International Journal of Computer Science and Applications,

Vol. 11, No. 3, pp. 21 – 44, 2014

LINKED DATA CONNECTIONS WITH EMERGING INFORMATION TECHNOLOGIES: A SURVEY

HELE-MAI HAAV

Institute of Cybernetics, Tallinn University of Technology, Akadeemia tee 21/1
12618 Tallinn, Estonia
helemai@cs.ioc.ee
http://cs.ioc.ee/~helemai/

In the nearest future, semantic technologies will significantly improve machine understanding and interoperability of data. In this paper, we review and evaluate four distinct areas of emerging Information Technologies: Linked Data, Big Data, Internet of Things and social media. For each of these technology areas we present concept, outline state-of-the-art and development trends, and show impact based on application scenarios. Furthermore, we show how these technologies are related and how Linked Data technology plays a role of integrator of other analysed technologies.

Keywords: Linked Data; Big Data; Internet of Things; Social media; Survey.

1. Introduction

Nowadays, humans and machines connected to the Internet produce a huge amount of data that need to be collected, analysed and meaningfully used. Humans create voluminous data using different social media technologies. Machines (e.g. sensors and devices) create data that is transmitted to computers via the Internet.

Various new technologies emerged to cope with ever increasing amount of data of different types and originating from disparate sources. For example, Big Data technologies contribute to solving problems related to volume, variety, variability, velocity, and veracity of data. Big Data analytics tries to find hidden patterns in data for decision making purposes. However, there is need for technologies that could solve semantic interoperability problems of data produced by humans and machines. In general, these problems could be solved by semantic technologies like Linked Data technologies and ontologies.

In the nearest future, semantic technologies will significantly improve machine understanding and interoperability of data. In this paper, we review and evaluate developments of emerging Information Technologies for capturing data from humans and machines like social media and Internet of Things (IoT). We also discuss Big Data technologies for management and analysis of these data and Linked Data technologies for semantic interoperability and linking data and technologies.

Results of large number of studies on research and technology trends show that due to ultrafast Internet connection there is emerging a future web that will be a web of linked data and devices enabling real-time aggregation and analysis of extra large amounts of data [Gartner (2012a, 2012b), PwC (2009), PwC (2011, 2013)]. For example, Cisco looks to these trends from the point of view of networks and calls this phenomenon Internet of Everything (IoE) as more and more people, places, and things are connected over IP networks [Cisco (2013)]. In order to make this new web a reality development and deployment of all above mentioned technology fields are important.

The main research question of this paper is as follows: What will be a role of Linked Data technology in a new emerging web - web of everything?

In order to answer to this question, we first briefly review and evaluate emerging technology areas needed to make this new web a reality. These are Linked Data, Big Data, IoT and social media. For each of them, we define concept, outline state-of-the-art and development trends, and show impact based on application scenarios. Furthermore, we show how these technologies are related with each other and a possible role of Linked Data in connection with other technologies.

Main contribution of this paper is an analysis of possible ways of deployment of Linked Data in the fields of Big Data, IoT and social media. In addition, central role of Linked Data in interlinking of these technologies is shown.

The rest of the paper is structured as follows. Linked Data technologies are evaluated in section 2. Sections 3, 4 and 5 are devoted to review of Big Data, IoT and social media developments. In section 6, technology connections and a role of Linked Data technologies are analysed. Conclusions are drawn and represented in Section 7.

2. Linked Data Technologies

2.1. *Concept*

Linked Data technology is one of the semantic technologies. Linked Data are distributed, linked and semantically interoperable data sets that are represented in uniform structured format (e.g. Resource Description Framework (RDF)), published on the web and accessible via query endpoints (e.g. SPARQL endpoints). Linked Data technology is built on existing standard web technologies. It is based on minimal required interoperability consensus for data representation on the web by using URIs[a] and RDF[b] and enabling access to data via HTTP. URIs are used for sharing information between computers for supporting machine readability as opposed to serving web pages to people. This in turn enables to link data from disparate sources and query this data [Bizer *et al*. (2009)].

One of the most famous Linked Data projects is DBpedia project[c] that publishes information extracted from Wikipedia[d] on the web in RDF format.

T. Berners-Lee introduced four principles of Linked Data in 2006[e] . These principles do not relate Linked Data to open data. The relationship between Linked Data and open data was introduced by Linked Open Data project of W3C and it was later amplified by

[a] http://www.w3.org/TR/uri-clarification
[b] http://www.w3.org/RDF
[c] http://dbpedia.org
[d] http://www.wikipedia.org
[e] http://www.w3.org/DesignIssues/LinkedData.html

EU project LOD2[f] . Thus, Linked Data could be open (e.g. government data) or closed (e.g. enterprise data).

Most important advantages Linked Data technology can provide are possibility to aggregate data that reside in different places without data warehousing and more straight and capable analysis of these data comparing to traditional methods.

2.2. *State-of-the art*

Basis for Linked Data technology was established by W3C by publishing RDF recommendation in 2004 and RDF query language SPARQL as W3C standard in 2008[g]. In 2006, T. Berners-Lee published linked data principles for open data as *5 Stars of Linked Open Data*[h] . Linked Data development was tightly related to open government data projects of USA and UK in 2009. Corresponding EU FP7 projects LOD2 and LATC[i] got funding in 2010.

State-of-the-art of linked data can be characterized by activities of W3C in 2013. W3C put a lot of effort to development of Linked Data technologies by publishing best practices and guidelines of linked data platforms [W3C (2013a)] and recommendation of SPARQL 1.1 [W3C (2013b)].

In addition, W3C has succeeded in standardization of some important vocabularies (ontologies) like people[j] and organization ontologies[k] . People ontology uses Friend-of-a-Friend (FOAF)[l] ontology for describing the concept of a *person* as an entity of type *foaf:Person*. People ontology includes more terms than FOAF vocabulary for describing people (e.g. their addresses, their connections to organisations, etc.).

We may conclude that technological basis for Linked Data technology is established. However, there is lack of tools that support all stages of Linked Data life cycle preventing rapid deployment of Linked Data. Another issue is unsatisfactory performance of these tools in the case of large scale data pools.

According to LOD2 project findings, only 5 % of web of data are linked meaning that a small fraction of linked data sets are connected via links. One of the reasons for that is lack of commonly accepted and open taxonomies (or ontologies).

W3C standards like people or organization ontologies can be used but their usage is limited to certain data sets and applications. The situation is better within concrete information systems because ontologies needed for data curation and linking are developed together with the system. Therefore, ontology engineering [Gómez-Pérez *et al*. (2004)] as a field of semantic technologies is very important in order to support Linked Data deployment and application development

[f] http://lod2.eu
[g] http://www.w3.org/TR/rdf-sparql-query
[h] http://5stardata.info
[i] http://latc-project.eu
[j] http://www.w3.org/TR/vocab-people
[k] http://www.w3.org/TR/vocab-org
[l] http://xmlns.com/foaf/0.1

On the other hand, it is hard work to develop standardized common ontologies as even creation of commonly acceptable domain ontologies has been very painful. Reasons are complexity of the task, its knowledge and time consumption as well as required commitment to created ontologies. One of the solutions will be machine learning of ontologies [Maedche and Staab (2001), Haav (2006)].

2.3. *Future trends*

Future developments of Linked Data technologies are related to semantic technology take-up and explosion of Big Data and open data. In addition, IoT and social media analytics can have influence to Linked Data deployment trends.

According Gartneri study *Hype Cycle for Big Data* [Gartner (2012a)] semantic technologies are to reach the plateau of "*Peak of Inflated Expectations*" in more than 10 years and social media analytics in 2-5 years. At the same time, some related technologies are close to the plateau of "*Technology Triggers*" as follows: content analytics (2-5 years), IoT (more than 10 years), and social network analysis (5-10 years). These technologies could make use of Linked Data approach.

Gartner top prediction for 2014 [Gartner (2013c)] brings to the front IoT as a bridge between machines and humans and neurocomputing (see also Section 4). Neurocomputing might become one of the supplements or replacements of Linked Data technologies. Gartner predicts that in 2017 due to deployment of neurocomputing 10% of computers rather learn than process information.

2.4. *Deployment*

Linked Data by nature are open data and therefore open governmental data provides good test bed for Linked Data technologies. This is one of the motivating factors of linked open data initiatives. Successful deployment of Linked Data is related to so-called start-up problem (i.e. exponential growth of Linked Data requires existence of a critical amount of Linked Data). Linked open data initiatives help to create this initial amount of data.

Linked Data technology is mainly used for data integration as consistent vocabulary (ontology) based linking of data reduces data integration costs and creates opportunities for application development. Nowadays, Linked Data technology is used in private and public sector.

In private sector, companies use this technology for linking their closed linked data to (linked) open data or results of mining open data. For example, Fujitsu Europe combines open and closed data for creation of a healthcare application based on usage of sensors [W3C (2013c)]. Linked Data technology is well used for development of semantic content management systems and media applications. For example, BBC is one of the first companies who started to use Linked Data technology with building data service for Olympic Games 2012[m].

[m] http://www.bbc.co.uk/blogs/internet/posts/olympic_data_xml_latency

Garlik from financial sector uses extensively closed and especially secured linked data. Sindicetech helps companies to create closed linked data clouds. They have clients like publishing company Elsivier, pharmaceutical company AstraZeneca[n] etc.

Though many private sector enterprises all over the world have been adopted linked data technology as an effective data integration infrastructure, there exist also pessimistic opinions, which conclude that ICT companies have not adopted Linked Data as well as Big Data technology[o]. Large software vendors like IBM, Microsoft etc. have not been interested in Linked Data so far. Linked data is a specific technological platform for them and they do not see profitable combined effect with their own products. However, Oracle provides RDF database solutions [Oracle (2014)]. At the same time, IBM and others have integrated open source Big Data tools like Hadoop and MapReduce with their own products.

On the other hand, there have been emerged specialised companies who offer Linked Data technology based solutions like MarkLogic, OpenLinks Software, Franz, etc.

In public sector, Linked Data technology is used for building information systems that integrate data from closed, open or both sources. Closed Linked Data are used in national security and defence domains. For example, recently USA re-constructed its healthcare portal[p] that publishes partially open data according to Linked Data technology principles. Portal of statistical data TheDataWeb[q] uses among other technologies also Linked Data principles for data publishing and integration. There are available open data sets for re-use by companies like OpenCorporates[r] Linked Data cloud that includes data about 50000 enterprises all over the world. Another example is Product Ontology [s] that classifies and includes information about more that 1 Million products. Many governmental open data portals publish also Linked Open Data (e.g. in USA, UK, EU etc.).

3. Big Data Technologies

3.1. *Concept*

Big Data is characterized by the following properties: volume, variety, variability, velocity, and also the complexity. Recently, the feature of the veracity has been added to the concept of Big Data. In general, data is considered to be Big Data if volume of data set is so massive (e.g. measured in petabytes) and data are so complex that traditional database techniques are not efficiently applicable for processing these data. Variety refers to different formats of Big Data, for example, structured and unstructured formats. Variability characterizes Big Data from the point of view of analysis of data pointing to variable options for analysis and interpretation of the results of analysis. Velocity

[n] http://semanticweb.com/sindicetech-helps-enterprises-build-private-linked-data-clouds_b30454
[o] http://blog.semantic-web.at/2013/06/04/theres-money-in-linked-data
[p] www.healthcare.gov
[q] http://thedataweb.rm.census.gov/TDW.html
[r] http://opencorporates.com
[s] http://www.productontology.org

describes how fast data are generated and delivered for analytical purposes. Usually, real time or near real-time delivery is considered in the case of Big Data. Complexity is an important feature of Big Data as data sets can be very complex linking data from multiple disparate sources. Veracity refers to such characteristics of data like reliability, lack of errors, consistency, etc.

Big Data technologies can be divided into two groups as follows:

- Big Data curation and storage technologies
- Big Data analytics. It is used for finding hidden patterns and correlations from collected data in order to apply obtained knowledge to better decision making. These technologies include high performance data mining, predictive analytics, text mining, forecasting and optimization.

3.2. *State-of-the art*

In 2010, Gartner [Gartner (2010)] predicted that volume of enterprise data of any format will grow about 650% in next 10 years. This data volume growth is a challenge for data processing technologies.

Comparing to relational databases, Big Data is not so well and uniformly structured. Big Data are created by machines (e.g. sensors) or by people with very different background and experience (e.g. via social media). This sets up new requirements for Big Data processing technologies, analytical tools, data mining algorithms and data visualization techniques.

Nowadays, Big Data reside in almost all sectors (e.g. banking, healthcare, insurance, manufacturing, marketing, transport, etc.). This in turn creates a market demand for technologies of Big Data processing and analytics having positive influence to development of corresponding technology sector. One of the main information technology challenges with this respect is management of volume of data and creation of corresponding scalable data processing architectures as traditional database technologies are not any more suitable for handling these new volumes of data. One of the unsolved problems related to Big Data is semantic interoperability as data variety and variability are growing. Semantic interoperability of Big Data is becoming one of the most important problems of effective Big Data management [Haav and Küngas (2014)]. Big Data are mostly unstructured data. Therefore, availability of (semantic) metadata is crucial. Linked Data technologies could help linking different Big Data sets and providing a wider view to data.

On the other hand, Big Data analytical capacity is very important as it gives business value to Big Data by digging deeper knowledge from existing data. T. Pellegrin [Pellegrin (2014)] considers macro and microeconomical value of Big Data. He argues that on macroeconomical level Big Data is rather political than technological issue similar to open data that began as a result of USA policy in 2009. After that, open data policy moved to EU, where it became a foundation of European Digital Agenda. Political motivation of open data was to stimulate the economy while Big Data role is to stimulate decision making process based on data analytics and usage of unstructured web data for national mission-critical applications. Therefore, USA announced Big Data Initiative in 2012 in order to express interest in Big Data technology research, development and application [Whitehouse (2012)]. Corresponding EU initiatives include Big Data technology related calls and projects of EU framework programmes (FP7 and Horizon 2020) and also wider forums like BIG Project[t].

Open source basic technologies for Big Data processing like Hadoop[u] , MapReduce[v] and related applications such as Cloudera[w] and Hive[x] and adoption of these by large software companies as Oracle and IBM have contributed to take-up of Big Data technologies.

At the present time, basic technology for Big Data analytics is parallel or NoSQL database extended with Hadoop connection. Hadoop is used for processing of unstructured Big Data. For example, Amazon is using Hadoop. Hadoop is widely used because of its scalability with respect to both growing volume of data and expanding network with new nodes. The latter creates good opportunities for developing applications that meet needs of companies. Another important aspect of Hadoop is that it is open source software.

3.3. *Future trends*

According to Gartner survey [Gartner (2013b)] among 720 companies, 64% of organizations planned Big Data projects for 2013. Media, communication companies and banks have constituted the top of the list. More than one third of media and communication companies told that they already have invested to Big Data analytics projects. Therefore, Gartner called the year 2013 Big Data experimentation and early adoption year.

At the same time, Gartner makes conclusion that Big Data technology take-up is at early adoption level as less than 8% of respondents answered that their organization already uses Big Data analytics solutions.

The question arises, what prevents companies to use Big Data technologies. According to Gartner research, many companies do not see business value of Big Data technologies. On the other hand, initial investments are not small and return of investments is coming after long period. In addition, Big Data analytics requires hiring highly paid data scientists in the situation where labour market has shortage of them. Usage of services of consulting companies makes projects costs higher.

According to Gartner 2012 analysis of Big Data technology adoption [Gartner (2012c)], technologies of Big Data were positioned just below "*The Peak of Inflated Expectations*" giving the 2-5 years for reaching to this plateau.

In Gartner 2014 report *Hype Cycle for Emerging Technologies* [Gartner (2014)], Big Data is not any more at the top of the Hype Cycle as it moved down to the level of

[t] http://www.big-project.eu
[u] http://hadoop.apache.org
[v] http://research.google.com/archive/mapreduce.html
[w] http://www.cloudera.com
[x] http://hive.apache.org

"*Trough of Disillusionment*". However, there is a new entry in the Hype Cycle, called Data Science that is related to Big Data analytics. According to Gartner, Big Data technologies "*moved beyond the top of the cycle because the market has settled into a reasonable set of approaches*". This may indicate that Big Data technologies will foresee a growth in the massive market and are not so interesting for early adopters any more.

On the other hand, fast introduction of Data Science to Hype Cycle may indicate that Big Data analytics will gain importance in decision making process.

3.4. *Deployment*

From the point of view of usefulness, Big Data technology has to some extent already proved its utility. However, take-up of Big Data technologies and its economic benefit in business and the public sector grow according to the rate of capacity of usage of these technologies for solving complex problems. In this relation, a bottleneck is not the technology but analytical competence of data processing staff. This situation creates a demand for data analysts or Data Scientists on labour market. For example, UK Economics and Business Research Centre forecasts that 36000 new start-ups and 60000 new job sites will be created in the field of Big Data analytics in UK by 2017 [Pellegrin (2014)].

Areas of usages of Big Data analytics include business resource planning, research and development activities, marketing and risks management. Many social media companies like Google, Twitter, Facebook, LinkedIN etc. apply Big Data management and analytics for enhancing their customer base.

A number of database management and data warehouse software vendors (e.g. IBM, Oracle, Teradata, SAP EMC, HP, Amazon, MS, Google, etc.) [Korolov (2013)] have integrated Big Data processing facilities, basically open source Hadoop and MapReduce, to their products.

On the other hand, there are companies like HaDapt, Platfora, YarcData, SiSense, Space-Time Insights, Zettaset etc. who are oriented specifically to Big Data technologies and analytics.

Recently, IBM Research created Accelerated Discovery Lab [IBM Research (2013)] for Big Data research and development to San Jose. Lab already manages 15 Big Data projects in fields like healthcare, pharmacy, geoinformation systems, water-resource management, social media, etc.

According to SAS Institute study *Big Data in Big Companies* [Davenport and Dyché (2013)], it is clear that big industrial companies are starting to use Big Data processing and analytics even if it needs time and new technological approach, new organizational and management structures and new skills. For example, large companies like UPS, General Electric etc. use Big Data technologies for optimization of their manufacturing process collecting data using sensors.

USA Central Intelligence Agency (CIA) is interested in Big Data technologies for collecting and analysing intelligence data [Petraeus (2012)]. Automatic data curation and analytics provide opportunities to find hidden patterns in data and relationships between data and events.

It is quite clear that adoption of Big Data technologies allows increasing economic growth. However, social aspects of Big Data need to be taken into consideration. These are security, privacy and copyrights insurance.

4. Internet of Things

4.1. *Concept*

Internet of Things (IoT) means equipment of physical objects (e.g. things, devices) with sensors that are connected via wired or wireless network using Internet Protocol (IP). Such sensor-networks or physical information systems create a huge amount of data that are sent to computers for analysis. IoT is a part of so-called Future Internet technologies.

K. Ashton who first introduced the term of "Internet of Things" in 1999, in his article *That "Internet of Things" Thing*, [Ashton (2009)] very well explained the concept as follows:

"*We need to empower computers with their own means of gathering information, so they can see, hear and smell the world for themselves, in all its random glory. RFID and sensor technology enable computers to observe identify and understand the world—without the limitations of human-entered data*".

4.2. *State-of-the art*

According to Wikipedia, in 2013, estimated 39% of 7.1 billion inhabitants of the world have been connected to the Internet [Wikipedia (2014)].

Today, we have reached a situation, where more devices (things) than people are connected to the Internet and it is predicted that this number soon exceeds the world population becoming 20-50 billion of connected smart devices. At the same time, this also means that the Internet protocol IPv4 that was designed for 4 billion addresses is exhausted [Ziegler *et al*. (2013)]. Adoption of new Internet protocol IPv6 is important in relation with expansion of IoT. According to EU project IoT6[y], 2012 was starting year of global transition to IPv6.

Adding network and sensor capacity to things has lost it's technological and price barriers. For example, this capacity could be added to consumer products like toothbrush in order to follow habits of oral hygiene (see e.g. BeamBrush[z]), etc.

PwC in its recent technology study on IoT [Baya and Parker (2013)] concludes that growth of IoT is very fast in coming years. They consider 3 layers of IoT as follows:

- Sensors that are embedded into devices or physical environment for collecting needed data or information about events

[y] http://www.iot6.eu
[z] http://www.beamtoothbrush.com/toothbrush

- Network and computing platform that shares information with sensors and works according to this information in order to affect the environment
- Service platform that aggregates and performs analysis of data, and takes care of overall customer experience.

On each of the layers, rapid development of new technologies and solutions are foreseen.

Sensors that have been used for many years in automotive industry, healthcare, and manufacturing are becoming small and inexpensive to be embedded in all devices and the physical environment. According to IC Insights [IC Insights (2012)], sensor sales worldwide will increase 18 % between 2011 and 2016. Sensor market is driven by smart phones that have embedded sensors for capturing contextual information like location, movement, light and other features of environment. These features in turn are used for provision of new services by service providers.

Network and computing platform provides processing, local memory and connectivity. IoT requires internet connection for combining sensor information with information residing on cloud. For example, CeNSE project[æ] by HP labs combines new solutions from material- and nanotechnologies as well as in the field of electromechanical microsystems in order to create a sensor-network that contains billions and trillions tiny, cheap and resistant sensors. In the project web site they tell that „*CeNSE consists of a highly intelligent network of billions of nanoscale sensors designed to feel, taste, smell, see, and hear what is going on in the world*.“ Potential usage of this type of sensor-networks will be in infrastructure objects like roads, bridges, buildings. In additions, these could be used in airplanes, manufacturing, healthcare, security assurance etc.

Service platform layer is needed for emerging use-cases of IoT. This layer includes middleware, analytics and application software in order to combine sensor-data with other important contextual information.

4.3. *Future trends*

According to Cisco predictions in [Evans (2011)], the number of devices connected to the Internet will be around 50 Billion by 2020.

The Editor of IoT World R. Quinell predicted significant growth of IoT and machine-to-machine (M2M) communication already for 2014 [Quinnell (2014)]. The prediction is based on Deutsche Telecom’s prediction for M2M technology developments for 2014 [Hase (2013)]. First trend is expansion of global alliances, because IoT and M2M need gross boarder interoperability solutions. These should include common protocols, data formats and access technologies. Second trend concerns integration of Big Data analytics and IoT solutions for providing new services. IoT and M2M solutions create a huge amount of data enabling new ways of data mining and innovative usage of data analytics.

Gartner study *Hype Cycle for Emerging Technologies 2013* [Garner (2013b)] concentrates to the relationship between humans and machines taking into account

[æ] http://www8.hp.com/us/en/hp-information/environment/cense.html#.Uue80BD8LRY emerging technologies like smart machines, cognitive computing and the Internet of Things. According to Gartner, there are three trends in usage of these emerging technologies as follows:

- Augmenting humans with machines (e.g. usage of wearable computing devices)
- Replacing humans by machines (e.g. a cognitive virtual sales-assistant)
- Co-operation of humans and machines (e.g. a human in cooperation with a robot).

Concerning development of cognitive computing, USA and EU have launched specific projects related to cognitive systems. EU has Human Brain project[ø] and USA started in 2013 BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies)[å] .

Components of cognitive systems are neurosynaptic chips which architecture is inspired by structure of human brain [IBM Research (2014)]. Such chips have been demonstrated already in 2011 during the presentation of SyNAPSE project by IBM and DARPA. Long term goal of IBM is to create a system of neurosynaptic chips that consists of 10 billion neurons and 100 trillion synapses. Neurosynaptic chips are intended to complement cognitive systems like IBM Watson [Ferrucci (2012)]. Watson will simulate activities of the left hemisphere of the human brain (language understanding and analytics) and cognitive chips will deal with activities of the right hemisphere of the human brain like cognition and image recognition. An aim of IBM Research is integration of these two capacities into one as we may see it in the case of human brain [IBM (2011)].

4.4. *Deployment*

Nowadays, IoT systems are used in many fields like smart homes or cities, transport, healthcare, etc. According to Gartner [Gartner (2012a)], value of IoT does not lie so much on devices and their connectivity but rather in data integration; data processing and decision making that make it possible to use data streams originating from devices or control devices from a distance.

Automobile industry is one of the main adopters of M2M technologies. Deutsche Telecom predicts [Hase (2013)] that in 2014 the trend is to continue and more M2M solutions will appear at the market. Current trend is that services that provide added value to the customers are not any more tightly related to only one certain automobile maker.

In the public sector, IoT usage in smart city solutions (e.g. smart control of traffic and street lights) will see a breakthrough in the nearest future.

Another interesting trend related to M2M is a breakthrough from business-to-business (B2B) market to consumer market. More and more smart consumer products will appear on the marketplace. Market of personal tracking devices and wearable technology is to grow fast in 2014. Prices of personal tracking devices are going down and various smart

[ø] https://www.humanbrainproject.eu
[å] http://www.whitehouse.gov/share/brain-initiative

watches, trackers of different characteristics of human body will come to the consumer market.

Extensive take-up of IoT has also downsides. Networked things could become target objects for spying. At the same time, IoT technology could be used for military and counterintelligence purposes.

In his speech on In-Q-Tel[aa] Summit in 2012, D. Petraeus listed technologies that are of interest of CIA as follows [Petraeus (2012)]:

- Item identification and tagging devices
- Sensors and wireless sensor networks
- Embedded systems
- Nanotechnology that enables to produce devices being small enough to operate anywhere.

These are also technologies that are driving development of IoT and Big Data.

5. Social Media

5.1. *Concept*

Social media[bb] is a general notion to denote a variety of web sites or services that allow users to create content and interact with each other.

The most widely used types of social media are as follows:

- *Social Network Sites/Services.* These are web sites or services that allow people to connect with other people of similar interests and background. We may say that these are social networks supported by web or sometimes also called social web networks. The most popular examples are Facebook and LinkedIn. It is interesting to mention that according to [boyd and Ellison (2008)], most of social web networks support already existing online or offline social relationships.
- *Publishing Services*. These are services that allow people to upload and share files of different types like audio, video, pictures, text, etc. These services can also have additional social features such as profiles, commenting, etc. Well-known and widely used examples are YouTube, Flickr and WordPress.
- *Microblogging Sites/Services.* These services enable their users to post small text messages that will update their personal profile and will also be delivered to their list of friends. Microblogging services became very popular due to Twitter, however there are other microblogging sites as well.
- *Blog Comments and Forums*. Online forums facilitate conversations between their members by posting messages. Blog comments are similar but the content is arranged around the topic of the blog post and as such allows users to publish their content to the audience interested in the topic.

[aa] https://www.iqt.org
[bb] http://www.techopedia.com/definition/4837/social-media

- *Crowdsourcing*. This is web based distributed model of problem solving and creation of products and services [Daren (2008)]. This is a process in which work is ordered or funding is received from the crowds acting online. One of the examples is OpenStreetMap project[cc].

In addition to the list above there are many other interesting social media types like playing, buying, localisation, etc.

5.2. *State-of-the art*

According to statistics portal Statista.com, as of 2014, more than 1.8 Billion Internet users have accessed social networks and it is estimated that about 2.33 Billion social network users will be in 2017. These figures indicate that user engagement in social media will continue to grow.

Concerning the social network penetration rate in 2013[dd], North America was ranked first with a social media penetration rate of 56%, followed by Western Europe with 44%. The global average penetration rate was 26%. Facebook users form 42.6% of global online population[ee].

Social media users in USA prefare mobile devices to spend their time in social networks. From overall time spent in social networks, in USA, 98% of Instgram users, 90% of Twitter users, and 68% of Facebook users did it via a mobile device[ff].

According to these figures and [Adler (2014)] multi-device social media user engagement is a current state of the art that might be changed to prevalence of mobile device engagement in the future.

5.3. *Future trends*

The main future trend in user engagement of social media is fast growth and expansion to new geographic areas. According to [Jones (2014)], growth of number of users of most popular social media sites like Facebook, Twitter, Google + etc. has been exponential.

According to F. Cavazza [Cavazza (2012)], during the last 10 years there have been three waves of activities of social media usage: the publishing (e.g. blogs), the sharing (e.g. social web networks like Facebook and Twitter) and the curation (e.g. platforms like Quora and Pinterest) waves. The shift in users' behavior to content curation is caused by increasing volume of users' provided content that makes it challenging to find really valuable content.

Another trend in social media users' behaviour is that they move away from the Facebook like social networks to the direction of apps that opposite to Facebook profile based representation let users represent themselves to different people in different ways at different times [MIT (2014)]. One of the examples is the mobile app Snapchat that

[cc] http://www.openstreetmap.org
[dd] http://www.statista.com/statistics/269615/social-network-penetration-by-region/
[ee] http://www.statista.com/statistics/241552/share-of-global-population-using-facebook-by-region/
[ff] http://www.statista.com/statistics/294445/minutes-spent-on-us-media-sites-by-platform/

enables to send photos and short videos ("snaps") to chosen recipients instead of sending text messages. At the same time, Facebook also has worked in this direction and turned back to forums and chatrooms by announcing new pseudonymous mobile app Rooms[gg] that enables users to set-up their own discussion space about any topic.

A growing trend related to social web networks is also sharing geographic information and location based features (e.g. Facebook, Twitter, OpenStreetMap etc.) using GPS equipped smartphones. In general, social networks that are equipped with user generated geographic information are called Location Based Social Networks and these became a challenging research subject [Roick and Heuser (2013), Gordon, and de Souza e Silva (2011)].

5.4. *Deployment*

Social media in general and social web networks in particular are highly exploited by marketing and advertising business. Growing number of social media users is in correlation with growing efforts in social media marketing.

Recent social media marketing study [Stelzner (2014)] provides results of survey that included about 2800 marketers that were asked questions about several aspects of social media marketing.

A significant 92% of marketers indicated that social media is important for their business. Facebook, Twitter, LinkedIn, YouTube, blogging, Google+ and Pinterest were the top seven platforms used by marketers. Facebook and LinkedIn are the two most important social networks for marketers. More than half of marketers (54%) chose Facebook as their most important platform, followed by LinkedIn (17%), Twitter (12%) and then blogging (8%). However, only 34% of marketers think that their Facebook efforts are effective. Surprisingly, 58% of marketers stated that original written content is the most important form of content, followed by original visual assets (19%).

Crowdsourcing is another growing area of deployment of social media. Crowdsourcing is ensured by creative communities which have been grown fast during 5 last years. E. F. Petavy highlights in his study *Five predictions for crowdsourcing in 2014* that an estimated number of 10 Million people all over the world contribute to crowdsourcing [Petavy (2014)]. The study points out that the world is changing to become multi-modal meaning that a person does not need to have only one job but after the ordinary work he/she can contribute to some creative crowdsourcing project. This phenomenon is denoted by a term *cloud labour*. These are virtual distributed workers who are available on demand. Using this kind of labour force creates also problems like copyright insurance, social guarantees, etc.

Deloitte in its study on industrial crowdsourcing highlights six most important crowdsourcing platforms as follows: Gigwalk, oDesk, Kaggle, Tongal, Quirky and Kickstarter [Deloitte (2014)]. For example, community size of Kickstarter was 5.4 Million and oDesk 4.5 Million. At the same time, Gigwalk had the biggest number of

[gg] http://www.rooms.me/]

completed tasks being 4 Million. Deloitte study concludes that crowdsourcing is at stage of early adoption.

Social networks analysis is widely applied on social web networks and other social media. Current social media analytics focuses to three main areas as follows [Gastelum and Whattam (2013)]:

- Content analysis (e.g. finding popular topics, mining sentiment or mood)
- Group or network analysis (e.g. classifying user groups, characterizing interactions between group members and identifying influential users)
- Prediction of real-world events or characteristics.

Different time critical application scenarios such as disaster response can benefit from monitoring geographic information from social networking sites. For example, Twitter messages about Sichuan Earthquake in 2008 appeared within seconds close to the event [Li and Rao (2010)].

Challenges of social media analysis are related to Big Data and Big Data analytics. Linked Data technology can be used to represent social network graphs and related information.

6. Interlinking Information Technologies via Linked Data

6.1. *Information Technology connections*

Linked data technologies are used as information infrastructure for integration of data from disparate sources. Comparing to mainstream technologies of relational databases, linked data technologies have advantages in cases where relational databases demonstrate slow performance and inefficacy, namely integrating (joining) data about a certain object. Another advantage of linked data technology is its ability to automatically integrate data from RDF database with data from external sources like web data or sensor data, etc. Linked data could provide structure (RDF triples) and metadata to unstructured data like textual data or several data streams.

At the same time, most RDF databases have currently slower performance (except join operation) than most commercial relational database systems. However, there are good commercial RDF datastores available (AllegroGraph, Big OWLIM, Marklogic, etc.) and also open source RDF database systems like Virtuoso, 4store, Bigdata etc. [Auer (2011)].

Comparing to web services technology that also enables to create data-integration services, linked data technologies are lighter, more flexible, faster and as such cheaper. Linked data technology is built on existing web standards (and open source RDF stores) that enables small and medium size companies to use it for development of data integration solutions. They might not have resources for building data integration solutions based on web services technology.

In the following Fig 1, most important connections of linked data technology with other information technologies considered in this paper are depicted. In addition, the figure also shows relationships between these other technologies.

As depicted in Fig. 1, Big Data explosion is driven mainly by the following technologies:

- IoT that provides various sensors and connected devices collecting data about environment or human activities
- Social web generating a huge amount of data (e.g. audio, video, click-streams, etc.) by different online activities. In addition, social web networks share a huge amount of data of different types.

Nowadays, the IoT data is not only sensor data. Data from social media and data about measurements or observations submitted by users are integrated to pure sensor data and form together IoT data [Sheth *et al.* (2013)]. On the other hand, social media data integrates a part of sensor or device data as people can easily share their measurements of real world objects or their observations about themselves via social media. If things are given capabilities to automatically create social media data or things will be reached via social media, then this phenomenon extends paradigm of Web of Things towards networking of things.

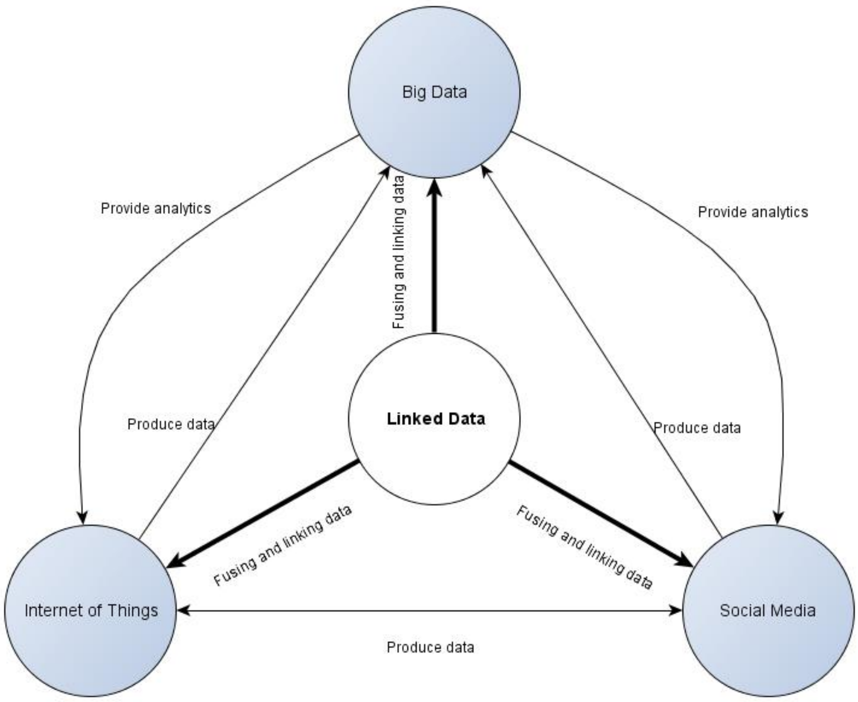

Fig. 1. Information Technology connections.

In this relation, a new paradigm called Social Internet of Things[hh] (SIoT) has emerged. The idea is to separate humans and things allowing things to have their own social networks i.e. social networks of smart objects [Atzori *et al*. (2014)]. However, these separate social networks of things should be controlled by humans. As reported by L. Atzoni et al. [Atzori *et al.* (2014)] there are a many open research issues concerning SIoT. One of these is how to define inter-objects relationships and build their semantic model. In this respect, Linked Data technologies could be useful.

Big Data analytics can be applied on both social media data (e.g. social networks analysis) and IoT data (e.g. integrating analytics to situation and context aware systems).

According to Fig. 1, linked data technologies play a central role of integrator of other technologies. Link data technology provides tools for the following types of data processing:

- Data fusion
- Structuring of semi- or unstructured data
- Providing of semantic metadata
- Linking of data
- Usage of common web standards for representation of data (e.g. RDF) and metadata or vocabularies (e.g. OWL).

Other technologies like Big Data, social media and IoT can benefit from applying linked data technology tools for processing collected data and developing new innovative applications. In the following subsections, linked data connections with each of analysed technologies are examined in more detail.

6.2. *Linked Data and IoT*

Linked data technology can provide semantic data model for IoT. Data models facilitate interoperability of M2M systems and machine understanding. They describe connected things to applications, drive discovery and linking, and allow abstraction of particular M2M protocols (e.g. M2M SOA, M2M HTTP, etc.). For example, M. Koster in [Koster (2013)] introduces a Semantic Proxy that uses a common data model for supporting M2M communication.

Semantic interoperability of IoT entities is very important for sensor networks. W3C standardization efforts in this area led to the Semantic Sensor Network (SSN) ontology[ii] that represents a semantic model for describing sensor devices, their operation, observation and measurement data, and process related attributes of sensors.

Linked data could be used for data integration or fusion of IoT data that originates from disparate sources (e.g. from different devices or a human). In general, IoT data is collected about an entity in the physical world. This data could be combined with other data or knowledge and integrated in order to create applications that support context and

[hh] http://www.social-iot.org
[ii] http://www.w3.org/2005/Incubator/ssn/ssnx/ssn

situation awareness. For example, W. Wang et al. [Wang et al. (2013)] created linked IoT data and corresponding ontologies that make use of external links to spatial data.

L. Atzori et al. [Atzori *et al.* (2010)] consider three visions of IoT: Internet-oriented, Things-oriented and Semantic-oriented visions. Semantic-oriented vision is related to utilization of Linked Data principles for uniquely addressing things (e.g. by URIs), representation and storage of data that is to be exchanged between things and between things and the rest of the world.

A future vision of IoT that is called ‘‘*Web of Things*” [Guinard and Vlad (2009)] creates new challenges for utilization of Linked Data technologies used so far in the context of “*Web of Data*”. According to the vision of „*Web of Things*” every device or tangible thing will be connected and integrated using Web standards. Linked Data technology could play central role in this new Web. URI-s are candidates for uniquely identify real world entities, RDF standard enables represent and store data about them and interlinking enables represent relationships between entities themselves as well as with external objects.

Some of the attempts to this direction are already done. For example, T. Ermilov and S. Auer [Ermilov and Auer (2013)] developed an approach that provides embedded and smart devices with a Linked Data interface. Their approach is based on mapping existing structured data on the device to vocabularies and ontologies and exposing this information as dereferenceable RDF directly from the device. The approach allows smart devices (e.g. tablets, smart phones, etc.) to identify and describe themselves by providing information in the form of Linked Data. The approach is implemented according to the concept of Embedded Linked Data Server, specifically as Android Linked Data Server.

6.3. *Linked Data and Big Data*

In the Big Data era where most important drivers of explosion of data volume are IoT and social media the role of Linked Data technologies cannot be underestimated. The Linked Data approach could solve some of the Big Data management problems.

First of all, using Linked Data technologies reduces variety aspect of Big Data or semantic heterogeneity of data enabling to cope with semantic data interoperability problems of Big Data. Due to volume and velocity issues of Big Data, manual curation of variety using methods established for traditional databases are not sufficient any more.

P. Hitzler and K. Janowicz [Hitzler and Janowicz (2013)] point to the following features of Linked Data that help to solve semantic interoperability problems:

- Agreement to use RDF for data representation
- Commitment to use vocabularies (or ontologies) created using standard formally well-defined languages (e.g. OWL)
- Existence of compatible open source Linked Data tools for accessing, storing, linking and querying data
- Linked Data are mostly object centric.

These observations are supported by experience with Big Data solutions in Fujitsu Europe presented by I. Mitchell and M. Wilson in [Mitchell and Wilson (2012)]. They found that main problems of Big Data management cannot be resolved only by deployment of Hadoop and integrating SQL and NoSQL database systems. The main question of Big Data is related to semantic interoperability or how to bring together structured and unstructured database models. I. Mitchell and M. Wilson point out that optimal solution would be the usage of Linked Data as the reference for big data systems in order to query, search and provide analysis of Big Data. Linked data enables to link traditional OLTP and OLAP databases without changing their existing database schemas. The Linked Data component could maintain relationships between databases to keep them consistent and existing applications continue to function whilst the data is exposed to other applications for consumption.

As Linked Data are structured data (RDF triples), then using Linked Data for querying data from multiple sources provides a highly scalable solution to be used as a mediator for variety of Big Data sources. Linked Data can be used for example for mapping and interconnecting and indexing of Big Data sources. In addition, data analytics could be run on the Linked Data.

However, Linked Data usage requires metamodels (i.e. vocabularies or ontologies) that should be managed under strong control. Otherwise, Linked Data structural properties do not create any benefits.

6.4. *Linked Data and social media*

Linked Data could be used in social media sites for many purposes like named entity disambiguation of keywords, categorization of information and improving of web site search and discovery.

For example, P. Singh and N. Shadbolt [Singh and Shadbolt (2013)] used Linked Data principles in crowdsourcing social network site for facilitating better search of topics and finding relevant topic experts.

Linked Data is utilized to support faceted search that can be widely used in web sites with very different purposes. For example, PoolParty[jj] is well-known by its faceted browser and search facilities based on Linked Data principles. Serendipity[kk] is used for faceted search of online courseware, Facete[ll] is used for library search and there are many other Linked Data based faceted search engines available.

Linked data is also widely utilized for linking and visual analytics of social media data like social web networks. For example, well-known social network company Facebook launched Open Graph protocol[mm] for accessing its social network. Facebook Open Graph representation uses RDFa and particular metadata. This is a step towards Linked Data and semantic web.

[jj] http://www.poolparty.biz
[kk] http://serendipity.utpl.edu.ec/index.html
[ll] http://aksw.org/Projects/Facete.html
[mm] https://developers.facebook.com/docs/opengraph

Another example is Semantically-interlinked online communities' initiative (SIOC) project[nn] . One of its main contributions is SIOC ontology for representation of social web network data in RDF facilitating utilization of Linked Data.

7. Discussion and Conclusions

A new *web of everything* is emerging as more people, places, and things are connected over Internet. Concerning this new web, developments of the following Information Technologies are important: Linked Data, Big Data, Internet of Things and social media. This paper surveyed the most important development aspects of these emerging Information Technologies. The paper has shown relationships between these technologies and pointed to the central role that Linked Data technology plays among them.

As a result, we conclude that Linked Data technology has the potential to serve as an integrator of these technologies by providing tools for interlinking data produced as Big Data, IoT or social media data. We have specifically shown what benefits the usage of Linked Data principles can provide for IoT, Big Data and social media.

For IoT, Linked Data can provide a semantic model and help to solve semantic interoperability problems of IoT. For Big Data, Linked Data technologies reduce variety aspect of Big Data enabling to cope with semantic data interoperability problems of Big Data. For social media, Linked Data could be used for representation and analysis of social web networks as well as for faceted search.

As a result, Linked Data technology is a good candidate for realisation of the vision of *web of everything* as URIs could be used for uniquely identify real world entities, RDF standard is good for data representation and interlinking could be used for representing relationships between entities and between entities and external objects.

There are challenges in deployment of Linked Data as an integrator of Big Data, IoT and social media. Linked Data is based on semantic web standards but IoT field is very complex and wide. IoT has already its own interoperability standards and it is not clear weather big vendors like to commit to semantic web standards. The same is about social media companies like the market leader Facebook. Even Facebook Open Graph is now available via corresponding API as RDF, it is not clear does Facebook really go for semantic interoperability in its wider meaning. Concerning Big Data, it seems that Linked Data well serves integration of different Big Data sets. However, providers of Big Data analytics are not so much interested in Linked Data technology.

Acknowledgments

This research was supported by the target-financed theme no. SF0140007s12 of the Estonian Ministry of Education and Research, and by the European Regional Development Fund (ERDF) through the projects MBJSDT and EXCS.

[nn] http://sioc-project.org

References

Adler, E. (2014): Social Media Engagement: The Surprising Facts About How Much Time People Spend On The Major Social Networks. http://www.businessinsider.com/social-media-engagement-statistics-2013-12. Accessed 30th of October, 2014.

Asthon, K. (2009): That 'Internet of Things' Thing. RFID Journal, http://www.rfidjournal.com/articles/view?4986. Accessed 11th of September, 2014.

Atzori, L.; Iera, A.; Morabito, G. (2010): The internet of things: A survey. Computer Networks, 54(15), pp. 2787-2805.

Atzori, L.; Iera, A.; Morabito, G. (2014): From 'smart objects' to 'social objects': the next evolutionary step of the Internet of Things, IEEE Communication Magazine, pp. 97-105.

Auer, S. (2011): The Semantic Data Web. LOD2 project, http://www.slideshare.net/lod2project/the-semantic-data-web-sren-auer-university-of-leipzig, Accessed 16th of October, 2014.

Baya, V.; Parker, B. (2013): The Thing Stack: Technologies that guide customers to their goals. Internet of Things: Evolving transactions into relationships, Fred Cripe, F. (Ed), PwC Technology Forecast.

Bizer, C.; Heath, T.; Berners-Lee, T. (2009): Linked Data—The Story So Far. IJSWIS, 5 (3), pp. 1–22.

boyd, d. m.; Ellison, N. B. (2008): Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13, pp. 210–230.

Cavazza, F. (2012): Social Media Landscape 2012. http://www.fredcavazza.net/2012/02/22/social-media-landscape-2012/. Accessed 20th of October, 2014.

Cisco (2013): Internet of Everything. Cisco, http://www.cisco.com/web/about/ac79/docs/IoE/IoE-AAG.pdf. Accessed 10th of October, 2014.

Daren, B. (2008): Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. The International Journal of Research into New Media Technologies 14 (1), pp. 75–90.

Davenport, T. H; Dyché, J. (2013): Big Data in Big Companies, SAS Institute http://www.sas.com/resources/asset/Big-Data-in-Big-Companies.pdf. Accessed 20th of September, 2014.

Deloitte (2014): Tech Trends 2014. http://dupress.com/articles/2014-tech-trends-crowdsourcing/?coll=6210. Accessed 20th of October, 2014.

Ermilov, T.; Auer, S. (2013): Enabling Linked Data Access to the Internet of Things. iiWAS2013, ACM, http://dl.acm.org/citation.cfm?id=2539157. Accessed 10th of October, 2014.

Evans, D. (2011): The internet of things: How the next evolution of the internet is changing everything. http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf. Accessed 20th of October, 2014.

Ferrucci, D. A. (2012): Introduction to “This is Watson”, IBM Journal of Research and Development, 56(3-4), pp. 1-15.

Gartner (2012a): Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data. http://www.gartner.com/it/page.jsp?id=1731916. Accessed 10th of October, 2014.

Gartner (2012b): Gartner, Hype Cycle for Big Data. https://www.gartner.com/doc/2100215/hype-cycle-big-data-. Accessed 20th of October, 2014.

Gartner (2012c): Gartner on Big Data. http://whatsthebigdata.com/2012/08/05/gartner-on-big-data/. Accessed 20th of October, 2014.

Gartner (2013a): Gartner Survey Reveals That 64 Percent of Organizations Have Invested or Plan to Invest in Big Data in 2013. http://www.gartner.com/newsroom/id/2593815. Accessed 20th of October, 2014.

Gartner (2013b): Hype Cycle for Emerging Technologies 2013. http://www.gartner.com/newsroom/id/2575515. Accessed 20th of October, 2014.

Gartner (2013c): Gartner Top Predictions 2014: Plan for a Disruptive, but Constructive Future. Gartner.

Gartner (2014): Gartner's Hype Cycle Special Report for 2014. http://www.gartner.com/doc/2816917?refval=&pcp=mpe. Accessed 20th of October, 2014.

Gastelum, Z. N.; Whattam, K. M. (2013): State-of-the-Art of Social Media Analytics Research, Pacific Northwest National Laboratory.

Gómez-Pérez, A.; Fernandez-López, M.; Corcho, O. (2004): Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web. Springer Heidelberg.

Gordon, E.; de Souza e Silva, A. (2011): NetLocality. Why Location Matters in a Networked World. Wiley-Blackwell, UK.

Guinard, D.; Vlad, T. (2009): Towards the web of things: web mashups for embedded devices. Proceedings of the International World Wide Web Conference 2009.

Haav H.-M.; Küngas P. (2014): Semantic Data Interoperability: The Key probleem of Big Data. Big Data Computing; Akerkar, R. (Ed), Chapman and Hall/CRC, pp. 245-269.

Haav, H.-M. (2006): An ontology learning and reasoning framework. Information modelling and knowledge bases XVII: Kiyoki,Y., et al.(Eds). IOS Press, FAIA, 136, pp. 302 - 309.

Hase, J. (2013): What are Deutsche Telekom's M2M predictions for 2014? http://www.telecomstechnews.com/news/2013/dec/19/what-are-deutsche-telekoms-m2m-predictions-2014/. Accessed 10th of October, 2014.

Hitzler, P.; Janowicz, K. (2013): Linked Data, Big Data, and the 4th Paradigm. Semantic Web, 4(3), pp. 233-235.

IBM (2011): SyNAPSE: a cognitive computing project from IBM research. http://www.ibm.com/smarterplanet/us/en/business_analytics/article/cognitive_computing.html. Accessed 20th of October, 2014.

IBM Research (2013): IBM Research to Accelerate Big Data Discovery. http://www-03.ibm.com/press/us/en/pressrelease/42169.wss. Accessed 5th of October, 2014.

IBM Research (2014): Brain Power. http://www.research.ibm.com/cognitive-computing/neurosynaptic-chips.shtml#fbid=-pHjUq5w8zr. Accessed 20th of October, 2014.

IC Insights (2012): New Embedded Features Will More Than Double Sensor Sales by 2016. IC Insights, Research Bulletin, http://www.icinsights.com/data/articles/documents/431.pdf. Accessed 20th of October, 2014.

IoT6 (2012): IoT6 European research project, Deliverable D1.5: IoT6 Reference Model. White Paper. http://www.iot6.eu. Accessed 20th of October, 2014.

Jones, K. (2014): The Growth of Social Media v2.0. Search Engine Journal, http://www.searchenginejournal.com/growth-social-media-2-0-infographic/77055/. Accessed 20th of October, 2014.

Kontokostas, D.; Auer, S.; and Lehmann, J. (2013): Crowdsourcing Linked Data quality assessment. http://jens-lehmann.org/files/2013/iswc_crowd_qa.pdf. Accessed 10th of September, 2014.

Korolov, M. (2013): 15 most powerful Big Data companies. Network World, http://www.networkworld.com/article/2288276/big-data-business-intelligence/114134-15-most-powerful-Big-Data-companies.html#slide1. Accessed 20th of October, 2014.

Koster, M. (2013: A Modular Open Source Platform for Web Scale IoT Interoperability. http://iot-datamodels.blogspot.com/. Accessed 10th of September, 2014.

Li, J.; Rao, H. (2010): Twitter as a Rapid Response News Service: An Exploration in the Context of the 2008 China Earthquake. The Electronic Journal on Information Systems in Developing Countries, 42(4) pp. 1–22.

Maedche, A.; Staab, S. (2001): Ontology Learning for the Semantic Web. IEEE Intelligent Systems, 16(2), pp. 72-79.

MIT (2014): Emerging trends in social media. MIT, http://newsoffice.mit.edu/2014/emerging-trends-social-media. Accessed 20th of October, 2014.

Mitchell, I.; Wilson, M. (2012): Linked Data. Connecting and exploiting big data. Fujitsu White Paper. http://www.fujitsu.com/uk/Images/Linked-data-connecting-and-exploiting-big-data-%28v1.0%29.pdf. Accessed 20th of October, 2014.

Oracle (2014): Oracle Spatial and Graph RDF Semantic Graph. http://www.oracle.com/technetwork/database/options/spatialandgraph/overview/rdfsemantic-graph-1902016.html. Accessed 30th of October, 2014.

Paquet, R. (2010): Technology Trends You Can't Afford to Ignore, Gartner Webinar. http://www.gartner.com/it/content/1258400/1258425/january_6_techtrends_rpaquet.pdf. Accessed 20th of October, 2014.

Pellegrin, T. (2014): Economics of Big Data: A Value Perspective on State of the Art and Future Trends. Big Data Computing; R. Akerkar, R. (Ed), Chapman and Hall/CRC, pp. 344-367.

Petavy, E. F. (2014): Five predictions for crowdsourcing in 2014. Crowdsourcing, http://www.crowdsourcing.org/editorial/eyekas-francois-petavy-five-predictions-for-crowdsourcing-in-2014/30116. Accessed 20th of October, 2014.

Petraeus, D. H. (2012): Remarks by Director David H. Petraeus at In-Q-Tel CEO Summit. https://www.cia.gov/news-information/speeches-testimony/2012-speeches-testimony/in-q-tel-summit-remarks.html. Accessed 20th of October, 2014.

PwC (2009): Spinning a data Web. Making Semantic Web connections. PwC Technologyforecast.

PwC (2011): Harnessing the Power of Crowdsourcing. PwC Technologyforecast.

PwC (2013): Internet of Things: Evolving transactions into relationships. PwC Technologyforecast.

Quinnell, R. (2014): Predicting 2014's IoT Trends. IoT World, http://www.iotworld.com/author.asp?section_id=3150&doc_id=562274. Accessed 20th of October, 2014.

Roick, O. & Heuser, S. (2013): Location Based Social Networks – Definition, Current State of the Art and Research Agenda. Transactions in GIS, 17(5), pp. 763-784.

Sheth, A.; Anantharam, P.; Henson, C. (2013): Physical-Cyber-Social Computing: An Early 21st Century Approach. IEEE Intelligent Systems, 28 (1), pp. 79–82.

Singh. P.; Shadbolt, N. (2013): Linked Data in Crowdsourcing Purposive Social Network. WWW, SOCM Workshop, ACM, http://sociam.org/www2013/papers/socm2013_submission_2.pdf. Accessed 20th of October, 2014.

Stelzner, M. A. (2014): Social Media Marketing. Industry Report, Social Media Examiner, http://www.socialmediaexaminer.com/SocialMediaMarketingIndustryReport2014.pdf. Accessed 20th of October, 2014.

W3C (2013a): Linked Data Platform Best Practices and Guidelines. https://dvcs.w3.org/hg/ldpwg/raw-file/default/ldp-bp/ldp-bp.html#bib-LDP-PRIMER. Accessed 20th of October, 2014.

W3C (2013b): SPARQL 1.1. Overview. http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/. Accessed 20th of October, 2014.

W3C (2013c): Open Data on The Web. http://www.w3.org/2013/04/odw/report#closed. Accessed 21th of October, 2014.

Wang, W.; De, S.; Cassar, G.; Moessner, K. (2013): Knowledge Representation in IoT: Semantic Modelling and its Applications. Automatika – Journal for Control, Measurement, Electronics, Computing and Communications, 54 (4). 388 - 400.

Whitehouse (2012): Obama Adminstration Unveils "Big Data" Initiative: Announces $200 Million in New R&D Investments. Big Data Press Release, Whitehouse, http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release.pdf. Accessed 13th of October, 2014.

Wikipedia (2014): List of countries by number of Internet users. http://en.wikipedia.org/wiki/List_of_countries_by_number_of_Internet_users. Accessed 15th of October, 2014.

Ziegler. S. et al. (2013): IoT6 – Moving to an IPv6-Based Future IoT. FIA 2013, Galis, A. and Gavras, A. (Eds.), LNCS 7858, pp. 161–172.