The Terminator's origins or how the Semantic Web could endanger Humanity*

Miel Vander Sande, Sam Coppens, Davy Van Deursen,
Erik Mannens, and Rik Van de Walle

Ghent University - IBBT
Department of Electronics and Information Systems - Multimedia Lab
Gaston Crommenlaan 8 bus 201, B-9050 Ledeberg-Ghent, Belgium
firstname.lastname@ugent.be

Abstract. The next ten years will contain some major technological shifts in the Semantic Web. Agents will become highly intelligent, data will be fully interlinked, content will be highly personalized and the Web of Things will be fully deployed. Although these shifts enable new powerful applications, the aftermath will have a negative side. Similar to former evolutions, computer criminals will find new ways to benefit from these shifts.
In this paper, we give a pessimistic view on the future Web by giving an overview of possible misuses. We discuss adopted or new ways humans can misuse the Semantic Web for their benefit. Next, we describe how evolved Artificial Intelligence could perform automated misuse and how this could lead to the extinction of humanity. As an example, we look for similarities in the scenario of the *Terminator* franchise.

1 Introduction

In the last ten years, researchers have been designing and constructing the Semantic Web. A lot has already been achieved: standards were developed, frameworks were constructed, ontologies were modelled. Furthermore, as a result of many efforts of the Linked Open Data community, many linked and structured data was published. With these initiatives, the Semantic Web will one day be a worldwide intelligent data source. This will lead to more relevant search results, simpler knowledge extraction, better integration and easier access of data sources. However, this will not all be beneficial. In former technology shifts of the Web (e.g., embedding of multimedia, user generated content), computer criminals have found new ways to misuse this technology to their advantage. In most cases, the techniques for illegal purposes have exposed features that have never been discovered by anyone else. Thus, while creating the Semantic Web, many zero-day vulnerabilities are yet to be exploited.

* The research activities as described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT), the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders (FWO-Flanders), and the European Union.

The popular *Terminator* franchise, a series of motion pictures initiated by director James Cameron, describes an artificial intelligence system called *Skynet*. After this global digital defence network became 'self-aware', it started the total extermination of humanity (e.g., the so called *Judgement Day*). Although this is a purely fictional scenario, could the future Semantic Web show indication of creating a similar disastrous outcome?
In this paper, we discuss a pessimistic view on the evolution of the Semantic Web in the coming 10 years. First, we discuss the possible misuses by humans to commit the next generation computer crimes. Next, we describe how these can lead to misuse by Artificial Intelligence in Sect. 3. Finally, we end with a conclusion in Sect. 4.

2 Misuse of the future Web by Humans

In this section we discuss how existing criminal methods, created by humans to benefit themselves, could adopt to the changes on the Web and which new additional methods could come into existence.

2.1 The Spam takeover

According to I. Davis, a system can be called popular from the moment it attracts Spam attacks[3]. Since Spam did not reach today's Semantic Web, it can not be called mainstream yet. However, Davis already identified seven possible Spam vectors in Linked Data. With future emerging of the Semantic Web and increasing popularity, future technology shifts will create even more new challenges.
A first shift is the continuing expansion of social networks. Current websites already indicate a rapid increase in social media integration in other website's authentication. What is currently an account on a specific website (e.g., Facebook, Twitter) turns into a complete identity, which includes extensive personal information. Furthermore, the social networks will be fully linked with the Semantic Web, enabling any piece of data to be connected to anybody's identity. This enables Web applications to deliver completely personalized content. However, this advantage is also a disadvantage. Spam messages become much more convincing by including deep personal information of the targeted user.
A second shift is the further Linked Open Data cloud expansion and the improvement of agents. These agents will have access to any data on the Web and will have perfect understanding of their semantics. Therefore, they can be used to assemble Spam messages that are harder to detect. Intelligent reasoning combining Linked Open Data and personalisation could obtain full context awareness. Also, future Web application's will mostly be fully automated aggregators of data, based on generic reasoners. Hence, by simply adding a few logic rules, their central technology can be reused to generate tons of Spam with little cost.
A third and final shift is the Read/Write Web. Today, the Semantic Web is mainly read-only. Luckily W3C's Linked Data Platform Working Group [2] predicts a future read-write architecture. Not only will clients

be able to write to the Web, machines will autonomous manage data and perform updates. However, this will also provide new ways for Spam to reproduce and reach its targets. Firstly, Spam messages can easily be added to existing resources (e.g., adding an *rdf:label* to an exiting resource). Secondly, agents will spread these false triples when harvesting them through their automatic update process.

2.2 Link spoofing on the rise

In the next 10 years, the Linked Open Data cloud is likely to grow rapidly, forming one huge data source. Over the last two years alone, the cloud has tripled in size. Also, with interlinking, data cleansing and reconciliation being hot research topics, the quality of the links will become a lot higher. Thus, clients become what they were intended to be: smart consumers of Linked Data. Many applications will depend on these links and request them blindly. This opens many new domains for link spoofing. In wireless communications, B. Kannhavong et al. describe routing attacks used today, based on link spoofing[6]. Mobile ad hoc networks form paths between source and target using devices in the network as routing nodes. A malicious node can advertise false links in order to hack the network or disturb its traffic. In Linked Data, similar approaches can be used. Injecting malicious links into data can trick automated processes in clients. Possible ways for doing this are adding a property in RDF, using *owl:sameAs* to link to malicious resources, changing an object of a triple or altering the redirection process of dereferenceable URIs. For example, one could replace an URI in the object of *foaf:img*, which the client uses for presenting an image to the user. When the image is requested, a hazardous Worm is downloaded instead.

2.3 The next generation identity theft

As mentioned above, social networks will be a fundamental part of the Semantic Web, putting a user's Web-Id (Web-Identity) into the centre. However, as these identities become more and more integrated, their vulnerability for illegal purposes increases. This deep integration will have three main uses: centralized authorisation, interconnection of devices and automatic provenance generation.

The first use is making a user's Web-Id the main authorisation system on the Web. Even today, Facebook and Twitter are already being implemented as a single sign-on service in many websites. It is clear that this bigger exposure makes account hacking a bigger threat.

The second use is, related to the previous one, the interconnectedness of all personal applications and devices. Everything that a person uses will be linked to his or hers Web-Id, providing optimal interoperability. As part of the Web of Things (which will be discussed later), any device can access another's data or send it instructions. (e.g., Apple already uses its iCloud services to link all of your Apple devices together). When hacked into a Web-Id, one has access to all of the linked applications, data and devices. Also, all currently existing documents will be digital

and, if necessary, signed with digital signatures based on the Web-Id. As these signatures will evolve, so will the forgeries. The first issues already occur today, as hacked cloud services (e.g., Apple, Amazon) gave remote access to all subscribed devices[5].
The third use is the automatic generation of provenance. Since the future Web will have a full Read/Write architecture, manageability of data becomes the most important issue. Users will need to know what data to trust. Therefore, the reliance on *proof* and *trust* systems will grow, which depend on widely implemented provenance generation to evaluate data. The Web-Id plays a key role in automatically generating provenance, making every move made traceable. Unfortunately, this comes with a bigger dependency on provenance quality. With provenance manipulation, *proof* can be falsely obtained, deceiving *trust* systems to allow malicious data into clients.

2.4 Malicious rule injection

Reasoning is one of the key features of the Semantic Web. Most of today's systems use first-order logic or probabilistic logic. As the research progresses, these agents will become extremely fast and capable of processing more complex logic[7]. Depending on rules or ontologies, the functionality of a generic Web agent can be customized by injecting a set of rules. This can fit multiple criminal purposes. First, false information can be spread. Reasoning based on false rules can extract incorrect knowledge from the data and present it to the client.
Second, Botnets (a collection of hacked internet-connected computers controlled by a malicious party) are easier to create. Rules can connect machines by defining links between them. Furthermore, they can add semantics that trigger certain tasks on the client. By combining the former and the latter, the functionality can be steered towards tunnelling requests, making it harder to trace hackers. Also, by adding enough logic, a client can be programmed to automatically generate requests to perform, for instance, DDOS (Distributed Denial Of Service) attacks. For example, the value of *rdfs:comment* is used to represent an object on a Web page, which is generated by server-side scripting. By replacing the value of that *rdfs:comment* property with a malicious script, user agents reading this Web page could be opening a Web socket without noticing it.

2.5 Introduction of Data viruses

Current computer viruses infect other programs by adding code to their executables. As discussed in the previous section, reasoners will handle all the processing on the future Web, with rule injection as potential treat. This evolution will also introduce data viruses. Using malicious reasoning, ontologies and triples, data can be modelled to reproduce itself or to force others to link to it. With the Semantic Web having reached huge coverage, tons of networks can be infected instantly. Examples of such viruses are Spam message generators, Flooders (e.g., producing triples until the server crashes) and Provenance Spoofers (e.g., adding false Provenance to gain or loose Trust).

3 Misuse of the future Web by Machines

In this section, we discuss new malignant methods introduced and performed by intelligent machines. We start by describing which milestones in the development of the Semantic Web result in fully automated clients driven by Artificial Intelligence. Next, we discuss possible forms of misuse by these automated clients and how this could have a *Terminator*-like scenario as possible outcome.

Acceptance of machine intelligence Many efforts in creating Artificial Intelligence are currently hold back. This is caused by the estranged feeling people get when observing computed logic and the answers it comes up with. Mistakes made by machines are seldom forgiven. In the next 10 years, developers and end-users will have a change in mindset about what computer intelligence should be. With IBM Watson as pioneer, next generation systems will accept computer intelligence as being different from human intelligence. This causes development to stop mimicking human logic and starting to embrace the way computers solve problems, resulting into far better results very fast. For instance, self-driving cars (as already being tested by Google) will cause accidents in a way no human would, but in total, the number of car crashes will be less.

Complete implementation of the Web of Things In the Web of Things, everyday devices and objects all become read(-write) information resources on the Web. This way, connected Things will become accessible for intelligent machines. Information coming from these 'things' has been identified as a major driver in many domains, from smart cities to environmental monitoring to real-time tracking. For instance, it is not unlikely that military equipment will be part of the Web of Things in the future. Autonomous computerized weaponry already exists (e.g., drones, Stuxnet)[4], and semantic technologies will be key to manage and control them.

Along the lines of *Judgement Day* As discussed in the previous section, the Semantic Web is working towards a complex distributed Artificial Intelligence network. This network performs data-driven tasks, fed by one huge Linked Open Data Cloud. This will be a result of years of (i) developing reasoners able to handle high levels of logic; and (ii) modelling knowledge and logic as ontologies and rules.
Although a *Judgement Day*-like scenario is obviously too far fetched, it does point out that building an intelligent Web requires caution. If agents would become advanced enough, after consuming a sufficient amount of description logic, they will be able to make autonomous decisions. Because of their ability to perform complex tasks and the acceptance of their intelligence (as described above), the main use of future agents will be managing digital devices in the Web of Things (especially tasks where human failability needs to be avoided). With the amounts of malicious

rules and links put into the Web (as discussed in the previous section), machines could take wrong decisions and start to misuse the Semantic Web in their own way. Also, when having control over dangerous devices, such an event could even cause harm to humans.

4 Discussion

In this paper we examined possible misuses of the Semantic Web in the next ten years. First, we discussed how computer criminals will adopt their methods to misuse new technological shifts. Next, we described how the evolution of intelligent agents could lead to misuse by Artificial Intelligence. Finally we explained how this could have disastrous results, with human extinction as possible outcome.
It is clear that internet security will have a new set of challenges when the Semantic Web goes mainstream. Therefore, it is crucial to proactively perform brainstorming about new possible crimes on the Web. Security techniques cannot lag behind and needs to adopt fast to technological changes.
The evolution of reasoning processes also requires some thought. Although a *Terminator* scenario seems very unlikely, adding intelligence to the Web opens powerful doors to people with bad intentions. Improved reasoning approaches need to be evaluated against misuse or malfunctions. Also, possible fail-safe approaches need to be investigated.

References

1. Spam Statistics and Facts. `http://www.spamlaws.com/spam-stats.html`, August 2012.
2. A. Bertails, S. Hawke, and I. Herman. Linked Data Platform (LDP) Working Group Charter. `http://www.w3.org/2012/ldp/charter.html`, May 2012.
3. Ian Davis. Linked Data Spam Vectors. `http://blog.iandavis.com/2009/09/21/linked-data-spam-vectors/l`, September 2009.
4. Peter Finn. How Apple and Amazon Security Flaws Led to My Epic Hacking. `http://www.washingtonpost.com/national/national-security/a-future-for-drones-automated-killing/2011/09/15/gIQAVy9mgK_story.htm`, September 2011.
5. Mat Honan. How Apple and Amazon Security Flaws Led to My Epic Hacking. `http://www.wired.com/gadgetlab/2012/08/apple-amazon-mat-honan-hacking/all/`, August 2012.
6. B. Kannhavong, H. Nakayama, Y. Nemoto, N. Kato, and A. Jamalipour. A survey of routing attacks in mobile ad hoc networks. *Wireless Communications, IEEE*, 14(5):85 –91, october 2007.
7. D. Lopez-de Ipina, A. Almeida, U. Aguilera, I. Larizgoitia, X. Laiseca, P. Orduna, A. Barbier, and J.I. Vazquez. Dynamic discovery and semantic reasoning for next generation intelligent environments. In *Intelligent Environments, 2008 IET 4th International Conference on*, pages 1 –10, july 2008.