

Business Process Optimization with Big Data Analytics Under Consideration of Privacy

Silva Robak

Uniwersytet Zielonogórski, ul.
prof. Z. Szafrana 4a, 65-516
Zielona Góra, Poland
Email: s.robak@wmie.uz.zgora.pl

Bogdan Franczyk

Uniwersytet Ekonomiczny we
Wrocławiu, ul. Komandorska
118/120, 53-345 Wrocław,
Universität Leipzig, Germany
Email: franczyk@wifa.uni-
leipzig.de,
bogdan.franczyk@ue.wroc.pl

Marcin Robak

Universität Leipzig, Germany
Email: robak@wifa.uni-leipzig.de

Abstract—One of the contemporary problems, and at the same time a big opportunity, in business networks of supply chains are the issues associated with the vast amounts of data arising there. The data may be utilized by the decision support systems in supply chains; nevertheless, often there are information privacy problems. The supply chains in cloud will need appropriate administration for support of privacy aspects of co-operating business units existing in big data ecosystems. In this paper we analyze the possibility of utilizing the big data technology for supporting business processes optimization with respect of the privacy regulations in supply chains under the usage of the big data analytics lifecycle. We present our approach on an example of a business process in logistics.

I. INTRODUCTION

THE emergence of Big Data is creating significant new opportunities for business to achieve added value and competitive advantage. Nevertheless the huge volume of data, the complexity of new data types and structures and the speed of new data creation cause problems in utilizing it in established business solutions like SCM Supply Chain Management [1] to attain the competitive advantages. From the business point of view using Big Data in the viable ways in the logistics requires gradual convergence between the Big Data Analytics Lifecycles process stages on one side and established ways of business process modeling in logistics on the other. What's more the emergence of data deluge and diversity of data types cause problems to be solved in the context of business process modeling. Additionally the security and privacy of data in logistics have to be treated in a way appropriate for business stakeholders.

In the paper we will approach the problem of utilization of vast amounts of data in supply chains with respect of the data privacy issues. We will investigate the big data ecosystem, and the process of big data analytics. We consider the big data analytics process from a supply chain perspective with emphasis on showing the differences between the traditional approach and the approach by using big data. The proposal of the modeling of data as an integration solution for business process management in supply chains networks we have already presented in [2], and [3]. We also have considered the research problems associated with big data utilization in logistics and supply chains design and management in [4]. In this paper we mainly review possible changes in

the business processes which result from the above stated requirements and further implications based on big data analytics lifecycle, regarding privacy aspects.

The supply chains define the network that comprehends all the organizations and activities associated with the flow and transformation of goods from the raw material stage, through to the end user, as well as the associated information flow [5]. In the paper we will concentrate on the networked supply chain activities and flow of information.

In the inter-organizational information systems, which link the companies to their suppliers, distributors and customers, a movement of information through electronic links takes place across organizational boundaries, between separately owned organizations. It requires not only the electronic linkage in form of basic electronic data interchange systems (as for purchase orders), but also interactions in complex cash and information management systems or by access to shared technical databases. So the problems with the privacy are very persistent in supply chains contexts.

A business process consists of one or more than one related activities that together respond to a business requirement for an action [5]. The processing steps in a workflow may undertake numerous transformations of data (geographic, technological, linguistic, syntactical and semantic transformations), communication is an important part of the process and (e-) business processes exist within certain environments. In the dynamic business environment, such as networks of venture participants involved in value chains in logistics, where data is forwarded among different enterprises, the appropriate arrangements of privacy aspects are essential.

Therefore, as stated previously, in our paper we will analyze how the big data analytics results can influence the business processes and their privacy aspects. For this aim the rest of the paper is organized as follows.

In Section 2 we characterize the main features of big data, the big data ecosystem elements. In Section 3 we summarize the big data analytics lifecycle process phases, key stakeholders and artifacts. In Section 4 we consider privacy regarded from the perspective of enterprises collaborating in logistics. We also give an example of a business process using big data analytics results in near-real-time in a supply chain. In the example we show how it can support privacy

from the perspective of enterprises. In the last Section we conclude our work.

II. BIG DATA

The big scale usage of available and generated data is made possible for organizations owing to cloud computing paradigms, such as Infrastructure as a Service (IaaS), Storage as a Service, which revolutionized the way the computing infrastructures are used [6]. As big data is referred to data that goes beyond the processing capacity of the conventional data base systems. In addition to this aspect that it is big, e.g. a huge number of small transactions, or (continuous) data streams from sensors, mobile devices etc., it may move too fast, or do not fit the structure of traditional (i.e., relational) database architectures.

According to [7] when we denote a big amount of data as “big data” it has to cover the three “Vs” (features) such as: volume, velocity and variety. Another authors (e.g. [8], [9]) add another V-features like value or veracity.

The first feature - volume of big data - denotes its massive character. The big volume of data is beneficial for the data analysts. It may improve the analytics models by having more cases available for forecasts and increase the number of factors to be considered in the models making them more accurate. Nevertheless, the volume feature bears potential challenge for the IT infrastructures to deal with big amounts of data, especially taking into account its second feature – velocity.

The second feature of big data is the velocity in which data flows into organization or the expected response time to the data. Big data may arrive quickly - in real-time, or near real-time. If data arrives too quickly the IT infrastructures of the organization may be not able to respond timely to it, or even to store all of it. Such situations may lead to data inconsistencies [10].

The third feature of big data is the variety of data. Big data may have diverse structures and forms, not falling into the rigid relational structures of SQL databases without loss of information. Some of data may be saved as blobs in inside traditional data bases. Therefore the IT infrastructures for big data are denoted as NoSQL, what means the data is “not only SQL” [9]. Several examples for diverse kinds of data are standard business documents, transactional records, and unstructured data in form of images, recordings, HTML documents (web pages), text messages and email messages, streams from meters and environments sensors, GPS tracks, click streams from Web queries, social media updates, data streams from machines’ communication or wearable computing sensors, and many others.

The data deluge possibly useful for enterprises, especially involved in SCM is driven nowadays by many factors like data created by the traditional IT devices, mobile devices, Internet and social network users, GPS systems and sensor nets.

The flood of data comes into existence with the images and videos uploaded in the www, video surveillance in enterprises and cities, medical information recordings, mobile devices of the users (phone calls, text messages, mobile ap-

plication usage, online games), smart devices (TV sets and receivers, smart private and industry buildings, electric grids), traditional devices (computers, game boxes, video games, e-books, etc.) and non traditional IT devices such as GPS navigation systems, earth data processing devices, radio-frequency identification (RFID) readers, ATMs, credit card readers, sensor nets and Internet of Things.

To each Gigabyte of the collected information the additional Petabyte meta-information is added. All this create a landscape of the emerging big data ecosystem.

The main player in the system are [11] the data devices, data collectors, data aggregators and data users, buyers. The data devices are the above stated originators of data deluge.

The main data collectors include information broker in Internet, the government using analytics services, medical institutions and their appliances, employers.

The marketers and private investigators can act as data aggregators by using the obtained data and transforming and packaging it for diverse stakeholders interested in conducting campaigns for users with high likelihood willing to get or buy a specific service.

The users and buyers of data are financial institutions, retail, phone and TV providers, media archives and so on. They also purchase data from third parties which enables more targeted marketing campaigns.

The collected data can be used for forecasting of the user behavior and suggesting and recommending services for user intended willing to pay for it.

Broadly speaking, the main types of data structures for big data are structured data (the most minor part), semi-structured data (slightly bigger ratio of big data), quasi-structured data (a big part of data) and unstructured data (building the majority of big data). The structured data type are such known from traditional databases and warehouse applications processing (OLAP, RDBMS, spreadsheets). The semi-structured data enables its relative easy parsing (e.g., XML-data files), while the quasi-structured data can be formatted only by using appropriate tools with much effort and possible inconsistencies occurring. On the other hand, the unstructured data does not possess an inherent structure.

III. BIG DATA ANALYTICS LIFECYCLE PROCESS

Conducting big data science ventures differs from approaches for projects using SQL data and Business Intelligence BI methods and tools applied for data analysis aims [3]. The approach used for big data is more explorative in its nature as well as there are additional key stakeholders involved in process stages.

The process can be easier handled by dividing it phases with clear defined milestones, involved stakeholder’s responsibilities and artifacts. It is according to widely accepted divide and conquer principle in computer science. It does not strictly mean a waterfall process with no returns to the previous phases but the process is a guideline for the development process which is iterative and incremental in its nature.

The process for big data projects can be divided in phases such as [11]:

- Data discovery,
- Data preparation,
- Model planning,
- Model execution,
- Communicating results
- Operationalizing the results.

As mentioned above, several of the phases can be performed simultaneously as a project work. The iterative process in the subsequent phases can proceed forwards or sometimes backward, dependent on the yes/no decision at the milestones, or a deeper insight of the problem not enough realized in the earlier phases. It may include not understanding of the problem domain, stakeholder requirements or insufficient data available to solve the given project problems. First we regard the process stakeholders, then we give a description of the process stages with regard to the stakeholder roles and their important activities and deliverables within the phases.

In each phase there are workflows to fulfill the project aims in each phase and artifacts to be delivered by the project stakeholders.

The key stakeholders roles involved in big data projects [4] and [11] are the:

- Project sponsor,
- Business user,
- Project manager,
- Business intelligence analyst,
- Database administrator,
- Data engineer
- Data scientist.

The project sponsor sets priorities and metrics for the projects and establish the desired project outcomes. The business users as the experts in the business domain can provide guidance to the project requirements recovery and operationalizing the results of the big data analytics and also are those user who directly benefits from the results of the big data project. The project manager ensures the proper project scheduling according to key project's objectives and milestones. The business intelligence analyst is responsible for creating dashboards and reports and have knowledge in proper data feeds and sources. The database administrator is responsible for provision and configuration of the database environment in order to support the analytics needs of the project; he enables the access to the needed databases and data sets and ensures the desired security levels of the data repositories.

The first five roles are good known from the usual software projects, however the roles of data engineer and data scientist are new. The role of the data engineer is needed in context of the usage of analytic sandbox workspace and fulfills the needs of data preparation for data manipulation (extracts, transformation and loading). The data scientist provides expertise needed for accurate application of analytical techniques and enables the right choices for given business problems. We already explained in detail the skills and

background of such roles as data engineer and data scientist in our paper [4].

The six stage process for big data analytics project starts with the first, the discovery phase. In this stage where the project team became acquainted with the business domain and accesses the resources needed for conducting the project. The team should familiarize itself with the project data. The accomplishment manner of this phase is dependent of the team experience in the past in similar projects. The main deliverables in this phase include framing of the business problem as an analytics challenge and formulating the initial hypotheses for data analytics.

In the second, the data preparation phase, the team prepares the analytics sandbox with the data required and conducts the ETL operations (Extract, Load, Transform, Load) on the data in the sandbox workspace [11]. In this phase the team further familiarize with the available data and if required have to decide how to obtain the required, but at the moment not available data. In this phase the usage of technologies like Hadoop [12] may be needed.

In phase three, the model planning phase, the project team needs to decide the usage of the methods, techniques and workflows to be followed in the subsequent phases, especially in the next one. The dependencies between the variables have to be established, the key variables have to be chosen and the most appropriate models types have to be selected.

In the next, the model building phase, the models recommended in the previous phase have to be developed and executed together with the appropriate data sets [4]. If the IT project environment will appear not to be sufficient for the project aims, the needed adaptations and hardware platform changes (parallelization or change to faster hardware) have to be fulfilled in this phase.

The fifth phase, which assignment is to communicate results consists of summarizing the project results such as key finding, quantification of business value and communicating the between the project stakeholders. In this phase the project aims success or failure will be decided according to the criteria determined in the first phase.

In the last operationalize phase, the final deliverables of the conducted project will be released in form of presentations, final reports, briefings, code, technical descriptions and documents etc. The form of the documents will be dependent on the recipient stakeholder type. The business value of the project should be conveyed to the key stakeholders. Moreover, after successfully running all the project phases the pilot project implementing the models in the production environment might be launched.

IV. PRIVACY FROM THE PERSPECTIVE OF ENTERPRISES

From the perspective of enterprises data protection in business processes can be seen from two different perspectives. The first one is considering the security and privacy of sensitive business data which belongs to the enterprise or its supply chain [13]. It could be enterprises important operational data or the processed data of the customers, which is aggregated while doing own business or won in other man-

ner like big data analytics. From this perspective, where it is enterprises who can fall prey to data leaks, it is vital for them to protect important information which is needed for successful operation or maintaining of a competitive advantage. This situation was recently addressed by several solution frameworks or software platforms which take care for the data exchange and access management in supply chains. We refer to the examples of PReSTiGE platform or Aniketos [14], which organize and manage data access rights in business processes. However, they only treat explicit data and do not provide strategies for management of big data and its analysis.

Other point of view the protection of the privacy of individuals [15], i.e. current or potential clients, employees and other actors whose data is aggregated by enterprises in the course of their business processes. From this perspective – where enterprises are seen as potential benefiter of extrinsic information – the important points are developing and staying compliant with the privacy policy of the enterprise and above all staying in accordance with law in order to make safe both the individuals as well as enterprises.

Governments and several organizations developed regulations, guidelines and proposals for dealing with personal data. These are legal laws, or frame conditions for professional work with personal data.

The European Union's Data Protective Directive is most advanced and restrictive among world's data privacy laws. It contains several points which must be followed by entrepreneurs. In the USA, the Federal Trade Commission proposed a three step framework, consisting of the demands of privacy by design, simplified choice and greater transparency. Other recent development of American administration is Privacy Bill of Rights which includes demands individual control, transparency, respect for context, security, access and accuracy, focused collection, and accountability.

In the literature there are several privacy requirements concepts defined which are proposed for the processes which deal with personally-identifiable information (PII) and are component of the legal regulations. Shown below important hallmarks for dealing with PII in enterprises identified in the literature [16], shall be regarded in the big data applications.

Authorization verifies who has the right to access the data or can use specific activities. In the field of business processes and especially big data it is not an easy task to draw a clear line about such requirements like separation and binding of duties since the data and possible operations on them can be so different in their nature that it is not possible to model every feasible use option. Nevertheless binding of duty can be used for setting additional responsibilities for the data or activity user of the business process. In the same way some actions can be excluded which may help maintaining privacy in the big data field, i.e. by lowering the chances of de-anonymization. Authentication is accountable for the verification if the current user is one of the user authorized for the data or service use. Confidentiality postulates to keep the architecture, process and the whole environment in such state that the data stays protected from all

non-authorized actors. Audit-ability is incorporated to that the process can be reviewed for keeping the privacy rights. Data integrity demands that the integrity of the original data has to be preserved after a processing failure. The data has to remain consistent, accurate and correct.

Since the demands regarding PII coming from the law regulations and internal privacy policies are very high therefore the procedures to fulfill them are correspondingly complex. One of the first questions which arise for enterprises is when the aggregated data has to be treated as personally identifiable. This is particularly complicated when big data comes in picture since with the ongoing technology development amounts of various data grow even more rapidly and cannot be thoroughly inspected as fast as they are collected. It demands elaborately designed analysis to assess if the conclusions drawn from the data will fall into the PII category. The technology advances enable easier identifying of specific person with growing amounts of data which only seem to have no correlation. In the end one can never be sure if apparently minor extension of the collected data won't enable the identification of individuals, and possession of such PII could harm the legal laws or cause negative effect on public opinion and therefore on the enterprises reputation.

Further step of this thinking is how the data can be processed that it will not fall in the PII category anymore.

In the next Section we propose the integration of big data analysis in business processes so that privacy will be regarded in everyday work of enterprises.

V. BUSINESS PROCESS MODELING PRACTICES FOR SUPPLY CHAINS ARCHITECTURES IN CLOUD

As stated in the above Section, privacy is an important part to regard in the business processes. While doing business the enterprises must adhere to legal law regulations and must comply with standards set in their own or in branch policies. It is often not apparent and not easily recognizable within standard business process if and when the stored data and usage of has to adhere to the above mentioned standards. It demands special steps within process which can help noticing potentially susceptible usage and unwilling consequences.

When thinking about the steps the first one would be to identify when the process and its data moves into the domain of PII. This could be done thanks to big data analysis, which can be seen a part of business process. Such analysis could be triggered by several stakeholders or events. Secondly if the data is identified as one which falls under the legislative of privacy, then actions must be undertaken in order to desensitize the contents and make it usable for business. Data after such processing must be replaced with the previous data.

It also must be analyzed why and how the data was aggregated to the level which does not comply with the standards and how this can be avoided in the future. It must be noted that the causes can be multifold and shall not be explicitly seen as enterprises misdemeanor, since these are often external circumstances which take the data out of balance. Such

circumstances are sourcing of data with previously not known contents, lack of overview on the available data, lack of thorough analysis of the new data in regard on relations with the already available data, changes in the permissions for the owned data, changes in the regulations and laws. As soon as the causes are known process changes must be made in order to avoid reoccurrence of similar situation in the future.

As an example let consider a web shop which plans to establish new service: delivery within one hour in a big city. The service itself will incorporate live tracking of traffic and weather situation, which will be used for delivery feasibility in the one hour time frame. To offer the service a warehouse will be established in the area which covers the service. The warehouse will be rather small, since in a big city area the costs are high. Because of the warehouse capacity restrictions, the products offered for sale will be studiously chosen. The options for opening other warehouse for better service or its expansion will be analyzed too. The goods shall bring good profit per warehouse area, so they must have high turnover or have a good profit margin. New opportunities for the goods will be continuously searched, i.e. the shop will also analyze their standard delivery shop and look for articles which are often ordered together. They will also review the preferences of the area population in order to offer better article range. The goods have to fit to the transport means (lorry, car, motorbike, cycle, etc.) which will be provided - at different cost - by external providers.

As described above the process uses big data and big data analytics at least two layers. First one is associated with the data related to transport used for the ordering process and assessing delivery time and service feasibility. Second is for optimization of product assortment at the warehouse or assessment if additional warehouse(s) would increase the profit.

First category extensively uses real-time data, needed for prompt delivery. This incorporates real time analysis of traffic data at the area of covered by the service, as well as observing current weather situation. Also the fleet of couriers is tracked at real-time, with detailed information about delivery duration. This feedback helps to choose right means of transport, e.g. a motorbike or a car. This data contains personal identifiable information – about the driver performance - as well.

Second level of big data analysis uses the data related to customers and predicts their buying preferences in order to adjust selection of the products available at the warehouse; it could show that new products or their varieties shall be offered, or that nearby area would be ideal candidate for expansion with additional transport means (i.e., e-bike) or with another warehouse. The data used for this analysis is won in multiple ways. It can be data feed in from social networks, where the engine is looking for activities of persons living in the area of warehouse delivery radius. It can look for the groups people living there belong, seek for their hobbies, music, films, sports, lifestyle, books, health interests. The machines can track and deliver the data of their click-streams, friend lists, follows, likes and tweets. Also text min-

ing of news sites and feeds, online newspapers, blogs, and public chats can be conducted in order to detect new social trends and needs. Although collecting such data may provide high-grade information about the needs of the inhabitants of the area, the risks for harming the privacy rights is clearly recognizable even for a layman. At the same time even for professionals it is very hard to interpret if privacy rights are harmed with ongoing data aggregation.

This shows that in both cases - transport and order data analysis, and population/trend analysis – it is not possible to assess in the real time, which information can be aggregated and stored, and when the thin border between preserving and breaking privacy is crossed. There is a need for deeper privacy rights compliance analysis for the aggregated data, which could be compared to the actual business analysis looking for business process optimization. In ideal case such privacy analysis should be done before the business analysis is conducted.

VI. CONCLUSION

Within the emerging big data ecosystem there are new groups of players, and also key roles for stakeholders. Big data analytics lifecycle is more exploratory in the nature as conventional processes, and requires a new approach. The process includes six main stages as: data discovery, data preparation, model planning, model execution, communicating results and operationalizing the results. The process itself is iterative and recurring while few phases can be carried out simultaneously as a project work.

The aim of this paper is to present the impact of big data analytics on business process modeling practices for supply chains architectures in which the modeling of privacy and security aspects play a significant role for the businesses, as they have to hold on to privacy laws like the European Data Protection Directive. The global players, as well as smaller businesses using big data, must be thoughtful about their data aggregation and analytics practices, in order to hold to those regulations. Some examples are capabilities of big data analytics which may interfere with privacy rights, like re-identification of (sensitive) data owners, profiling, amassing granular information about a person, and other uses, which go beyond the purpose and use restrictions, or the requirement of data minimization, data security, etc. Other open questions and discussions are the use of derived information based on personal data, as well as empowerment of consumers to manage their data. Supply chains are using nowadays huge amounts of data available in batch-time, online, as well as other diverse data.

As stated in [4] the data analytics involves descriptive analytics, predictive analytics and prescriptive analytics. Business Intelligence methods and tools using structured data, manageable data sets and traditional data sources can be utilized for diverse queries and providing answers for common questions of what happened in the past and why it did. Going beyond structured data requires usage of data science methods for conducting predictive analytics and data mining techniques for optimization, predictive modeling and forecasting. It will not only foster resolving reporting questions,

but also forecasting of what and why will happen. Moreover it may also support the operationalization of the key outputs of the analytic process.

Considering the data analytics process for big data described in Section 2, the sharing of the results within organization is conducted in the phase 6 (operationalization of analysis results), which is aimed at effective passing the outcomes of the analysis to the stakeholders, who are responsible to address them with appropriate actions and changes in the existing business process, so that the proposed changes and customizations can be efficiently integrated. The results of the analysis in form of describing and reporting change instructions and proposals for the business intelligence analysts will have a technical character and be in form of technical graphs like density plots, histograms etc.

If the data analysts shall detect that privacy rights may be harmed, then the suspect data sets shall be removed from the analysis. One must realize that removing parts of data will lead to results with lesser granularity but this is a price which must be paid to stay on the safe side and compliant with the privacy rights.

In the future we will investigate how big data analytics of privacy aspects could influence the established business process modeling methods, models and tools (e.g. BPMN [17]).

REFERENCES

- [1] H. Baumgarten, *Das beste der Logistik*. Springer Verlag, Berlin 2008, <http://dx.doi.org/10.1007/978-3-540-78405-0>.
- [2] S. Robak, B. Franczyk, and M. Robak, *Applying Linked Data concepts in BPM*, FedCIS 2012, IT4L. IEEE Conference Publications pp. 1105-1110.
- [3] S. Robak, B. Franczyk, and M. Robak, *Applying Linked Data concepts in Supply Chains Management*, FedCIS 2013 IEEE Conference Publications pp. 1215-1221.
- [4] S. Robak, B. Franczyk, and M. Robak, Research Problems Associated with Big Data Utilization in Logistics and Supply Chains Design and Management. FedCSIS 2014. Warsaw: Polish Information Processing Society, 2014. Annals of Computer Science and Information Systems, Vol. 3, pp. 92-93, <http://dx.doi.org/10.15439/2014F472..>
- [5] M. P. Papazoglou, and P. M. A Ribbes, *E-business: organizational and technical foundations*, John Wiley and sons. London 2006, pp.88-90.
- [6] D. Agrawal, S. Das and A. E. Abbadi, *Big data and cloud computing: current state and future opportunities*. EDBT 2011, March 22-24, 2011, Uppsala, Sweden. ACM 978-1-4503-0528-0/11/0003, <http://dx.doi.org/10.1145/1951365.1951432>.
- [7] E. Dumbill, *What is big data? An introduction to the big data landscape*, Strata O'Reilly, 11 January 2012, <http://strata.oreilly.com/2012/01/what-is-big-data.html>
- [8] S. Wrobel, *Big Data – Vorsprung durch Wissen*, Fraunhofer-Institut für Intelligente Analyse- und Informationsverarbeitungssysteme IAIS. Presentation, www.iais.fraunhofer.de
- [9] I. Mitchell and M. Wilson, *Linked Data. Connecting and exploiting big data*, Fujitsu Services Limited, March 2012, www.fujitsu.com.uk.
- [10] N. Marz and J. Warren, *Big data. Principles and best practices of scalable realtime data systems*. Manning Publications, MEAP Edition, Manning Early Access Program Big Data version 7, 2012.
- [11] Data Science and Big Data Analytics. Discovering, Analysing, Visualizing and Presenting Data. EMC Education Services – Ed. Wiley 2014.
- [12] The Apache Hadoop Project. <http://hadoop.apache.org/core/>, 2009.
- [13] Schwarzbach, B., Glockner, M., Pirogov, A., Rohling, M. M., & Franczyk, B., *Secure service interaction for collaborative business processes in the inter-cloud*. In Computer Science and Information Systems (FedCSIS), 2015 Federated Conference pp.1377-1386), <http://dx.doi.org/10.15439/2015F282>.
- [14] Brucker, A. D., Malmignati, F., Merabti, M., Shi, Q., & Zhou, B. *A framework for secure service composition*. In Social Computing (SocialCom), 2013 Int. Conference pp. 647-652, <http://doi.ieeecomputersociety.org/10.1109/SocialCom.2013.97>.
- [15] Pearson, S., *Taking account of privacy when designing cloud computing services* In Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing pp. 44-52, <http://doi.ieeecomputersociety.org/10.1109/CLOUD.2009.5071532>.
- [16] Mülle, J., Von Stackelberg, S., & Böhm, K. *Modelling and transforming security constraints in privacy-aware business processes*. In Service-Oriented Computing and Applications (SOCA), 2011 IEEE International Conference pp. 1-4, <http://dx.doi.org/10.1109/SOCA.2011.6166257>.
- [17] OMG Consortium, *Documents Associated with Business Process Model and Notation (BPMN) Version 2.0*, Release date: January 2011, www.omg.org/spec/BPMN/2.