services, as well as the creation of new functionalities. The identified risks mainly concern unauthorised access and manipulation of data; autonomous initiation of actions, and the development of scripts and languages. Risks will be mitigated by control procedures (examples include encryptions; access control; filtering; language and ontology development control procedures; education of consumers and usage policies). The findings will assist management to identify the key focus areas when implementing this new technology.*

**Keywords:** Web 3.0; Risks; Controls; Safeguards; Semantic Web

## 1. INTRODUCTION, RESEARCH OBJECTIVE AND METHODOLOGY

### 1.1 Introduction And Research Objective

Many organisations consider technology as a significant asset necessary to generate income and control cost (Brynjolfsson & Hitt, 2000). The World Wide Web (henceforth referred to as the Web), is recognised as the fastest growing publication medium of all time, containing well over 1 trillion URLs (Alpert & Hajaj, 2008). With an estimated growth rate of 566% in Internet usage in the last twelve years (Internet World Stats, 2012), the Internet has become the main source of communication worldwide. With ever increasing growth rates, the technology supporting the structure of the Internet is evolving. Keeping abreast with technological trends creates new opportunities for organisations, as well as challenges.

The Web, acting as an enabler for technological advancement, matured in its own unique way. Initially there were the static informative characteristics of Web 1.0 which progressed from a passive to an interactive experience that Web 2.0 provides. The next phase of Web evolution, Web 3.0, is already in progress. The evolution of the Web will bring forth new opportunities. Web 3.0 will change the way people interact with devices and

networks, and how companies use information to market and sell their products, and operate their businesses (Booz & Company, 2011). Web 3.0 calls for a complete reconstruction of Internet and IT infrastructure. Organisations need to start preparing for the changes, otherwise they may be unable to satisfy customer needs, capitalise on emerging trends, and seize new opportunities. Before rushing to realise these opportunities organisations need to fully understand the impact the technology will have on business operations. The objective of this study is to investigate the risks an organisation will be exposed to when interacting with Web 3.0 technologies (referred to as Web 3.0). The study proposes to provide managers, boards of directors, IT professionals and information managers insight into the risks arising from the use of Web 3.0, and to recommend possible safeguards to mitigate these risks to an acceptable level.

The purpose is not to discuss the underlying technologies of Web 3.0 in detail, but rather to highlight the risks arising from the use of these technologies. The study investigates the impact Web 3.0 will have on broad based business drivers which are applicable to most industries and companies, and not on industry-specific business drivers. The focus of the research is on incremental risks specifically pertaining to Web 3.0, and not pre-existing internet risks. Some of the prevalent risks associated with previous Web generations will be reinvestigated since the underlying technology creating these risks, has changed.

### 1.2 Methodology

A non-empirical study reviewing papers published in accredited research journals, popular articles, whitepapers and Websites was conducted. In order to add scientific rigour to a literature review, a four stage approach is suggested by Sylvester, Tate and Johnstone (2011). This four stage approach was followed by the researcher, and each stage was repeated and performed interactively. A wide selection of articles and readings were selected at the beginning stages to enable a comprehensive understanding of the underlying literature, and the selection was narrowed down to more specific areas at the latter stages. The literature was selected within a timeline between 1996 and 2013. Literature was reviewed in order to develop an understanding of *inter alia*, *'Web 3.0'; 'semantic Web'; 'next generation Web service'; 'Technologies driving Web 3.0'; and 'defining Web 3.0 technologies'.* Thereafter *'control framework for IT governance',* was considered. Due to the fact that very little research has been conducted on the subject matter, the reputational value of the articles was originally not taken into account during the selection. An in depth reading of the narrowed down selection enabled the researcher to develop a concept of Web 3.0 and underlying technologies, and to elaborate on the impact these technologies have on business operations. Thereafter the following structured approach was followed:

1. **Define Web 3.0.** A widely accepted definition for Web 3.0 did not exist as minimal research has been performed on the subject matter. Therefore, a formal definition is required. This definition was used to categorise the technologies that would fall within the ambit of Web 3.0.
2. **Obtain an understanding of the technologies underlying Web 3.0.**
3. **Select an appropriate control framework.** Based on a literature review of various control frameworks, an appropriate control framework to be used to identify the risk applicable to Web 3.0 was selected. Control Objectives for Information and related Technology (COBIT) was selected because of the low implementation cost, and the fact that it is openly available. Moreover, it is widely acceptance and international recognised by various international organisations. It covers a wide range of IT processes which ensures easy alignment with other international frameworks and standards, thereby ensuring sound controls and regulatory compliance.
4. **Perform a detailed study of COBIT control framework.**
5. **Use the COBIT framework to identify risks.** The technology was mapped against the COBIT framework and associated processes and control objectives. These objectives were used to identify relevant risks associated with the changes in business operations due to the impact of Web 3.0. A control framework was used to identify the risks because implementing IT governance practices will lead to:
    - The development of a more complete risk and control framework.
    - Strategic alignment between IT and business goals that will create a competitive advantage.
    - Better risk management procedures and a better understanding of IT.
    - Greater compliance with governance requirements, laws and regulations.

6. **Formulate mitigating safeguards**. The impact of each risk was evaluated and suitable internal control measures formulated. The mapping is available on request.

By implementing the methodology above, a more detailed understanding of Web 3.0 and the underlying technologies, risks and controls was obtained.

## 2. LITERATURE REVIEW

The latest evolution in Internet, Web 3.0 will not only restructure Internet communication, but give rise to new business drivers while redefining existing business drivers. The exact interpretation of what Web 3.0 will ultimately entail and how it will influence the Web experience, is not clear, but an array of opportunities arise for innovative services and applications with the introduction of these technologies (Knublauch, Fergerson, Noy & Musen, 2004). Missing opportunities are not the only concern for organisations when analysing the risks promoted by these new technologies.

### 2.1 Historic Review

Historical research on the evolution of the Web shows patterns and goes through phases. Initial research focused on defining the technology, understanding its benefits, and how it will have an impact on business environments regarding opportunities and challenges (Clearswift, 2007; O'Reilly, 2009). Research investigating user behaviour and privacy issues (Lawler & Molluzzo, 2010) focusing on knowledge of personal information gathering, and sharing techniques on Web technologies, has also been undertaken. As the Web evolved and became more popular, the focus shifted to security risks, including business risks (Grossman, 2007). The latest research into Web evolution conducted by Benjamins, Contreras, Oscar, Corcho and Gómez-Pérez (2002), focuses on defining and predicting the challenges arising from Web 3.0. Related research by Lu, Dung and Fotouhi (2002) investigates possible opportunities and complications Web 3.0 might offer, and how an enterprise can gain business value from using these applications.

Various attempts have been made to develop an organisational framework to help businesses to mitigate the risk arising from the use of Web technology (not including Web 3.0). Dawson (2007, 2008) tried to develop a framework in an effort to help businesses not only to understand and mitigate risks, but also to add business value from using Web technologies. Rudman (2010) specifically considers the incremental risk arising from Web 2.0 technologies, and the creation of a control framework to mitigate the security risk in Web 2.0.

The majority of research completed on Web 3.0 was performed by independent private organisations like Booze & Company; Verizon; Gartner, Clearswift and SEM Logic. Most of the research consists of whitepapers and articles with very little academic peer-reviewed articles. Most of the articles aim to define Web 3.0, and rarely address advantages and disadvantages arising from use of Web 3.0. A study that focuses on defining Web 3.0, identifying the business opportunities and risks arising from the use of this technology, and the creation of a control framework to mitigate these risks, has not been conducted.

### 2.2 Stages In The Evolution Of The Web

To understand in what direction the Web is heading and what impact it will have on organisations, it is necessary to define the various stages of the Web evolution. It was difficult to differentiate between the various stages of the Web (O'Reilly, 2007).

#### *2.2.1 Web 1.0*

Web 1.0 was a platform through which information could be published in a static form designed with text and images (DCruz, 2009). It portrayed an environment where information and data were static, and displayed with no interaction between the information and the consumer. Content could be viewed, but not created by users.
The protocols associated with this generation were Hypertext Transfer Protocol (HTTP) and Hypertext Markup Language (HTML). The HTTP protocol transfers information between a Web server and a Web browser. HTML

protocol communicates with the browser, and informs it how to display whatever text, graphics and images transferred by the HTTP protocol.

*2.2.2 Web 2.0*

In an effort to clarify the paradigm shift, Cormode and Krishnamurthy (2008) stated that the main difference between Web 1.0 and Web 2.0 is not only the underlying infrastructure of the Web, but rather the ability of consumers to create, share and interact with content on the Web. New technological aids made it possible for consumers to create and share content. Getting (2007) describes it as the greater collaboration between consumers, programmers, service providers and organisations, which enables them to re-use content. Web 2.0's applications have the ability to harness collective intelligence, and in doing so combine and integrate Web content and services to improve the end user`s experience (Giannakos & Lapatas, 2010). This sharing was facilitated by online software with the ability to deliver rich interfaces operable on any device or platform without the need of additional software installation.

*2.2.3 Web 3.0*

Web 3.0 is not represented by the emergence of a new Web but rather an extension of the technologies already present in Web 2.0. Internet content is becoming more diverse, and the volume of data becoming more openly available (Bergman, 2001). The Web is becoming a platform for linking data, and by making connection between similar data characteristics, the data itself becomes more valuable (Tarrant, Hitchcock & Carr, 2011). Computers still cannot automate the function of harvesting this data, or of performing complex tasks with it (Intervise, n.d.). The need for data structuring and integration is important to enable the Web to evolve into its next phase.

As with previous versions of the Web, consensus on the definition for Web 3.0 varies (Farah, 2012),and names include, amongst others: *Web 3.0*; *Semantic Web*; *Transcendent Web* and *Web of Things* (henceforth collectively referred to as *Web 3.0*). Even though the names differ, all these phrases have the same basic fundamentals.

Wolfram (2010) stated that Web 3.0 is where the computer, rather than humans, generate new information. This is supported by Morris (2011) that integration of data is the basic foundation of Web 3.0, and by using metadata (a term used to describe data within data, which provides information about a item's content) imbedded in Websites, data can be converted into useful information, and be located, evaluated, stored or delivered by software programs designed to collect information based on the users' interaction with the Web (known as Intelligent Agents (IA's)). They can also act on behalf of the user to perform certain tasks. In order for IA to understand the information gathered, expressive languages that describe information in forms understandable by machines, need to be developed (Lu *et al.*, 2002). With the development of expressive languages Web 3.0 has the capability to use unstructured information on the Web more intelligently by formulating meaning from the context in which the information is published (Verizone, n.d.). There is a need for Web 3.0 to express information in a precise, machine interpretive form, so that IA can process this data and not just share it, but understand what the terms describing the data mean (Noy, Sintek, Decker, Crubézy, Fergerson & Musen, 2001).

Booze & Company (2011) stated that recommendation engines will focus on habits and preferences of users, and in doing so will produce more complete and targeted information. The information of habits and preferences used on a recommendation engine will be collected and stored in a hierarchical manner by IA. This is what will give Web 3.0 the ability to gather, analyse and distribute data which can be turned into information, knowledge, and, ultimately wisdom (Evans, 2011).

The key elements of Web 3.0 are:

- The **introduction of new programming languages** with the ability to categorise and manipulate data in order to enable machines to understand data, and the phrases describing data.
- The capability of **obtaining contextual information** from a Web search and storing it in a hierarchical

manner, according to similar characteristics for easy and specific retrieval.

- The ability to obtain information from a **bigger and wider variety of sources**, including previously walled application.
- The **ability to create and share all types of data** over all types of networks by all types of devices and machines.

Web 3.0 will ultimately entail an integrated Web experience where the machine will be able to understand and catalogue data in a manner similar to a human. The data collected will be categorised in a hierarchical manner in order to link data with similar characteristics, and retrieve consumer specific data effectively and efficiently. This will facilitate a worldwide data warehouse where any format of data can be shared and understood by any device over any network.

With the adoption of new technologies, organisations lack the understanding on how to define it, extract value or govern it.

### 2.3 Corporate Governance And Control Frameworks

With the ever increasing reliance on IT and emphasis placed on information by legislation, such as the Protection of Personal Information Bill; Protection of Information Bill, Regulation of Interception of Communications and Provision of Communication-Related Information Act and codes such as The King Report on Corporate Governance III, the responsibilities of managing and controlling IT risks and the information organisations possess, have become a vital part of corporate governance. The process of managing IT has become more than mere technical functions carried out by IT experts, and now forms part of the essential management function within an organisation (Stoneburner, Goguen & Feringa, 2002). The King III report contains the corporate governance principles for South African organisations. The governance principles reported in King III led to the introduction of IT governance in South Africa.

In order to ensure effective IT governance, organisations should make use of already established and internationally accepted frameworks as guidance. The Control Objectives for Information and related Technology (COBIT) provides a detailed framework, and describes the controls which need to be implemented in order to have a sound IT governance structure and controls (Hardy, 2006).

The purpose of COBIT framework is to guide management with controlling and managing information and related technology. The framework describes the importance of IT resources (people, applications, technology, facilities and data), and how the information, with the assistance of IT processes, created by the resources, should be delivered in order for it to fully support the business objectives (Hussain & Siddiqui, 2005). This delivery is controlled through 37 high-level control objectives or processes contained in five domains. These are as follows:

- **Monitor, Evaluate and Assess:** entails the evaluation of whether business and IT processing goals are achieved. It ensures that all IT processes are monitored in a timely manner, and the results are measured against the expected outcomes.
- **Align, Plan and Organise:** addresses strategic aspects, and concerns the implementation of a proper IT infrastructure.
- **Build, Acquire and Implement:** covers identifying, developing and acquiring an IT solution. It also assists in creating IT maintenance policies to control changes and maintenance of the system.
- **Delivery, Service and Support:** covers the actual delivery of required services, including service delivery; management of security and continuity, as well as training and technical support for users.
- **Evaluate, Direct and Monitor:** All IT processes need to be regularly assessed for effectiveness and ability to meet the business objectives. This includes assessing performance management and monitoring of internal control, regulatory compliance and governance (ISACA, 2012).

Each of these five domains and underlying processes will assist an organisation in implementing the controls needed to mitigate the risks identified and associated with the adoption of new Web 3.0.

## 3. FINDINGS

In order to assess the impact Web 3.0 will have on business operations, the first step is to define the different technologies associated with Web 3.0. After obtaining an understanding of Web 3.0, COBIT control processes will be used to identify risks associated with Web 3.0. After the risks have been identified, appropriate controls will be identified to mitigate these risks.

### 3.1 Defining Technologies Associated With Web 3.0

According to Berners-Lee, Hendler and Lissila (2001), Web 3.0 will rely on a variety of different technologies, some of which are still in development, while others are already, to varying degrees, being implemented on the Web. In order to obtain a better understanding of what Web 3.0 consists of and how it functions, one needs to be familiar with how specific terminology (such as these presented in Figure 1) associated with Web 3.0 interact. The following sections discuss some of these terminologies and technologies. These technologies can be classified in terms of Identifiers, Structures and Languages.

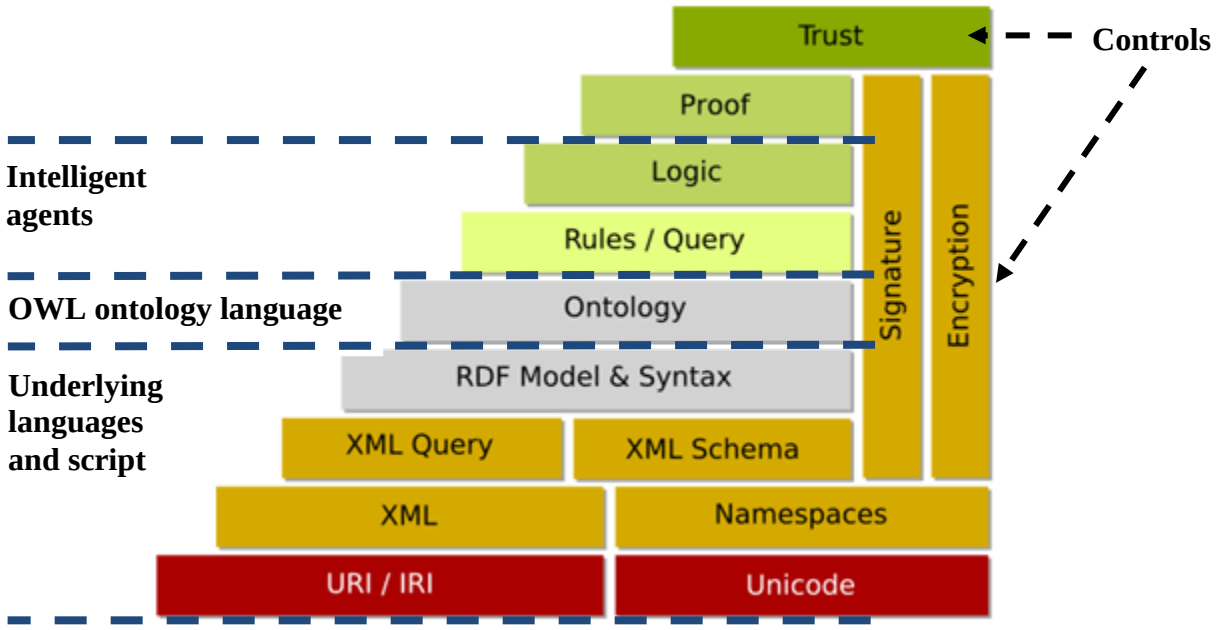


**Figure 1.** The Technological Layout Of Web 3.0

*3.1.1 Extensible Mark-up Language*

Unicode is an extensive method of defining characters electronically to ensure internationalisation of applications. It is based on similar principles as ASCII code, and is used by XML to describe characters. Extensible Mark-up Language (XML) is used to define documents with a standard format that can be read by any XML compatible application. The language can be used with HTML pages, but XML itself is not a mark-up language. Instead, it is a "meta-language" that can be used to create mark-up languages for specific applications. XML lets everyone create pieces of information, also known as tags. These tags are used to describe parts of a Website or sections of text. These tags of information can be used by other scripts or programs in very sophisticated ways if the script writer understands what the page writer used the tags for. XML provides a way for page writers to include extra tags of information on their pages, but does not explain what the tags are used for (Tim Berners-Lee *et al.*, 2001).

*3.1.2 Simple Object Access Protocol*

Hyper Text Transfer Protocol (HTTP) is a protocol used on the Web to enable different machines with different operating systems and software to communicate with each other over the Internet (Rouse, 2005). One of the main characteristics of Web 3.0 is internationalisation of Web languages and interoperability of all machines on the Web. Current Web applications run secure protocols, and HTTP has not been designed to bypass these firewalls, which lessen interoperability of new applications. Simple Object Access Protocol (SOAP) uses XML and HTTP to communicate between applications and bypass firewalls. SOAP is a XML-based protocol that enables communication between different applications.

*3.1.3 Resource Description Framework*

Resource Description Framework (RDF) is a mechanism through which information about data is captured. It acts as a mechanism for Web page writers to add semantic information to their pages. This type of data collected by RDF is called *metadata*. Metadata is a term used to describe data within data. It provides information about a item's content. RDF is the specification that defines how metadata, or descriptive information, should be formatted. The RDF model uses a standardised subject-predicate-object format.

RDF creates statements about particular resources on the Web by means of a triple expression in the form of subject-predicate-object. The subject represents the resource, while the predicate refers to an attribute of the subject, and the object is what is referred to in the predicate. This structure is the natural way to describe the vast majority of the data processed by machines (Tim Berners-Lee *et al.*, 2001; Decker, Melnik, van Harmelen, Fensel, Klein, Broekstra, Erdmann & Horrocks, 2000).

RDF uses Uniform Resource Identifiers (URI) to specify subjects and predicates. URI's are used to identify all resources on the Web. URI identifies the name and location of a file or resource in a uniform format. URI's provides a standard protocol for resources to be accessed by other computers across a network or over the Web. After the information has been identified, the method of locating the resources is called Universal Resource Locater (URL). URL is the address of a specific Website or file on the Internet and specifies where to obtain resources.

XML and RDF have some similar qualities on querying documents on the Internet. XML has the ability to query information in a document, while RDF has the ability to extract the "meaning" of information in a document, and query that which will be essential in the development of Web 3.0 (Berners-Lee, 1998).

*3.1.4 Resource Description Framework Schema*

Resource Description Framework Schema (RDFS) is a set of classes with certain properties using the RDF extensible knowledge representation language, providing basic elements for the description of ontologies, otherwise called RDF vocabularies, intended to structure RDF resources. RDFS is an extension of the RDF vocabulary. It has the ability to collect a range of properties and relate the RDF classes and properties into taxonomies using the RDFS vocabulary.

*3.1.5 Structured Query Language and Simple protocol and RDF query Language*

Structured Query Language (SQL) is a mechanism that enables communication with a database. It is the standard language for relational database management, and has the ability to perform tasks such as retrieving data from a database, or updating a database (Melton & Eisenberg, 2001). It is a widely used standard on the Internet. Due to the complexity of data storage within RDF and RDFS, a more complex and integrated query language had to be developed. Simple Protocol and RDF Query Language (SPARQL) was created in order to enable RDF database manipulation. It has the ability to make RDF data available through a standard interface, and query it, using a SQL (Quilitz & Leser, 2008).

*3.1.6 Ontology Web Language and Web Ontology Language for Services*

According to Sowa (2009) ontology is the study of the categories of things that exist or may exist, and describing their relationships in a certain domain. For the Web, ontology is about extracting descriptions of Web information, and understanding relationships between Web information. Ontology Web Language (OWL) is the language that enables a machine to process information contents on the Web in a universal manner. OWL was created to give machines, the ability to process and read information on the Web. OWL is a set of mark-up languages which are designed for use by applications that need to process the content of information, instead of only presenting information to humans. OWL ontologies describe the hierarchical organisation of ideas in a domain, in a way that can be understood by software. OWL has more facilities for expressing meaning and semantics than XML, RDF and RDFS, and thus OWL goes beyond these languages in its ability to represent machine interpretive content on the Web.

OWL has a lot of characteristics similar to those of RDF, but is much stronger with greater machine interpretability, a larger vocabulary and a stronger syntax (Webschool, 2013). Similar languages have been developed in the past, but only for specific user communities (examples include science and company-specific e-commerce applications). These earlier adoptions of ontology languages were user-specific, and not designed to be compatible with the architecture of the Web.

The ability of machines to harvest information and understand the meaning thereof, is a crucial characteristic needed in order to develop Web 3.0. OWL will create the opportunity for machines to adopt these characteristics and, through RDF linking, enables a Web where information is categorised by machines in a meaningful manner and in a universal format that can be queried by any other scripts (Horrocks, 2004).

OWL-S is an extended version of OWL, based on the same principles and annotation processes, but has a greater ability with respect to expressive properties, extends support for data types, enables metamodeling and extends annotation (Golbreich & Wallace, 2012).

*3.1.7 Intelligent agents*

One of the major features that need to be available in order for Web 3.0 to be more accurate and useful, is the establishment of interlinking data between data sets across the Web. Intelligent agents (IA) with the use of ontologies will enable interlinking. IA's are enclosed computer systems consisting of specialised computer architecture and programming (Lewis, 2008). These IA are programmed to function in a manner similar to humans browsing the Web. Agents will be omnipresent on Web 3.0, and will be able to harvest and collect information based on the users' interaction with the Web. They can also act on behalf of the user to perform certain tasks. The precision and applicability of the information harvested and utilised by IA's, will be greatly improved through the use of ontologies (Lu, Dong *et al*., 2002).

The ability which OWL technologies provide to IA's will enable them to create meaningful reasoning about information on the Web, which will equip IA's with knowledge about data, and increase their intelligence and mobility. The current computing paradigm on the Web is based on the client/server initiative. With IA's gathering more intelligence and mobility, the paradigm will shift to agent based distributed computing. IA's will be able to fulfil their assignments autonomously and precisely by migrating from one site to another, carrying their codes, data, running states and intelligence (Lu, Dong *et al*., 2002). IA's will act as an electronic assistant by automating repetitive tasks, harvesting and summarising complex data, and being able to learn on behalf of the user by analysing the users' interaction with the Web. The information gathered from this analysis, will give an agent the ability to make recommendations to the user (Gilbert, 1997).

IA's ability to harvest information collectively and autonomously and to structure it has been limited due to the fact that they are not able to harvest information from the Web in a meaningful way, and reason with this information. With the introduction of technologies outlined above, IA's will gain the ability to understand information in a meaningful way. The technologies discussed will form the core operatives to enable Web 3.0, and mainly consist of scripts and programming codes/languages. The required infrastructure, usage, data and

information are already available on the present Web 2.0 platform. These technologies will form the fundamental structures which will enable machines to exploit the rich content available on the Web and create opportunites.

**3.2 Opportunities or possible uses for Web 3.0**

In present business and other environments, the Web is an essential resource, and Web 3.0, with metadata annotated information, will be even more vital for completing information based tasks. By combining the technologies discussed in section 3.2, the Web has the potential to become the location of every possible information resource, person, and organisation, and all the activities relating thereto (Sheth & Meersman, 2002). Through Web 3.0 and IA, processes will become more automated, producing information much faster and precisely, at an improved level of access (Bakshi & Karger, 2005). The enhanced ability Web 3.0 will offer to machines to categorise and add meaning to information, will increase the range of uses for the Web, and will bring forth new opportunities.

- Overall increased collaboration between consumers, developers and machines. With Web 3.0 enabling machines to read, understand and reason with information on the Web, information will become integrated and precise, readily available, and will become more valuable to the consumers of the information.
- Autonomous characteristics of Web 3.0 will lighten the work load of data management, and enable new intuitive and personalised Web services. Web services, in this context, refer to websites which not only provide static information and allow the user to interact and contribute information, but also have the ability to create new Web services based on user preferences.
- Agent based distributed computing refers to the shift in the computing paradigm from information distributed through a client/server initiative, to information harvested and distributed autonomously.
- Web 3.0' ability to integrate and structure data autonomously, will increase the accuracy and availability of research data repositories. This will increase knowledge management, and enable more effective and collaborative research with less common restrains (such as language barriers). It will also assist in enhancing business intelligence or big data capabilities.
- They will overcome the short coming of traditional search engines that lack coherence in two major areas, namely, the reliability of the resources, and the relevancy of the information found. Natural language processing and Web 3.0 will enable a search engine to organise information based on the context within the document, and not just recognition of phrases. This, combined with information collected by IA's, will better define the users' preferences and make searches more precise and personalised than is possible with current algorithms (Verizon, n.d.).
- Web 3.0 and their ability to express meaning to data, will open an area of exploitation in eLearning and research (Naeve, Lytras, Nejdl, Balacheff & Harding, 2005). The facilities provided by Web 3.0 will enable learners to create, annotate, share and discuss content over the Web (Ghaleb, Daoud, Hasna, ALJa'am, El-Seoud & El-Sofany, 2006). According to Sampson, Lytras, Wagner and Diaz (2004), Web 3.0 will enable the creation of hypermedia systems. These hypermedia systems are portrayed as silos of information with the ability to adapt to the changes in its environment. The ability for a database to adapt is a crucial factor in the area of eLearning, especially taking into account the different needs of learners.
- IA's harvesting personal habits and information of consumers, will create a personalised Web experience which will amount to countless opportunities for inbound marketing schemes. Inbound marketing involves the distribution of information to consumers who value the information (Prescott, 2012). The main objective of inbound marketing is to target specific consumers based on their semantically related market segments, and build an electronic relationship with consumers by personalising their economical browsing experience.

*3.3 Risks Associated With Web 3.0*

Web 3.0 creates the opportunity for collaborative and autonomous integration and distribution of data on the Web. Autonomous machine communication, harvesting of data and creation of information, presents opportunities, as well as risks that need consideration when evaluating Web 3.0 (McGraw, 2008).

*3.3.1 Pre-Existing Web Risks*

Rudman (2010) explained that the risks associated with the different stages of Web evolution, are incremental. Vulnerabilities which were present during the first generation of the Web, had an impact on the second generation as well. The same notion will be applicable to the third generation of the Web. Some of the homogeneous vulnerabilities that organisations are, and will be, exposed to are:

- Unauthentic electronic intrusion.
- Unwanted application performance due to continuous updates.
- Over-reliance on services offered by third parties, or only relying on server side security.
- The loss of confidential and personal information due to malicious attacks.
- Unproductive use of organisational resources.
- Non-compliance with regulatory governance, and the possibility of loss due to legal action.
- Shortage in experienced technicians to ensure effective operation and monitoring of complicated systems and applications (Rudman, 2010).

The next section of the study explains the risks incrementally associated with Web 3.0.

*3.3.2 Unauthorised Access To Sensitive Information*

The exponential growth and availability of information on the Internet, and the new technologies offered by Web 3.0, will make data a crucial information resource. The ability of Web 3.0 to personalise Web usage, and IA to harvest browsing history and personal information in order to automate the Web experience, will bring forth a new level of privacy concerns. According to Nematzadeh and Pournajaf (2008) securing the Web is not just preventing unauthorised access, but also the prevention of unauthorised modification of data and use of resources. Kumar, Prajapati, Singh and De (2010) divide unauthorised access and data manipulation into four categories:

- **Unauthorised access**. The intrusion and capturing of sensitive information on a system by an entity without authentication. Many authorisation vulnerabilities exist, when no authentication is being used, or when password authentication is present, but it gets passed in plaintext format through SOAP headers. Another threat is when basic authentication is being implemented, but the data is transferred over unencrypted channels, or when the system accepts default passwords.
- **Parameter manipulation** is the unauthorised interception and tampering of data while it is being transferred over a network between the consumer and the publisher. These vulnerabilities which exist when data packages are not digitally signed or encrypted to provide privacy and tamper proofing.
- **Network eavesdropping** is the ability of an third party to listen in on conversations and obtain confidential information without the knowledge of the communicators (McGraw, 2008). This is usually accomplished by using monitoring software to obtain privileged information contained in SOAP headers. This occurs if the system contains vulnerabilities like minimal encryptions on both message and transport levels, or if credential data is stored in plaintext in SOAP headers (Kumar *et al*., 2010).
- **Message relay** attack enables an unauthorised person to intercept data sent over a network, and relay it back to the publisher. Generally the attacker will change crucial information in the message, such as the delivery address, and then relay it back to the publisher without the knowledge of the consumer. Vulnerabilities in a system include messages without ID numbers, unencrypted messages and messages that are not digitally signed.

*3.3.3 Hyper-Targeted Spam*

Hayati, Potdar, Talevski, Firoozeh, Sarenche and Yeganeh (2010) define spam as the unsolicited distribution of large amounts of content (such as e-mail; instant messages) via a network to a variety of consumers without their consent. Spam has the ability of carrying infected scripts like malware, adware and viruses. The ability of Web 3.0 to integrate and link vast amounts of available metadata in a machine interoperable format, will create opportunities for a new enhanced form of spam attacks. Web 3.0 changed the method of distribution and the

intensity of the distribution of unsolicited content. Hasnain, Al-Bakri, Costabello, Cong, Davis and Heath (2012) and Ferrel (2008) describe the methods of exploitation of the platform provided by Web 3.0, as follows:

- **Application pollution.** Applications running within Web 3.0 will use the entire Web`s resources as a database. Spammers can infect a universal resource that acts as a specific database for an application. With the infection of an application`s database, spam can be distributed directly inside the application.
- **Personalisation of Web content** will enable spammers to gather more private and personal information about users. The use of this information will make the differentiation from legitimate communication increasingly difficult.
- **Improved ranking.** Ranking is the method used to determine what rank a search result will obtain when entered into a search engine. Web 3.0 will empower search engine capabilities, which will allow spammers to manipulate the ranking of malicious resources by creating triples containing malicious literal values that will be able to influence term based metrics. Complicated algorithms in conjunction with linked data are used to calculate the rank of the resource. Spammers will also attempt to exploit these algorithms by creating fake external links to resources in order to improve the resource rank.
- **Hiding.** Web 3.0 is based on open source which will enable IA's to automatically harvest information about anti-spam software, and will empower spammers to improve their method of hiding malicious content from anti-spam software.

*3.3.4 Identity Theft And Social Phishing*

Web 3.0 will introduce richer metadata, and with ontologies the integration capabilities of metadata will increase. Consumers may lose track of sensitive data available on the Web, and where it is stored, and this may lead to an increase in the precision and volume of inference attacks. Inference attack is a form of intense data mining where confidential information is harvested and disclosed by integrating non-sensitive data with metadata (Farkas & Hunhs, 2002). Phishing is a socially engineered crime through which confidential information is harvested by an unauthorised party impersonating a trusted third party. A similar threat is, identity theft, which is the process of harvesting personal information with fraudulent intent by means of exploiting information available on electronic communications mediums. Both these risks existed previously, but the precision and volumes of these threats will increase. The main threat is the ability of script writers to exploit sensitive information distributed in metadata.

**3.3.5 Autonomous Initiation Of Instructions And Malicious Script Injections**

Web 3.0 relies on different levels of languages, each with its own individual characteristics. The most common attack on Web languages takes place in the subset of Query/Update languages (Orduña, Almeida, Aguilera, Laiseca, López-de-Ipiña & Goiri, 2010). The most widely used query language in the development of Web 3.0 is SPARQL. Orduña *et al*. (2010) introduces three new types of Web 3.0 query injections:

- **SPARQL injections** are a technique used by malicious attackers to take advantage of vulnerabilities occurring in Web applications by gaining unauthorised access to the back-end layer of a database by passing non-validated SPARQL commands through a Web application (Su & Wasserman, 2006). Attackers manipulate the execution of Web application commands by structuring specific queries that enable them to harvest sensitive information within the applications' database.
- **Blind SPARQL injections.** The query languages used with Web 3.0 will consist of complicated and high level structures which will make retrieval through injection attacks much more difficult. Through blind SPARQL injections the attacker queries the database, and receives Boolean result. By querying the database repeatedly, the attacker can harvest sensitive information through true and false error messages provided by the database.
- **SPARUL injections.** SPARUL is the update version of SPARQL, and allows not only reading query abilities, but writing them as well. This creates a new threat for manipulations and extraction of data from a database, since the entire ontology can be modified through queries.

*3.3.6 Development Of Ontologies*

Ontologies (being the carriers of meaning of information available on the Web) will need to be developed to be able to interpret unified meanings of information in order to integrate information gathered from a variety of sources. An adequate infrastructure needs to be set in place to support ontology development; mapping; annotations referring to them, control over adjustments and creation of new ontologies. Benjamins and Contreras (2002) explain that the major concerns that need to be addressed in order to manage the risks associated with the development of ontologies, are the creation of kernel ontologies which will act as a unified top level dictionary. The ontologies development process will also need the necessary methodological and technological support and configuration management in order to control the creation of different versions of ontologies, and to manage the association between the ontologies and annotations.

The creation of a new technology like ontologies also poses the threat of exploitation due to inefficient knowledge of the subject during its development phase. Script writers might try to take advantage of vulnerabilities prompted by inexperience with the technology. Vulnerabilities include, but are not limited to, hidden malicious script within ontologies, and manipulation of ontologies in order to obtain sensitive data.

*3.3.7 Proof And Trust Standardisation*

Due to the ability of Web 3.0 to autonomously harvest and integrate data and convert it into information, all statements on Web 3.0 need to be considered as claims before they can be trusted (Gil & Artz, 2007). Only when these claims have been established, should trust be put in the information provided. In order to be able to trust harvested information, the source of the information, as well as the policies available on the source, needs to be obtained and analysed (Medić & Golubović, 2010).

IA's can use both the context and reputation of sources to determine the level of trust that can be put in a source (Gil & Artz, 2007). IA will be able to communicate among themselves without human interaction to determine if a source (i.e. the agent of the source) can be trusted. This creates an opportunity for malicious attackers to write scripts which impersonate a trustworthy agent, and enable them to perform unauthorised actions and inject harmful scripts.

Web 3.0 will rely heavily on semantic tagging. Script writers can manipulate semantic tagging by providing inaccurate information, and by doing so improve their Website ranking. With higher rankings more users will visit these Websites, which can be infected with various types of malware and harmful scripts.

*3.3.8 Internationalisation – Multilingualism*

The challenge of sharing information from different geographical areas in a common language understandable by all participating entities, has been prevalent in Web 2.0. Web 3.0 will also be affected by the risk of multilingualism. According to Benjamins *et al*. (2002) multilingualism will affect some of the following areas of Web 3.0:

- **Ontologies** will be one of the cornerstones of Web 3.0, and developers will need to develop ontologies in their native language.
- **Annotation** and the description of content will be a bigger challenge since most of the Web community will take part in annotating its specific content. These annotations will have to be in detail and precise in order for Web 3.0 to realise. Consumers will only be able to become contributors if proper support is created to enable them to annotate in their native language.

Language boundaries will suppress the ability of consumers and contributors to describe and integrate content. This will increase the risk of non-interoperability and miscommunication between applications and IA's.

Many of the risks identified in this section were prevalent in Web 2.0, but due to the addition of new technologies with original and unknown structures, additional risks will arise. New controls and methods of

regulating activity on Web 3.0 will need to be adopted in order for these new technologies to operate effectively and accurately.

### 3.4 Safeguards And Controls To Mitigate Risks

Web 3.0 will be an environment where data and information will be shared openly and interactively. Information will be much more valuable to organisations and consumers, and it needs to be treated and protected as an asset. The consumption of information by machines will affect the level of control an organisation can exercise over securing valuable information on the Web. Controls need to be revisited and modified in order to mitigate the new risks.

#### *3.4.1 Controls*

The threats can be controlled through the use of technological methods with the inclusion of an administrative component.

##### 3.4.1.1 Technological Control

The following technological controls can be implemented:

- **Encryption and authentication.** Communication between a Web server and a consumer is controlled by input and output parameters. These parameters need to be encrypted. This can be accomplished by using a semantic mark-up that specifies the security characteristics of the Web servers' input and output parameters (Kagal, Paolucci, Srinivasan, Denker, Finin & Sycara, 2004). This will enable the data to keep its structure while not revealing the values, after which matchmaking services can select the service required by using this meta-information. Lee and Whang (2006) stress the fact that encryption mechanisms like Secure Sockets Layer (SSL) and Transport Layer Security (TLS), which are commonly used during transmission of HTTP protocols, are not sufficient for XML or RDF transmissions. XML Encrypt (XMLEcn) and XML Signature (XMLSig) are standards used to encrypt XML data and ensure that data which is transmitted, is authenticated through the use of digital signatures. Partial RDF encryption (PRE), as suggested by Giereth (2005), recommends an algorithm which encrypts fragments of RDF content to ensure secure transmission of data.
- **Access control.** Web 3.0 will grant access to resources through SPARQL query language, which will also be the main exploit malicious attackers will use to gain unauthorised access. Social Semantic SPARQL Security for Access Control (S4AC), is an access control vocabulary for SPARQL query language (Villata, Delaforge, Gandon, & Gyrard, 2011). It enables users to invoke access control policies for their RDF data by using a SPARQL 1.1 ASK clause that enables them to specify certain conditions within metadata tags that need to be met before access is given to resources. ROWLBAC (Finin, Joshi, Kagal, Niu, Sandhu, Winsborough & Thuraisingham, 2008) and SecurOntology (García-Crespoa, Gómez-Berbísa, Colomo-Palaciosa & Alor-Hernández, 2011) are other methods to ensure secure access of OWL ontologies.
- **Guaranteed e-delivery**. This method promises the secure delivery of electronic information, and ensures through follow-up controls that the information has not been tampered with during transmission. It is based on the combination of secure hosting and e-mail notifications. The recipient of the information will receive an e-mail that informs him that information has been sent, and that delivery is pending, while the sender will receive notification when the information has been retrieved. This enables crosschecks of information sent electronically (Gilbert, Abrams, Linden, Mogull, Orans & Wald, 2001).
- **Effective spam filtering.** The ability of ontologies to enable a machine to understand and reason with content, creates an opportunity to filter incoming and outgoing data more effectively. Youn and McLeod (2007) suggested controls can be implemented by mapping a decision tree into an ontology, thus enabling a server to understand incoming and outgoing messages, and based on rules set, filtering out spam emails. Eyharabide and Amandi (2012) introduced a next level of integration between the use of ontologies and spam filtering by not only using ontologies to filter spam, but also to personalise filter based on user preferences harvested by ontologies.

Dietzold and Auer (2006) suggest the implementation of a Lightweight Framework for Access Control which consists of an information query filter that is based on the quality and trust properties of resources. The framework controls access through a query engine that is limited by a rule processor which decides whether the query filter can trust a resource. The implementation of a Semantic Data Crawler which only harvests documents and information from predefined semantic Web data sources, can also assist in filtering out data retrieval from untrustworthy and malicious resources (Lašek & Vojtáš, 2011).

- **Anti-malware software.** Anti-malware software, including anti-spy and -virus software, should be installed in order to eliminate the threat of malicious infections of a system of both inbound and outbound electronic transmissions (Rudman, 2010). The implementation of a structure like Taiwan Malware Analysis Net (TWMAN) will mitigate the risk of infections associated with Web 3.0. TWMAN consists of two IA, a malware behavioural analysis agent and an ontology agent. By integrating the information collected by the ontology agent, the malware behavioural analysis agent collects malware behavioural information and builds a malware behavioural ontology with malware behavioural rules. This creates a ubiquitous software agent that examines malware behaviour, and autonomously creates rules to protect the system from infection in real time (Huang, Lee, Kao, Tsai & Chang, 2011). Chiang and Tsaur (2010) introduce the same concept of ontology based malware protection for the safeguarding of mobile devices.
- **Monitoring of composite events.** User history logs that define all types of network activity, which includes bandwidth usage, sites visited and files downloaded should be kept and reviewed regularly. Complex and composite events within a Web service should be monitored. This will simplify the process of sharing events and content between IA's, and will contribute to the standardisation and internationalisation of Web content (Vaculín and Sycara, 2007; Rudman, 2010).
- **Network and underlying infrastructure security.** Most of the security surrounding access control is focused on the application domain. Action needs to be taken to ensure that the users' system also protects the transmission of electronic information. Foley and Fitzgerald (2008) suggest the implementation of firewall IA's which act like IA with parameters set to support firewall configuration. The firewall IA's will negotiate the firewall settings, controlling access based on a knowledge repository that is gathered from ontologies. This repository will be constantly updated and controlled by facts collected by IA's as they harvest information from new knowledge and inferences.
- **Validation of input.** The most common attack to bypass input validation is SQL injections, which grant the attacker unrestricted access to database information. Halfond, Viegas and Orso (2006) suggest the implementation of a New Data Validation Service (NDVS) using Web 3.0. The process of validating input starts by using ontology to describe all the data in the Web application through RDF annotation. When the user requests a HTML, the interceptor will intercept it before it gets processed at the server side. The RDF annotations are extracted from the RDF ontology by the RDF extractor, and compared to the user inputs by the validator to ensure valid inputs. If the validation is correct the request is processed.

3.4.1.2 Keep users and developers informed about the risks

A predominant threat when adopting a new technology is users engaging in these technologies without training and guidance as to what the technology entails, and why there are risks associated with it. Developers as well as consumers need to be educated and made aware of these risks and should cover:

- **A better understanding of the issues**. Both consumer and developer need to understand the risks arising from Web 3.0. Collaboration is needed between the two parties to ensure that not only technical solutions are implemented, but consumer contributed solutions like Web interfaces, ease of use of applications and informative applications, are investigated as well.
- **Developers design methodologies.** Designers must be aware of the preceding risks before they start developing Web applications. During development stages technical solutions must be integrated into the design of the application to ensure secure and accurate execution of Web applications.
- **Education of consumers.** Consumers need to be educated on the issues surrounding Web 3.0, which includes technical understanding, identification of potential risks, and social rules and expectations. This can be introduced through the use of different approaches, like long term learning and maintenance campaigns, or by incorporating it into Web interfaces, which offers contextual tours to consumers to

highlight risk areas within the application. Consumers also need to be educated about the major safeguards that are associated with the adoption of new technologies, like the identification of risks relating to the new applications; refraining from interacting with suspicious applications and attachments of communication from unknown sources, and using security features embedded in applications, browsers and underlying infrastructure (firewalls, anti-malware software).

3.4.1.3 Policies And Guidelines Controlling Use

Organisations need to adopt regulating policies and guidelines to ensure users are educated and informed on why and how they are allowed to use certain technologies. When adopting these policies, organisations should not only consider written policies, but also automated policy creation through the use of semantic technologies. Policies implemented should not only support the organisations' strategic objectives, but also comply with regulatory governance. The policies should be clear and communicated to all levels of employees in a non-technical language. The policies should be written in a manner which makes it adaptive to changes, and it should be reviewed and adapted by top management on a regular basis.

3.4.1.4 Setting Electronic Parameters

Kagal *et al*. (2003) suggest the implementation of distributed policy management through the use of a semantic policy language. This policy language will be based on ontologies written with specific parameters. The language will have some domain independent ontologies, but will also require specific domain ontologies. The language will consist of two constructs that will be able to select meta-policies, which will be invoked to resolve conflicts. The first construct is a modality preference (negative over positive), and the second is priorities construct (when there is more than one policy applicable, this construct will invoke the policy that enjoys priority). The use of ontologies to create policies will enable policies to be much more adaptive to real time changes, and by integrating it with IA, ensure that policies are more effective and applicable.

Organisations are always eager to adopt the new technologies based on the opportunities they might offer. With the adoption of all new technologies new risks will arise which are rarely taken into account when making these decisions (Hall & Khan, 2002) or when implementing controls.

## 4. CONCLUSION

Modern organisations operate in a highly technological environment, rapidly and continuously evolving where technology plays a vital part in an organisations success. Much of this technology is driven by the Web. The introduction of Web 3.0 will create new opportunities that will assist organisations in reaching their objectives. With the adoption of new technologies, developers, consumers management tend to focus on the benefits, and ignore the risks associated with the implementation of these technologies. The research shows that in order to understand the true impact of implementing new technologies, the underlying infrastructures that support these technologies, must be defined. COBIT 5 was used as a governance framework to identify the risks associated with Web 3.0 and to recommend control processes that can lower the threat to an acceptable level. Table 1 maps the incremental risks associated with Web 3.0 identified during the research, and possible safeguards that can be implemented to mitigate these risks.

**Table 1.** Risks And Safeguards Associated With Web 3.0 Identified

| SAFEGAURDS \ RISKS | Unauthorised access to sensitive information | Parameter manipulation | Network eavesdropping | Message relay | Application pollution | Unauthorised ranking improvement | Hiding from anti-malware software | Personalisation of web content | Identity theft and social phishing | SPARQL injection | Blind SPARUL injection | Development of ontologies | Proof and trust | Standardisation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **XML and RDF encryption** | ■ | ■ | ■ | ■ | ■ | | | | | ■ | ■ | | | |
| **Access controls (ROWLBAC and S4AC)** | ■ | ■ | ■ | ■ | ■ | | | ■ | ■ | ■ | ■ | | | |
| **Guaranteed e-delivery** | ■ | | ■ | ■ | | | ■ | | | | | | | |
| **Spam filtering** | | | | | | ■ | ■ | ■ | | | | | ■ | |
| **Ontology and agent driven anti-malware (TWMAN)** | ■ | ■ | ■ | | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | |
| **Monitor and review logs** | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | | ■ | ■ |
| **Filtering and semantic data crawlers** | ■ | ■ | ■ | ■ | ■ | | | ■ | ■ | ■ | ■ | | ■ | |
| **Firewall agents** | ■ | ■ | ■ | ■ | ■ | | | ■ | ■ | | | | ■ | |
| **Input validation through NDVS** | ■ | ■ | ■ | ■ | | | | ■ | ■ | ■ | ■ | | ■ | |
| **Education of consumers and developers** | | | | | | | | ■ | ■ | | | ■ | ■ | ■ |
| **Implementation of Web 3.0 usage policy and agent based policies** | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | |

Web 3.0 is not a separate or isolated technology, but rather a compilation of already existing principles amalgamated with new programs and scripts. The underlying technologies are accompanied by risks specifically linked to these technologies. The main risks identified are as follows:

- Unauthorised access to sensitive data, or data manipulation by unauthorised persons.
- New and more complicated electronic attacks, such as SQL injections, malware, hyper targeted spam and internet ranking manipulation.

- Personalisation of Web content creates a situation where personal and sensitive data will be more widely available on the Web, thus creating an increased risk of identity theft and social phishing.
- The development and standardisation of new Web ontologies and languages increase the probability of releasing inferior or easily targeted software and application, due to a lack of knowledge and insufficient testing for risks associated with the technology.

A multilayer approach of technical and non-technical safeguards should be implemented before Web 3.0 are adopted in order to ensure that the risks associated with Web 3.0 are mitigated to an acceptable level, include:

- Technical safeguards include encryption and authentication; anti-malware software; physical and logical access controls; effective filtering and monitoring, and validation controls.
- Non-technical safeguards include ensuring that users are educated continuously about the underlying risks associated with the use of Web 3.0, and the communication of methods on how to avoid threatening situations.
- A policy regarding the use of Web 3.0 should be implemented that pins down responsibility, and elaborates on what actions are expected from consumers when using Web 3.0.

This framework program above outlines principles and procedures that could be used as a starting point to mitigate these 'new' Web 3.0 risks to an acceptable level.

## AUTHOR INFORMATION

**Mr Rikus Bruwer** is a lecturer at Stellenbosch University, South Africa. He lectures at an under- and post-graduate level. His area of interest lie in business risks associated with new technological developments and the management thereof. E-mail: rikus.bruwer@gmail.com

**Mr Riaan Rudman** is a Senior Lecturer at Stellenbosch University, South Africa. He lectures at an under- and post-graduate level. He spesialised in Financial Institutions before joining academia. His areas of interest lie in business management and acceptable corporate behaviour in an electronic environment and new technologies. E-mail: RJRudman@sun.ac.za (Corresponding author)

## REFERENCES

Alpert, J. & Hajaj, N. (2008), *We knew the Web was big*. Retrieved from: http://googleblog.blogspot.com/2008/07/we-knew-Web-was-big.html.

Bakshi, K. & Karger, D.R. (2005), *End-user application development for the semantic Web*, research paper, ISWC Workshop on the Semantic Desktop - Next Generation Information Management and Collaboration Infrastructure, pp. 123–137.

Benjamins, R. & Contreras, J. (2002), *Six Challenges for the Semantic Web,* white paper,Intelligent Software Components, Intelligent Software for the Networked Economy, April. Retrieved from: http://oa.upm.es/5668/1/Workshop06.KRR2002.pdf.

Bergman, M. (2001), *The Deep Web: Surfacing Hidden Value*, white paper, The Journal of Electronic Publishing, August. Retrieved from: http://quod.lib.umich.edu/cgi/t/text/text-idx?c=jep;view=text;rgn=main;idno=3336451.0007.104.

Berners-Lee, T (1998), *Why RDF model is different from the XML model*, World Wide Web Consortium. Retrieved from: http://www.w3.org/DesignIssues/RDF-XML.html.

Berners-Lee, T., Hendler, J. & Lissila, O (2001), *The Semantic Web [Preview]*, Scientific America. Retrieved from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=539724.

Brynjolfsson, E. & Hitt, L.M. (2000), *Beyond Computation: Information technology, Organizational Transformation and Business Performance*, Journal of Economic Perspectives, Vol. 14 No. 4, pp. 23-48.

Chiang, H. & Tsaur, W. (2010), *Mobile Malware Behavioral Analysis and Preventive Strategy Using Ontology*, IEEE International Conference on Social Computing/IEEE International Conference on Privacy, Security, Risk and Trust, pp. 1080-1085. Retrieved from: http://www.lasr.cs.ucla.edu/classes/239_1.fall10/papers/ontology.pdf.

Clearswift (2007), *Demystifying Web 2.0*, white paper, Clearswift Limited, July. Retrieved from:http://resources.clearswift.com/ExternalContent/C12CUST/Clearswift/9514/200707.

Cormode, G. & Krishnamurthy, B. (2008), *Key differences between Web 1.0 and Web 2.0*, First Monday, Vol. 13 No. 6. Retrieved from: http://firstmonday.org/ojs/index.php/fm/article/view/2125/1972.

Dawson, R. (2007), *Web 2.0 framework*, [blog]. Retrieved from: www.rossdawsonblog.com/Web2_.

Dawson, R. (2008), *An enterprise 2.0 Governance Framework-looking for input!*, [blog]. Retrieved from: http://rossdawsonblog.com/Weblog/archives/2008/02/an_enterprise_2.html.

DCruz, T. (2009), *Difference Between Web 1.0, Web 2.0 and Web 3.0*, Enzine Articles, Retrieved from: http://ezinearticles.com/?Difference-Between-Web-1.0,-Web-2.0-and-Web-3.0&id=2941533.

Decker, S., Melnik, S., van Harmelen, F., Fensel, D., Klein, M., Broekstra, J., Erdmann, M., & Horrocks, I. (2000), *The Semantic Web: The Roles of XML and RDF*, IEEE Internet Computing, Vol. 4, pp. 63-74. Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.81.6109&rep=rep1&type=pdf.

Dietzold, S. & Auer, S. (2006), *Access control on RDF triple stores from a semantic wiki perspective*, The Semantic Web Workshop at 3rd European Semantic Web Conference (ESWC). Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.3708&rep=rep1&type=pdf.

Evans, D. (2011), *The Internet of Things: How the Next Evolution of the Internet Is Changing Everything*, white paper, Cisco Internet Business Solutions Group (IBSG), April. Retrieved from: http://www.cisco.com/Web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf.

Eyharabide, V. & Amandi, A. (2012), *Ontology-based user profile learning*, Applied Intelligence, Vol. 36 No. 4, pp. 857-869. Retrieved from: http://dl.acm.org/citation.cfm?id=2011270.

Farah, J. (2012), *Predicting the Intelligence of Web 3.0 Search Engines*, International Journal of Computer Theory and Engineering, Vol. 4 No. 3, pp. 443-445. Retrieved from: http://www.ijcte.org/papers/503-G1326.pdf.

Farkas, C. & Huhns, M.N., 2002; *Making Agents Secure on the Semantic Web*, IEEE Internet Computing, Vol. 6 No. 6, pp. 76-79. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1067741.

Finin, T., Joshi, A., Kagal, L., Niu, J., Sandhu, R., Winsborough, W. and Thuraisingham, B. (2008), *Rowlbac: Role based access control in OWL*, ACM Symposium on Access Control Models and Technologies (SACMAT). Retrieved from: http://ebiquity.umbc.edu/_file_directory_/papers/391.pdf.

Foley, S.N. & Fitzgerald, W.M. (2008), *Semantic Web and Firewall Alignment*, First International Workshop on Secure Semantic Web (SSW'08). Retrieved from: http://www.cs.ucc.ie/~simon/pubs/ssw08.pdf.

García-Crespoa, A., Gómez-Berbísa, J.M., Colomo-Palaciosa, R. & Alor-Hernández, G. (2011), *SecurOntology: A semantic Web access control framework*, Computer Standards & Interfaces, Vol. 33 No. 1, pp. 42-49. Retrieved from: http://www.sciencedirect.com/science/article/pii/S0920548909000798.

Getting, B. (2007), Basic Definitions: Web 1.0, Web. 2.0, Web 3.0. Retrieved from: www.practicalecommerce.com/articles/464/Basic-Definitions:-Web-1.0,-Web.-2.0,-Web-3.0/.

Ghaleb, F., Daoud, S., Hasna, A., ALJa'am, J.M., El-Seoud, S.A. & El-Sofany, H. (2006), *E-Learning Model Based On Semantic Web Technology*, International Journal of Computing & Information Sciences, Vol. 4 No. 2, pp. 63-71. Retrieved from: http://www.ijcis.info/Vol4N2/pp63-71.pdf.

Giannakos, N., & Lapatas, V. (2010), *Towards Web 3.0 Concept For Collaborative e-Learning*, research report, International Multi-Conference on Innovative Developments in ICT. Retrieved from: http://www.academia.edu/417958/Towards_Web_3.0_Concept_for_Collaborative_e-Learning.

Giereth, M. (2005), *On Partial Encryption of RDF-Graphs*, The Semantic Web – ISWC 2005 Lecture Notes in Computer Science, Vol. 3729, pp. 308-322.

Gil, Y. & Artz, D. (2007), *Towards Content Trust of Web Resources*, Journal of Web Semantics: Science, Services and Agents on the World Wide Web, December. Retrieved from: http://www.isi.edu/~gil/papers/gil-artz-jws07.pdf.

Gilbert, D. (1997), *Intelligent Agents: The right Information at the Right Time*, white paper, IBM Corporation, May. Retrieved from: https://fmfi-uk.hq.sk/Informatika/Uvod%20Do%20Umelej%20Inteligencie/clanky/ibm-iagt.pdf.

Gilbert, M.R., Abrams, C., Linden, A., Mogull, R., Orans, L. & Wald, B. (2001), *Emerging Technologies for Managing Content*, research report, Gartner, 28 September. Retrieved from: http://my.gartner.com/portal/server.pt?open=512&objID=260&mode=2&PageID=3460702&resId=341672&ref=QuickSearch&sthkw=Guaranteed+e-delivery.

Golbreich, C., & Wallace. E.K. (2012), *OWL 2 Web Ontology Language: New Features and Rationale (Second Edition)*, World Wide Web Consortium. Retrieved from: http://www.w3.org/TR/owl2-new-features/.

Grossman, J. (2007), *Seven Business Logic Flaws That Put Your Website At Risk*, white paper, WhiteHat Security, October. Retrieved from: https://www.whitehatsec.com/assets/WP_bizlogic092407.pdf.

Halfond, W.G., Viegas, J. & Orso, A. (2006), *A Classification of SQL-Injection Attacks and Counter Measures*, The International Symposium on Secure Software Engineering, March. Retrieved from: http://www.cc.gatech.edu/fac/Alex.Orso/papers/halfond.viegas.orso.ISSSE06.pdf.

Hall, B., & Khan, B. (2002), *Adoption of new technology*, D.C. Jones (Ed.), New Economy Handbook, Elsevier Science, pp. 230–251. Retrieved from: http://emlab.berkeley.edu/~bhhall/papers/HallKhan03%20diffusion.pdf.

**Copyright by author(s); CC-BY** 

Hasnain, A., Al-Bakri, M., Costabello, L., Cong, Z., Davis, I. & Heath, T. (2012), *Spamming in Linked Data*, Third International Workshop on Consuming Linked Data (COLD2012). Retrieved from: http://ceur-ws.org/Vol-905/HasnainEtAl_COLD2012.pdf.

Hayati, P., Potdar, V., Talevski, A., Firoozeh, N., Sarenche, S. & Yeganeh, E. A. (2010), *Definition of spam 2.0: New spamming boom*, 4th IEEE International Conference on Digital Ecosystems and Technologies, pp. 580-584. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5610590&tag=1.

Horrocks, I. (2004), *OWL Rules, OK?*, World Wide Web Consortium. Retrieved from: http://www.w3.org/2004/12/rules-ws/paper/42/.

Huang, H.D., Lee, C.S., Kao, H.Y., Tsai, Y.L. & Chang, J.G. (2011), *Malware behavioral analysis system: Twman*, IEEE Symposium on Computational Intelligence for Intelligent Agents. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5953604.

Hussain, S. J. & Siddiqui, M. S. (2005), *Quantified Model of COBIT for Corporate IT Governance*, First International Conference on Information and Communication Technologies, Karachi. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1598575&tag=1.

Internet World Stats (2012), Internet usage statistic. Retrieved from: http://www.internetworldstats.com/stats.htm.

Intervise (n.d.), *The Semantic Web Information with Knowledge*, white paper, Intervise Consultants Incorporated. Retrieved from: http://www.semantic-experts.com/galleries/default-file/White%20Paper%20Semantic%20Web.pdf.

ISACA (2012), *COBIT 5 Enabling Processes*, ISACA. Retrieved from: http://www.isaca.org/COBIT/Documents/COBIT5-Ver2-enabling.pdf.

Kagal, L., Finin, T. & Joshi, A. (2003), *A Policy Based Approach to Security for the Semantic Web*, 2nd International Semantic Web Conference (ISWC2003), September. Retrieved from: http://www.csee.umbc.edu/~finin/papers/papers/iswc03b.pdf.

Kagal, L., Paolucci, M., Srinivasan, N., Denker, G., Finin, T. & Sycara, K. (2004), *Authorization and privacy for semantic Web services*, Spring Symposium on Semantic Web Services. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1333035&tag=1.

Knublauch, H., Fergerson, R.W., Noy, N.F., & Musen, M.A. (2004), *The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications*, Lecture Notes in Computer Science, Vol. 3298, pp.229-243.

Kumar, S., Prajapati, R.K., Singh, M. & De, A. (2010), *Realization of Threats and Countermeasure in Semantic Web Services*, International Journal of Computer Theory and Engineering, Vol. 2 No. 6, pp. 919-924. Retrieved from: http://www.ijcte.org/papers/264-G796.pdf.

Lašek, I. & Vojtáš, P. (2011), *Semantic Information Filtering - Beyond Collaborative Filtering*, 4th International Semantic Search Workshop. Retrieved from: http://km.aifb.kit.edu/ws/semsearch11/11.pdf.

Lawler, P. & Molluzzo, C. (2010), *A Study of the Perceptions of Students on Privacy and Security on Social Networking Sites (SNS) on the Internet*, Journal of Information Systems Applied Research, Vol. 3 No. 12, pp. 1-18.

Lee, J. & Whang, K. (2006), *Secure query processing against encrypted XML data using Query-Aware Decryption*, Elsevier, Information Sciences,pp. 1928-1947. Retrieved from: http://dm.kaist.ac.kr/jaegil/papers/infosci06.pdf.

Lewis, D.J. (2008), *Intelligent agents and the Semantic Web*, developerWorks. Retrieved from: http://www.ibm.com/developerworks/library/wa-intelligentage/.

Lu, S., Dong, M., Fotouhi, F. (2002), *The Semantic Web: Opportunities and challenges for next-generation Web applications*, Information Research, Vol. 7 No. 4. Retrieved from: http://informationr.net/ir/7-4/paper134.html.

McGraw, G. (2008), *Software [In] security: Securing Web 3.0*, InformIT. Retrieved from: http://www.informit.com/articles/article.aspx?p=1217101.

Medić, A. & Golubović, A. (2010), *Making secure Semantic Web*, Universal Journal of Computer Science and Engineering Technology, Vol. 1 No. 2, pp. 99-104. Retrieved from: http://www.unicse.org/publications/2010/november/Making%20secure%20Semantic%20Web.pdf.

Melton, J. & Eisenberg, A. (2001), *Sql multimedia and application packages (sql/mm)*, SIGMOD Record, Vol. 30 No. 4. Retrieved from: http://delivery.acm.org/10.1145/610000/604280/p97-melton.pdf?ip=146.232.41.252&id=604280&acc=ACTIVE%20SERVICE&key=C2716FEBFA981EF194FAE17EEFF9EA9FD4E9E3A22F5A9081&CFID=256988813&CFTOKEN=82882342&__acm__=1382950300_f55d44317d0764a211ffd33048214af9.

Morris, R.D. (2011), *Web 3.0: Implications for Online Learning*, TechTrends, Vol. 55 No. 1, pp. 42-46. Retrieved from: http://download.springer.com/static/pdf/260/art%253A10.1007%252Fs11528-011-0469-9.pdf?auth66=1382776409_7aa9dbcb343b548f43658d5f1dc6709a&ext=.pdf.

Naeve, A., Lytras, M., Nejdl, W., Balacheff, N. & Harding, J. (2006), *Advances of semantic Web for e-learning: Expanding learning frontiers*, British Journal of Education Technology, Vol. 37 No. 3, pp.321-330. Retrieved from: http://www.noe-kaleidoscope.org/public/pub/news/0505/CFP-BJET-SW-for-EL.pdf.

Nematzadeh, A., Pournajaf, L. (2008), *Privacy Concerns of Semantic Web*, Fifth International Conference on Information Technology: New Generation, .pp.1272-1273. Retrieved from: http://www.mathcs.emory.edu/~lpourna/papers/2008_itng.pdf.

Noy, N.F., Sintek, M., Decker, S., Crubézy, M., Fergerson, R.W., & Musen, M. (2001), *Creating Semantic Web contents with Protege-2000*, IEEE Intelligent Systems, Vol. 16 No. 2, pp. 60-71. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=920601&tag=1.

O'Reilly, T. (2007), *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*, Communications & Strategies, No. 1, p. 17, First Quarter 2007. Retrieved from: http://ssrn.com/abstract=1008839.

O'Reilly, T. (2009), *What is Web 2.0*, California: O'Reilly Media Incorporated. Retrieved from: http://books.google.co.za/books?hl=en&lr=&id=NpEk_WFCMdIC&oi=fnd&pg=PT1&dq=defining+Web+1.0&ots=OXSFR8jzH-&sig=t6ZW8qlOh1q6Nbve1LA7f6VbVok#v=onepage&q=defining%20Web%201.0&f=false.

Orduña, P., Almeida, A., Aguilera, U., Laiseca, X., López-de-Ipiña, D. & Goiri, A.G. (2010), *Identifying Security Issues in the Semantic Web: Injection Attacks in the Semantic Query Language*, research paper, DeustoTech. Retrieved from: http://www.morelab.deusto.es/publications/2010/orduna2010identifying.pdf.

Prescott, B. (2012), *Business Sense: Inbound marketing*, Times Standard. Retrieved from: http://www.times-standard.com/business/ci_19898286.

Quilitz, B., & Leser, U (2008), *Querying distributed RDF data sources with SPARQL*, Proceedings of the 5th European Semantic Web Conference, ESWC 2008, Tenerife, Canary Islands, Spain. Retrieved from: http://dl.acm.org/citation.cfm?id=1789443&CFID=256988813&CFTOKEN=82882342.

Rouse, M. (2005), *SOAP (Simple Object Access Protocol)*, SearchSOA. Retrieved from: http://searchsoa.techtarget.com/definition/SOAP.

Rudman, R.J. (2010), *Incremental risks in Web 2.0 applications*, The Electronic Library, Vol. 28 No. 2, pp. 210-230. Retrieved from: http://www.emeraldinsight.com/journals.htm?articleid=1853058.

Sampson, D. G., Lytras, M. D., Wagner, G. & Diaz, P (2004), *Ontologies and the Semantic Web for E-learning*, The Journal for Educational Technology & Society, Vol. 7 No. 4, pp. 26-28. Retrieved from: http://www.ebiblioteka.lt/resursai/Uzsienio%20leidiniai/IEEE/English/2006/Volume%207/Issue%204/Jets_v7i4.pdf#page=87.

Sheth, A. & Meersman, R. (2002), *Amicalola Report: Database and Information Systems Research Challenges and Opportunities in Semantic Web and Enterprises*, ACM SIGMOD Record, Vol. 31 No. 4, pp. 98-106. Retrieved from: http://www.sigmod.org/publications/sigmod-record/0212/R1.Amicalola_Final_Report.pdf.

Sowa, J.F. (2009), *Building, Sharing, and Merging Ontologies. Retrieved from*: http://www.jfsowa.com/ontology/ontoshar.htm.

Stoneburner, G., Goguen, A., & Feringa, A. (2002), *Risk management guide for information technology systems: Recommendations of the national institute of standards and technology*, National Institute of Standards and Technology (NIST), Special Publication 800-30. Retrieved from: http://www.security-science.com/pdf/risk-management-guide-for-information-technology-systems.pdf.

Sylvester, A., Tate, M. & Johnstone, D. (2011). *Beyond synthesis: re-presenting heterogeneous research literature*, Behaviour & Information Technology. Retrieved from: http://www.tandfonline.com/doi/abs/10.1080/0144929X.2011.624633#.UmZQO3B_PQh.

Tarrant, D., Hitchcock, S., & Carr, L. (2011), *Where the Semantic Web and Web 2.0*

Verizon (n.d.), *Web 3.0: Its Promise and Implications for Consumers and Business*, white paper, Verizon Business. Retrieved from: http://www.verizonenterprise.com/resources/whitepapers/wp_Web-3-0-promise-and-implications_a4_en_xg.pdf.

Villata, S., Delaforge, N., Gandon, F. & Gyrard, A. (2011), *Social semantic Web access control*, 4th International Workshop Social Data on the Web (SDoW2011). Retrieved from: http://sdow.semanticWeb.org/2011/pub/sdow2011_paper_5.pdf.

Webschool (2013), *Web Service Tutorial*, Webschool.com. Retrieved from: http://www.w3schools.com/Webservices/default.asp.

Wolfram, C. (2010) Interviewed by Nicole Kobie on *Communicating with Apps in Web 3.0*, IT Pro, 17 March. Retrieved from: http://www.itpro.co.uk/621535/qa-conrad-wolfram-on-communicating-with-apps-in-Web-30.

Youn, S. & McLeod, D. (2007), *Efficient spam email filtering using adaptive ontology*, International Conference on Information Technology, pp. 249-254. Retrieved from: http://www.academypublisher.com/jsw/vol02/no03/jsw02034355.pdf.