CHAPTER II

LITERATURE REVIEW

2.0 INTRODUCTION

The chapter of literature review forms the basis of analysis of the empirical data and outlines for the basic concepts which are used in this study. In short, it defines a search and evaluation of the available literature specifically in the areas of ontology and social engineering. The chapter process involves survey synthesis, analysis and presentation of information obtained from various literatures to define what the social engineering is, and what constitutes its taxonomy.

2.1 WHAT IS SEMANTIC WEB

As a huge information space, the web should be useful not only for human-human communication, but it should also allow machines to participate and help. However, nowadays most of the information on the web is designed for human consumption and the structure of the data is not evident for a robot who might be browsing the Web. There are two distinct approaches to enable machines to automatically manipulate the information in the Web. One approach, which comes from artificial intelligence, is machine learning. The machine is trained to behave like a person. However, this approach depends on domains and requires a huge training

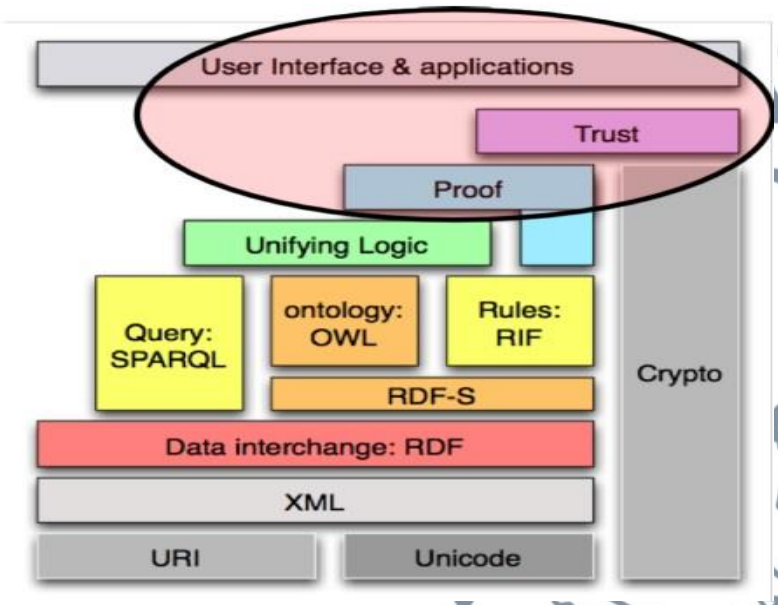
process. The semantic web (Berners-Lee et al, 2001) approach instead develops language for expressing information in a machine processible form.

The W3C gives the following definition for the Semantic Web:

"The Semantic Web is an extension of the current Web in which information is given a well-defined meaning, better enabling computers and people to work in cooperation".

The technologies used for the semantic web are usually represented by a stack of layers; each layer depends on the layers below it shown in figure 2.1 (Antoniou and Harmelen, 2008). Nevertheless, the basic premises remain the same; everything is built upon the basic web markup conventions (Unicode, URI) and data formats (XML, RDF). On this the foundations for machine intelligence are built using semantic web logic, rule and query languages and technologies (RDFS, OWL, RIF, SPARQL), applied data subjectiveness (proof, trust) and finally the user interfaces and the ways for a human to actually dive in to the world of intelligent data. Throughout the stack is the regard for privacy and security: data encryption and other cryptographic measures concern many different levels on the stack.

FIGURE 1: Semantic web layers.



As a user searches for the content of those files such as ones that talk about people, sports, companies, events, science, countries, food, evolution, etc also there are many threads talking about individuals or whatever, in one word "entity". The purpose of intelligence and development here is to enable the computer and web to know what is located inside those files. This comprises what is called the Semantic Web (Thomas Baker, 2005).

a) Semantic Web is a new revolution in the world of the Web, where information and data are logically become treated by computer programs. So that information and data turn into a meaningful data network, "computer programs are able to know what this data means".
b) Semantic Web is a set of methods and techniques used to make the machines which are able to understand the meanings or "semantics" of information on the World Wide Web.

c) Semantic Web is network of data that can be processed by machines, either directly or indirectly.

d) Semantic Web helps to make the Web more understandable by machines.

e) Semantic Web helps to build suitable infrastructure for intelligent agents to carry out a complex operations for their users.

f) Semantic Web is an explicit definition of the information found in many web applications, and integrates information in an intelligent manner, to provide a semantic access to Internet, and extract knowledge from texts.

2.2 WHAT IS ONTOLOGY

In this research, we can further our study of the meaning of ontology. Ontology is an explicit specification of a conceptualization (Gruber, 1993). Ontology definition is proportionate with the usage of ontology as a concept of definitions set, but is more general. And it is undoubtedly a different meaning of the word in computer science than its use in philosophy.

An improving of the world wide web has led to researchers finding a tremendous amount of knowledge and information. One of the significant problems that face the world of knowledge is how to reuse and retrieve these resources. Tim Berners promised that there will be a proposed approach to rehabilitation of the web as it exists in relation to the semantic web. As Berners says, "The Semantic Web is an

extension of the current Web in which information is given well-defined meaning" (Berners et al, 2001).

Within ontology, a conceptualization refers to a facilitated view of the objects and things that we desire to represent for different purposes. It is an abstract of the world view. All systems of knowledge, whether they were level agents or based systems, are obliged to some conceptualization, implicitly or explicitly. In other words, an obligation to common ontology is a warranty of uniformity and consistency but not to that wholeness.

Ontology languages are formal languages that are used to construct ontologies in computer science and allow the knowledge encoding about specific domains. There are a number of languages for ontologies, both proprietary and standard-based. As a language for the World Wide Web, XML is easy to parse, its syntax is well defined and it is human and machine readable. They allow users to define their own tags and attributes, define data structures, extract data from documents and develop applications which test the structural validity of XML documents. It also has a system schema that can be used to define input parameters and constraints (Mello and Xu, 2006).

An RDF data model is developed by the W3C for the creation of metadata describing resources on the web (Beckett, 2004). According to Heflin and Hendler (2000), 'RDF has as its aim to specify semantics for data based on the XML format in a standardized, interoperable manner and to define a mechanism for describing resources that makes no assumption about a particular application domain or the structure of a document containing the information'. It is written in XML and designed to be read

and understood by machines. The basic constituent of the RDF data model is triple, which has a subject-predicate-object structure. Each part of the triple is referred to the URI (Uniform Resource Identifier), which provides universal identifier to name these resources. The subject of the triple is a real-world object or abstract concept. The object can be either an object or a literal or XML Schema data type such as string and integer. The predicate is the property or relation that relates the subject and object resources. Predicates define specific aspects, characteristics, attributes or relations used to describe a resource. The object of the triple assigns a value for the property of triple's subject (resources) (Klyne and Carroll, 2004).

RDF is not enough to make a relationship among concepts. That's why RDF Schema has an important role on that point. The RDFS data model allows defining the relationships between properties and other entities like resources. It also provides main elements for ontology development to create sharable, controlled and extensible vocabularies.

OWL, which stands for Web Ontology Language, is a language for Web knowledge processing. It is designed to process Web information and easily read by machine. RDF and OWL show some similarity, but OWL is more advanced and stronger. Since it is written in XML, information in OWL can be transformed in different systems (W3C OWL Working Group, 2009). Although XML, RDF and RDFs are the basic elements in Semantic Web, OWL is more expressive because of its ability to describe knowledge of domain and to represent machine readable content on the Web languages (Deliiska, 2007).

According to the Semantic Web ontology languages, ontologies do not include only terminological knowledge, definitions of the terms used to describe data, and the formal relations between these terms, but may also include the knowledge bases themselves, i.e. terms describing individuals and ground facts asserting the state of affairs between these individuals. This holds to be true even though such knowledge bases are often not regarded as being ontologies, (Obrst et al, 2007).

2.3 WHAT IS SOCIAL ENGINEERING

The most popular definition of the social engineering is the way to obtaining unauthorized access, sensitive information or secret data from any business location, or computer system by tricking the human element that has knowledge about the required target, commonly through cunning and charm. It represents a serious threat to even the most technologically secure networks, because it exploits the human element (Bahirwani et al, 2008).

In the field of information security, social engineering indicates to psychological manipulation of people to do actions or leaking sensitive information. A social engineer uses this type of trick for information gathering or to get access to a system, and has many complex methods used in psychological manipulation for people. Before the computer age, this term has been associated with social sciences, but it is now widely used in computer science and information security.

The cutting edge of information and communication technology (ICT) century under the rapid growth of ICT facilities have led to the heavy usage of electronic data in our daily life. The current phenomena of using internet services, networks, and other forms of innovative gadgets has stimulated the change of entire economic and lifestyle to a more technological and simplified way to access information. Tremendous numbers of profit and non-profit organizations have been creating incentives to encourage people to use the Internet for handling business transactions, placing orders, selling and event rewarding are all under one roof service. However, the risk of using the Internet for data transmission and individual users' awareness of information security are still large areas of concerns.

Nevertheless, the growth of information communication technology has also fostered the growth of information security technology. Despite this, Internet fraud cases are still mushrooming in line with the growth of technology. Regardless of whether they are private or governmental sectors, optimal concerns for protecting information are being expressed. For instance, the Malaysian government has enforced the rules of Personal Data Protection Act 2010 (PDPA) which is a set of regulations in governing the processing of personal data in regards to commercial transition. However, there are still many cases of organizations selling customers' personal data to other non-related organizations for commercial activities. Therefore, the primary organizations from which the personal data has emanated are making vast investments in protecting data with a comprehensive rigid system software and hardware to control the access of information. However, the biggest threat to organizations' information security is the social engineers' manipulation of employees who are vulnerable to their attacks. (Richard Brody, et. al, 2012).

Furthermore, Luo et al, (2011), highlighted a social engineer's ability to exploit the human organizational policies and cognitive biases that allow these engineers to get access to required resources.

In view of the discussion of social engineering claimed different view on this term, some cited "Art of getting people to respond with your wishes" (Sarah, 2001).

2.3.1 SOCIAL ENGINEERING MOTIVATIONS

Social engineering has many motivations and various reasons that compels social engineers to continuously develop victim controlling techniques for such motives. So they are not secret to anyone who follows social engineering motivations. Allen (2007) mentioned that at least one possible motivation is, for example (but not limited to):

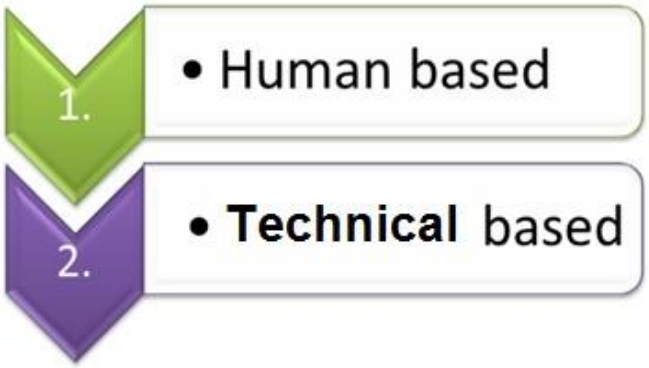
a) Financial gain: an example of this, a desire to get money, or transfixed to get more than he/ she earns by normal methods.
b) Self-interest: a need for access to certain information, and the desire to be adjusted for personal interest.
c) Revenge: for special reasons, individuals themselves had become detached enough to attack friends, relatives or even neighbors for obtaining certain information in order to take revenge.
d) External pressure: also in this aspect, for individual, may receive pressure from others as friends, organized crime syndicates, or perhaps one of the closest

people to him/ her such as family, for financial reasons or self-interest or even revenge.

2.3.2 SOCIAL ENGINEERING ATTACKS

Throughout the literature review research, there were lot of definitions and interpretations for social engineering and its complex attitudes towards obtaining sensitive information from targeted people. Of course it is a method to control someone by evoking his or her emotions. The usual route is that the striker composes a reasonable story, in order to gain the trust of his victim. Social engineers usually use deceptive methods or manipulation. At the same time there are people who reveal personal information science despite their awareness of the risks inherently involved. Actually, there are two kinds of social engineering attacks: human based attacks and technical based attacks.

FIGURE 2: Social Engineering Types.



2.3.2.1 HUMAN-BASED ATTACK

This type of attack requires interaction with humans. Also, this type can be called person-to-person contact in order to obtain desired information Many techniques can be used by the social engineers until their victims are controlled.

For Granger (2001), classified different ways of human-based techniques that allow the social engineer to gather information about victims, such as by phone, dumpster diving, on-line social engineering, persuasion, impersonation and conformity. All of these options had an effect on sources of desired sensitive information, whether they were sensitive information, harvest passwords, organization data security or individual settings.

In contrast to that, Gulati (2003), demonstrated that social engineering human-based attacks can be done even by a direct approach, such as spying and eavesdropping, technical expert person, support staff, or the voice of authority, in addition to dumpster diving.

On the other hand, researcher Guenther (2001) integrated ways of social engineers to gather information for the same type of human-based attack. Whereas he believed that, impersonation, pretending to be important user, dumpster diving, third-party authorization, in person, tech support, and shoulder surfing are all categorized under human-base attack. Because of that, all ways of information gathering could also affect and control behaviour of the victims to leak sensitive information needed by social engineer.

Unlike mentioned above, Heary (2009) turned in his topic to this type only. He indicated that there were five most popular social engineering exploit techniques, such as familiarity exploit, gathering and using information, creating a hostile situation, getting a job there and reading body language. Of course, there are many popular techniques still being used nowadays, but Heary's classification reflected the situation at that time.

Furthermore, Whitaker (2009) revealed the danger of using social engineering in reverse as one on the human-based attacks. His clarification consisted of three categories: sabotage, advertising, and assisting. In sabotage, usually social engineers aim to sabotage a network, send malicious software by email or give the impression that the network is sabotaged. Through the second method, the social engineer may advertise services as a security consultant whether by dropping a business card, or by sending email to create a problem in the network in order to place him-self in a position of help. In this way, the social engineer puts himself generally in the position of assisting in the same mode as a type of help desk in order to have first impressions to fix problems, then gains the chance to plant malicious software or steal sensitive information. Furthermore, Whitaker added a piggyback risk which had the effect of bypassing physical security control, tech support but in case of stealing passwords, personnel phishing who focus on bank accounts and neurolinguistic programming to get a company's secrets, sex sells and get smashed which lead to stealing trade secrets or getting access to a building (Whitaker, 2009).

For Allen (2007), in his research titled under "Social Engineering A Means to Violate a computer system" classified human-base attacks to nine sub-types such as

shoulder surfing that are used to steal access codes or get a password, checking the rubbish for data gathering about the victim, mail-outs which leads to individual/ organization information, forensic analysis, using direct approach, pretending to be an important user to allow remote access to a server, posing as a helpless user to get from assigned training a desired piece of sensitive information, technical support personnel and reverse social engineering to corrupt workstations or networks or stealing confidential information.

Twitchell (2009), introduced and defined threats which affect information systems. Also, he introduced a taxonomy which classifies social engineering attacks, depending on the dimensions of the targets, used media, attacks fit and methods/ techniques used to apply attacks. Twitchell classifications were, asking for favours, dumpster diving, cold calling, contriving situations, giving out free software, impersonation, photography, pharming, phishing, pre-texting, reverse social engineering, reconnaissance, shoulder surfing, simple requests, surveys, tailgating, theft and trojan horses. He discussed social engineering countermeasures in addition to mapped attack types to those countermeasures and turned to the trends and technologies that were supposed to aid in defending against social engineering attacks, respected that those trends should be mitigated and reduced the risks associated with social engineering attacks.

Since social engineering used malicious methods to gain access to required information by exploiting a bug in humans, Luo et.al, (2011) pointed to the most prominent attacks which are exposed to the human factor. Also, they dealt with the psychological aspects that help social engineers to influence the victims in an attempt

to penetrate them. They stated in their study five social engineering techniques, on the grounds that modern methods and recently innovative are: pretexting, online social engineering, phishing, and the oldest forms of social engineering attacks such as dumpster diving and shoulders surfing.

Nohlberg, (2008) also refers to some social engineering techniques and divides them into three basic categories, which are physical attack, social attack and a combination of technical and social attack methods. Nohlberg's categorization of physical attack contained dumpster diving, theft (of physical information), extortion and desktop hacking. Nohlberg added that the physical attack usually implemented was in the process of carrying out an advanced attack or to facilitate another attack, which could take the form of social attack by phone, pretext, reverse social engineering, and email. He classified phishing and websites as a technique that used a combination of both technical and social attack.

Stergiou, (2013) illustrated that the social engineering attack process starts with information gathering which usually depends on using social media and dumpster diving. The next phase consists of the development of a relationship between the attacker and the victim, depending on different techniques such as pretexting, micro expression, body language and neuro- linguistic programming. Stergiou classified attack methods into two categories, human-based and technical-based. In technical-based attacks, trojan horse, phishing, pop-up windows and reverse social engineering were employed. In the human-based attack, fear, diffusion of responsibilities, chance of ingratiation, guilt, and overloading were used. Stergiou mentioned about the principles of influences which are described within the two

contexts, human interaction and online interaction such as, reciprocation, commitment and consistency, social proof, liking, authority, and scarcity.

The classification of Gupta and Agrawal, (2012) included physical techniques and psychological techniques. The physical techniques contain dumpster diving and shoulder surfing. The other one which also refers to personal social engineering includes, by phone, online social engineering, persuasion, impersonation, and reverse social engineering. They respected that all techniques such as the voice of authority, IT support, network administrator and service provider, all these are categorized under impersonation and required pretending of social engineers.

Hasan, et.al (2010), focused on the techniques for persuasion and they organized these techniques in to impersonation, playing on users in order to pretend a certain role or work from inside / outside the company, fear or what is also called intimidation tactics, hoaxing, creating confusion, dumpster diving, reverse social engineering, mail that contains malicious codes, and phishing. They also indicated to other social engineering attacks that are used to spread spyware in order to help the social engineer in information gathering, these include options such as, mail, spam mail claiming anything to lure the target, piggyback software installation, and fake anti spyware.

Alan and Roderic (2006) did not much differ from others but they classified social engineering types in to Syntactic and Semantic instead of using human-based and technical-based. Syntactic refers to methods that depend on deploying viruses, malicious software and denial of service attack. These techniques are bogus surveys,

pop-up windows, spyware, phishing, vishing, and pharming. The other type is semantic which contains techniques that need direct interaction with the victim such as important user, helpless user, technical support personnel, reverse social engineer, by phone, eavesdropping, dumpster diving, and shoulder surfing.

Janczewski and Lingyan Fu, (2010) presented their study to demonstrate underlying constructs and present the major aspects of social engineering. They followed usual categories of social engineering classification. So they assigned the methods of human-based including impersonation, dumpster diving, shoulder surfing, reverse social engineering, and questionnaire. They found that the most widespread mode is, in person attack by phone. As long as technical-based attack is closer to techniques of traditional hacking, then the study classification of technical-based attack methods include: email attachments, popup windows, phishing, online social engineering, and rogue security software (Malicious). Off course attack such as social engineering represents a dangerous and wide range of threats. So it's important to have a multifaceted approach to defend against these attacks effectively.

As has been made clear, social network sites have become the fastest-growing online service. A lot of people prefer to share personal photos, files, and videos with family, friends or others through these social networks. Therein lies the danger of social engineering, where some of these social network features can be abused. Irani et.al (2010), discussed the danger of reverse social engineering as one of social engineering taxonomy, and elucidated how this type of abuse can happen online by exploiting friendship requests depending on displayed personal information.

The issue of risk mitigation of social engineering attacks has preoccupied many researchers. Spinapolice (2011), mentioned about these methods that must be considered, and focused on individual awareness. Matthew listed a taxonomy of social engineering as, dumpster diving, pharming, phishing, pretexting, impersonation (Quid Pro Quo), reverse social engineering, shoulder surfing, and trojan horse.

Kee (2008) discussed in his paper which was entitled "Social engineering, manipulating the source" accomplishment of social engineering. Kee mentioned some common techniques that lead to accomplish social engineering attacks such as by phone, online " refers to email, online chat session", dumpster diving, shoulder surfing, reverse social engineering and persuasion. Also he talked about who is affected by these attacks, and presented a chain of countermeasures that could prevent the occurrence of social engineering attacks.

Mandy, (2005) talked about information bandits and how they use a lot of techniques to obtain sensitive information. The paper discussed many social engineering techniques such as impersonation, dumpster diving, tailgating, name dropping, bribing, eavesdropping, phishing, stealing, shoulder surfing, and reverse social engineering with its approaches "sabotage, advertising, support". The paper concluded that it is difficult to find good real life examples of social engineering attacks.

Gragg, (2002) discussed social engineering basics, and elucidated the psychological triggers that respect good condition for social engineering success. Gragg focused on psychological triggers but also talked about some of the social

engineering techniques such as pretending to be an employee, reverse social engineering, exchange of favors, by phone, websites, and using voice of authority.

Lieu (2002) posited that humans are the weakest link referring to the ease of persuading and deceiving them. As he mentioned about social engineering techniques as dumpster diving, by phone, email, and massaging which also may called vishining.

Chan,(2006) saw that the techniques of manipulating individuals to give sensitive information depend on those who have a desire to be helpful, with a tendency to trust others, who fear getting in to trouble, and those with a willingness to cut corners tend to let general information into open view, which allow social engineers to easily exploit it. Olivia also classified social engineering types in to two categories, human based and technology based. Human-based contains impersonation, third party authorization, in person, dumpster diving, shoulder surfing, and finally, key-ghost which is a device used to capture everything typed on a keyboard, although it is possible to be classified as a technical-based. Technology-based contains popup windows, phishing, trojan, and websites.

Cheung, (2012) went on to talk about social engineering psychology, referring to the four human nature aspects that are suggested by Thomas Peltier. Actually, these aspects make the victim behave according to social engineering desires. Peltier classified these aspects as the tendency to be trusting, the desire to be helpful, the tendency to cut corners, and the fear of offending others. Alvin classified social engineering attacks into two categories, human-based and technology-based components. In human-based there are listed techniques such as impersonation, third

person authority, in person, dumpster diving, and shoulder surfing. Techniques listed under technology-based such as, software exploits, email phishing, and website phishing.

In a recent study, the researchers Algarni and Xe, (2013) observed that the social engineer needs to complete eight phases in order to accomplish successful social engineering attacks. One of these phases is how the social engineer could determine the suitable tactic and his/ her ability of developing an attack plan. In this phase, the writers mentioned that there are a lot of used techniques, but they listed as examples, phishing, persuasion, shoulder surfing, spam, dumpster diving and reverse social engineering attacks.

On another topic, Arief and Besnard (2003) talked about human and technical issues which are generally involved in the security of a computer-based system. In their research, they mentioned social engineering as one of the more prominent security issues. The methods reviewed of social engineering were mostly different from what was actually mentioned in other topics. These methods were impersonation, false authority, inconspicuous occupation, sympathy, reward, personal stake, boosting egos, and trojan horses.

Heikkinen (2006) highlighted social engineering in an emerging world of communication technologies and discussed the difference of social engineering aspects. Heikkinen posits that the new technologies provided a lot of ways of contacting which leads to affect people's decisions in their lives, and leads them to be exploited such as this contacting to execute social engineering attacks. Seppo

discussed some of the social engineering methods such as dumpster diving, shoulder surfing, surveys, persuasion, impersonation, tailgating (piggybacking), reverse social engineering, and referred to technical-based methods such as email, website, and phishing.

Buetler (2009) classified social engineering into certain terms included, by phone, dumpster diving, phishing, email, impersonation, and reverse social engineering.

Tovstukha and Laaneots, (2013) found that social engineers always use different techniques and may apply more than one attack to make people give or lead him to sensitive information. Furthermore, social engineers must apply psychological tricks on his victims, exploiting their human nature. Tovstukha and Laaneots discussed the common social engineering attack techniques such as pretexting, phishing, shoulder surfing, tailgating, dumpster diving, in addition to the Quid Pro Quo technique. They implied that some of the social engineering techniques need to be prepared in advance and some other techniques need multiple stages of execution.

Cazier and Botelho, (2007) researched the field of threats that still threaten human elements by social engineering attacks. As long as social engineers depend on exploiting the trust and fear of victims to obtain as much sensitive information as possible, Cazier and Botelho depended on an experimental survey to indicate these treats precisely. Techniques that were more threat and prevalent are, by phone, persuasion, dumpster diving, phishing, and internet websites.

2.3.2.2 TECHNICAL-BASED ATTACK

This type of attack is implemented by tricking the victims without having to interact directly with them. These attacks might involve similar principles that are found in face to face attacks. But technical-based differs from human-based in that the execution happens en masse, depending on the technical platform for implementation.

Gulati (2003), listed two common types of technical-based social engineering attacks, the first one is trojan horse which reflects same use as mentioned in Greek mythology. In this situation the attacker will send random emails to be appearing as harmless and containing attachment of worms or viruses. When an unsuspecting victim opens that innocuous attachment, the virus or a worm launches through the entire network to be infected. The most popular of these uses are "ANNA KOURNIKOVA" worm and "I LOVE YOU" virus. Usually this type is used to bring the network down and to put an exponential load on network resources. The second type is popup window, where the social engineer's rogue program will generate a popup window, to falsely state that there are network problems and due to that the application connectivity was dropped. Now if the user re-enters their ID and password in order to continue the session, then the attack will take place without realizing that, that he was the one who opened that attack gate.

For Guenther (2001), classified technical-base attacks to four subtypes such as, popup windows, mail attachments, spam, chain letters and hoaxes while they do not commonly cause harm, but they do cause a loss of productivity, where they use valuable network recourses, and websites which mean that the attacker will use a

common ploy to win a sweepstakes on a website or it could be to offer something free. But to win, it requires entering an email address and password.

According to Allen (2007), listed three subtypes of technical-base social engineering attacks such as, email by using a topical subject in order to trigger an emotion. In this attack, a malicious code or virus will be sent to clog mail systems. A second technique is using a website to promote a fictitious competition, which requires entering an email address and a password. Phishing is used to entice recipients to visit counterfeit websites by using specially crafted emails.

Whitaker (2009) mentioned about two types of social engineering technical-base attacks. The first is social (engineer) networking, while the most popular social networking sites are considered to be a social engineer's paradise. People use these sites and post their information about work, bands they like, friends and more. A social engineer can detect and find out about victims from these sites, and can impersonate a friend, find out popular hang-outs, discover personal information or be added as a friend to build a relationship in order to build trust. Then that trust will be exploited later to get information from the same person or it could be used to launch another attack. The second is "Catch Me a Vish" or SMS cell phone, vishing refers to an attack that uses the phone to perform deception which is the equivalent of a phishing attack. For example, you may receive a strange call tells you that your credit card has been compromised, and you should contact a specific phone number, after responding, the social engineer will ask you to enter the card number, PIN, address, and all that he needs to collect information about the victim. Another way of using a vishing attack is to send a message to a cell phone instead of direct calling.

Lvaturi and Janczewski, (2012) sought to focus on online social engineering which represents technical-based attacks. Although their study indicated three classifications based on broad criteria, the one most discussed is online social engineering. The taxonomy contained phishing, money laundering, and then malware which includes three methods such as malicious downloads malware through pop-ups and search engine poisoning. The fourth technique is clickjacking which refers to counterfeit links (malicious) where the victim use them based on wrong information given by social engineer. Another form of attack is malvertising, which means the using of malicious agents to attract users to malicious websites.

Maan and Sharma, (2012) contends that the basic goal of implementing social engineering attacks is the human element, although attack processes are needed to use technology in some cases to tightly control the victim. In their study, they indicated piggybacking, tailgating and telephonic cheating under the human-based attacks category. On the other side, they classified phishing, fake mail and pop-up windows attacks under technical-based social engineering attacks.

Most organizations suffer the harm of their careless individuals, Greitzer et al, (2014) reports that unintentional insiders' threats directly serve social engineers if the organization employees do not have sufficient awareness. In their taxonomy classification, they depended on the description of a social engineering attack whether it needs interpersonal interaction or not. One kind of attack that has no need for interpersonal interaction is intelligence gathering, which depends on open source research such as Facebook, organization sites, and dumpster diving. The other kind of attack that needs interpersonal interaction is divided into two types which are non-

electronic means and electronic means. Shoulder surfing, impersonation and reverse social engineering are listed under non-electronic means. Trojan horse, website, pretexting and phishing are listed under the electronic means sub-category.

2.4 RELATED STUDIES ON ONTOLOGY

There are many previous studies that have been conducted on the issues of ontology, and the factors that explain the existence of things, or to achieve satisfaction that reaches to truth, and describes relationship between these topics and the other sciences.

Information is scattered everywhere and it is not easy to search and find relevant information, therefore Raskin, et al (2010), conducted a study about computational systems to detect intentional reasons in casual, unwanted individuals who are responsible to lack sensitive information for unauthorized people. They explained protection methods of secret and sensitive information from insider threats affected by social engineering and tried to build obstacles against unauthorized access people. According to their research, they highlighted the use of ontological semantic technology to make this process as a computational system. Furthermore, the system discussed in this study, increased the level of awareness for users and extracted the information that is given away unintentionally, (Victor Raskin, et.al, 2010).

In contrast, For Feresa, et.al, (2011), illustrated a study about a generic taxonomy of social engineering attack, which aims to be used as a guideline for any proposed generic taxonomy of social engineering. As explained by their study, detecting social engineering is not easy without knowing how to manage secure sources and sensitive information. A weakness in this manner will unquestionably lead to a negative impact, especially if the lost data is associated with financial data. That's why this study reviewed and analyzed on both categories, management and technical aspects of social engineering attacks to produce a generic taxonomy that is composed of human-based and technical-based social engineering attacks. They classified few types of attacks listed under these categories, compared with the types that have been classified in the previous studies. Of course their study discussed different ideas on how to identify types of social engineering attacks, although these types are still unclear and are growing day by day, (Feresa et al, 2011).

Due to what has been done by others, and to avoid duplication in the development process, Bergey et al (2014) focused in their study on how to categorize social engineering tactics rather than focusing on ones that have been recently developed.. So their study achieved characterization of the various efforts that have already been implemented in the field of social engineering, employing tactics that have been adopted since 2001 by three of the particular authors in the study.

After progress in previous taxonomies review, they decided to develop one aim to achieve a conformal with a strict class hierarchy, clarify the nomenclature, comprehensive and neutralize to vulnerabilities and mitigation techniques. To achieve that, they analyzed the cases to determine if documentation were related to the

findings. Further, they examined the cases for any techniques that relate to contributing, targets, methods, factors, and attacks progression. According to their study findings, at best there is a weak correlation between social engineering receptivity and various demographic factors. A weak security policy in addition to defects in some organizational factors can generate a gap for system vulnerabilities and can be exploited by adversaries to implement attacks. There are many reams of social engineering academic research which have specified the possibility of human elements involved in (unintentional insider threat) to social engineering susceptibility. That's the reason for adaptation of conceptual modeling and analysis of collected research and case studies disclose a series number of common factors that may inform the improvement of strategy or vulnerability mitigation tools (Bergey et al, 2014).

Moreover, the study of Ivaturi and Janczewski, (2011) aimed to classify social engineering methods through taxonomy. This was done in order to lead to a better understanding of social engineering attack methods. Furthermore, as a good taxonomy is one that is comprehensive and comprehensible, they tried to make their taxonomical structure in these ways. In their study, they divided social engineering attack methods into two main categories, person-to-person attack methods and person-to-person via media attack methods. Then they listed many types of methods under these categories which formed a good organization for a social engineering taxonomy and to understand the attacks vectors better. Real person impersonation and fake person impersonation are listed under person-to-person attack methods, and contain several types such as, pretexting, reverse social or quid pro quo and tailgating. Text, voice and video are listed under person-to-person via media attack methods, and contain many

types described in the state of methods. Their classification is more clear, but it can be improved upon. (Ivaturi and Janczewski, 2011).

For the purpose of ensuring a certain level of understanding for the meaning of the data and to achieve a good knowledge of sharing and reuse**,** for Fuertesa et al, (2008) observed that the ontology development depends on the previous domain's definition, according to included terms and their classification system. They aimed by their study to reduce interoperability and the problems that occur with information exchange, constructing a hierarchal structure of all areas and an interrelated system to connect them, (Fuertesa, et al, 2008).

For Fonseca, (2007) made clear that the first essential purpose of much research on the ontology in modeling usage is required to understand what a system of information is, he has showed in his study that classification of all domain' terms and their uses that are covered in literature reviews will give a comprehensive understanding for ontologies. However, his study provided a better understanding of what ontologies are in addition to explaining a dual role of ontologies. Fonseca tried to offer a good level of knowledge sharing and reuse by reusing and updating existing ontology in a good manner, and reviewed an example of both Milton & Kazmierczak, (2004) which presented a model of using information systems ontology. As an external reference to their philosophical ontology usage, they made a differentiation of some modeling languages. Furthermore, for better understanding of the basics nature of languages of data modeling (conceptual modeling grammar), they reused Chisholm's ontology and analyzed the grammar by their own procedures.

On the other hand, the study of Fonseca, (2007) focused on researches through a distinction between those concerned with ontology construction and those concerned with the use of ontologies. So he has shown that the development of information systems and their usage is strongly connected and related to the use of the term. Illustrating the importance of enhancing ontology terms and to be more understandable, he divided research areas into two parts: the first part indicated ontology research, a powerful tool to guarantee that (conceptual modeling grammar) of data modeling languages are correct. The second part, the availability of ontologies research for information systems, are useful to guarantee that the (conceptual modeling scripts) of conceptual schemas are correct, (Fonseca, 2007).

As long as ontologies provide a good way for knowledge sharing, reusing, and updating existing data, it can also provide the structured vocabulary and semantics which can be used in the markup of web resources to provide machine understanding. As a web holds a vast amount of information, Gaihua, et.al, (2005), believe that ontologies have a major role in semantic web research. To enhance current web documents with Meta data, they developed an ontology of spatial query expansion in an information retrieving area. They thought that a project such as this will definitely support document retrieval that suggested being spatially pertinent to queries of different users.

Generic spatial queries applied in consideration in this project as one of the important factors. For this reason they used techniques that support spatial query expansion by generic factors consideration. However, iterative spatial query expansion is supported by these techniques in case of low search results, if a query

footprint produced that. According to features that would be achieved by iterative spatial query expansion, so the proposed research techniques will give additional support. The reason for this vital matter, as follows, is the existence of inappropriate documents on the web to characterize them. Some of the ontology information is in encryption case, then that would not be as valid as they assumed to be, mostly if the processes to obtaining these values are more expensive. Derivation of query footprint will take some generic factors in to account. If the initial results of search do not satisfy user needs, then it is eligible for spatial query expansion to be performed, (Gaihua et al, 2005).

2.5 ONTOLOGY TOOLS

There are many tools used for constructing ontologies, some of them are free, and others need a license.

a) Top Braid Composer (http://topquadrant.com/products/TB_Composer.html**)**

This tool is classified as a powerful modeling environment and one of the best tools of an IDE for building semantic applications. Compatible with W3c standards, able to offer support for developing, managing in addition to testing configurations of linked data and ontologies.

b) Pool Party (http://poolparty.biz/**)**

Is a tool of thesaurus management, which is designed to allow creation and enrich thesauri and concepts by supporting the analysis of documents and gathering more information from different data sources. Because the domain experts are almost never semantic web experts, this tool focuses on holding access to low contribution by supported with wiki style interfaces.

c) Text Editor (http://gnu.org/software/emacs/**)**

GNU Emacs is customizable text editor, is extensible, and offers more features. Supported to be an interpreter for Emacs Lisp, which is the Lisp programming language to support text editing with its extensions. GNU Emacs include the following features:

1) Content-sensitive editing modes, syntax coloring, support file types of, HTML, source code, and plain text.
2) Fully built-in documentation, in addition to a new user tutorials.
3) Complete Unicode support for all human languages and their scripts.
4) Support using of a graphical interface or Emacs Lisp code.
5) A lot of extensions that append Varity functionality, including a calendar, a project planner, debugger interface, mail and news reader and a lot of distributed with GNU Emacs extensions.

d) Hozo - Ontology Editor (http://ei.sanken.osaka-.ac.jp/hozo/eng/index_en.php**)**

Is a graphical editor for ontology construction especially created to produce well thought out and heavy weight ontologies. It is one of the editors that participates in the creation of collaborative knowledge.

e) ROO: Rabbit to OWL Ontology Authoring Tool (http://sourceforge.net/projects/confluence/)

ROO is an ontology creation tool based on Protégé 4 that assists domain experts to build conceptual ontologies. ROO uses Rabbit to enable domain experts to automatically formalize their knowledge in OWL. ROO provides easy-to- understand suggestions and task-specific messages to help the user enter correct-controlled natural language (CNL) constructs. Appropriate feedback is given to help users recognize concepts, relationships and individuals when writing CNL sentences. Syntax highlighting based on the parsed structure helps the user recognize CNL patterns. Although ROO is based on technologies such as OWL and natural language processing, ROO avoids exposing technical terminology to ontology authors. ROO prefers to use conceptual terminology that may not be well-defined in a technical sense, but which is easier to understand for novice users. In the case of OWL, ROO will avoid introducing terminology such as Object Property, opting to use relationship instead. In the case of the natural language processing, ROO avoids using linguistic terminology such as determiners or adjectives as much as possible. When technical terminology is introduced, ROO tries to give specific examples, preferably coming from the domain.

2.6 CONCLUSION

Many experiences were gained during the review of the previous studies. Studies also show that social engineering is a real challenge which harms and leads to great damage for individuals and organizations. There are many motivations for Social Engineers; the most notable is the financial gain. Social Engineers select victims by deceptive methods or manipulation. Most of the previous studies classified Social Engineering attacks into two types, human-based attacks and technical-based attacks. This classification makes it easier to put in place countermeasures and mitigate risks. The advantage of reviewing previous studies make our task of avoiding redundancy easier, as well as processing and enhancing the developing process by knowing what has been done by others.