Sharing Agreements and Quality attributes in Data Manufacturing

Konstantina Spanaki
(K.Spanaki@lboro.ac.uk)
School of Business and Economics,
Loughborough University,
Loughborough, UK

Erisa Karafili
(e.karafili@imperial.ac.uk)
Department of Computing,
Imperial College London,
South Kensington Campus, UK

Emil C. Lupu
(e.c.lupu@imperial.ac.uk)
Department of Computing,
Imperial College London,
South Kensington Campus, UK

Abstract

Data utilized in organizations as a resource, can provide multiple advantages for value creation and innovation. Nowadays, value creation is focusing on the data extraction, generation, collection and exploitation strategies formed around data as a raw material. In our study, we suggest argumentation and abductive reasoning for data processing while we extend the theoretical background of data manufacturing. The aim of the study is to provide a parsimonious approach for data processing with a focus on the quality attributes, to propose a research agenda for data value creation and new strategies around data used as a resource for competitive advantage.

Keywords: data sharing agreements (DSAs); data quality; data manufacturing; argumentation reasoning

Introduction

Contemporary advancements in technology coupled with the "Big Data" evolution provided access to enormous sets of data in an unprecedented way, as well as opportunities to innovate through the generation, collection and transformation of this data to actionable knowledge (Delen and Demirkan, 2013). The "Big Data" evolution has altered also the way firms compete and work for value creation (George et al., 2014; Lycett, 2013), and has reshaped the manufacturing landscape toward value co-creation among firms. New data-intensive ways of creating value are introduced and acquire additionally to analytical tools and techniques, sense making skills and that is the main challenge of organizations for data-based decisions. Nowadays, data can be considered as a raw material, a resource that can trigger supply chains (Data Supply Chains) and create innovative opportunities for organizations (Spanaki et al., 2017, 2016), therefore new ways of data processing are required, but also innovative strategies around this 'raw material' focusing on the quality aspects of the data (Karafili et al., 2018).

Literature reviews about the manufacturing context reveal that the data evolution has influenced radically also the production and service industries (Mishra et al., 2016; Samuel Fosso Wamba et al., 2015; Wang et al., 2016). The last decade there is a hype of service industries around data (for storing, analysing, processing etc.), where data can be shared as well as stored, used and transformed, not only by users but also by organizations, public sector and third parties. Data processing challenges are not new, however previous eras of data management were focusing solely on the quality of data of a single database and data warehousing techniques within the borders of a single organization or database (Hazen et al., 2014). What is new is the expansive way the data processing works; across the immediate boundaries of the firm, the industrial sector or even the country context.

The proposed approach extends the data manufacturing analogy and focuses on data processing by various entities and the associated data sharing agreements (DSAs). The study proposes an expressive policy analysis language for representing the DSAs through argumentation reasoning, enriched with data quality attributes to capture specified aspects of the data. The introduced analysis permits the construction of correct and efficient DSAs that can be applied in different contexts during the processing of data.

Theoretical Background

Data Manufacturing Analogy

Recent studies of Jones- Farmer et al. (2014) and Hazen et al (2014) have revisited the data manufacturing analogy and re-initiated the discussion about data quality within the context of ''Big Data''. Data manufacturing analogy was re-introduced to address the need for continuous improvement of data production throughout the SCM processes (Hazen et al., 2014), suggesting a framework for controlling the quality of the datasets in Data Analytics projects. In the aforementioned studies, data manufacturing process is investigated through a data quality lens and expands the areas around ''Big Data'' to a new research agenda of topics related to the supply chains of data.

The analogy of the data production process to a manufacturing process is prevalent in data management studies (Karafili et al., 2018; Spanaki et al., 2017). There are some significant differences between the data production process and the manufacturing process, but also noticeable similarities between these two (Karafili et al., 2018).

- The manufacturing process involves the processing of raw materials as inputs to that process, the materials are refined, and the result of this process is the output in the form of a manufactured product.
- The data manufacturing (production process) involves the data as raw materials and the input which triggers the process, the data is refined in a data product as the output of this process.

The main difference of these two processes is around the quality of data. Data of bad quality used through the data manufacturing process should be cleaned, and tailored until the quality is improved and the data product qualifies to be sold or marketed (Hazen et al., 2014; Jones-Farmer et al., 2014; Yeganeh et al., 2014).

The data manufacturing framing presented in data management literature of the previous decades could be applied in the context of data processing and production. The major concern though, is the fact that it should be extended in a boundary-less context between firms and the target should be innovative outputs, as these are mostly the results of the data evolution (Spanaki et al., 2017). Data manufacturing analogy was based on the data artefact as a unit of analysis; however, data era requires novel techniques focusing on the processing of data but not solely on their processing mechanism, but also the quality and sharing attributes associated with them (Karafili et al., 2018). Data sharing control was not the major focus of data management in previous decades; as the data were mostly shared within the boundaries of a company or between single databases. Dealing with problems of a sole database or a single company, we imply that the trust and security issues were solved by the individual sharing entities and the associated agreements between interested parties.

Based on the analysis above, we claim that data processing revisits the analogy of data manufacturing to product manufacturing, while it expands this view to represent a contemporary concept of the so called “Big Data” evolution. Data manufacturing is a distinct type of manufacturing process (or even a sub-category) depicting a transformational process of data to valuable outcomes. This view implies a dynamic interplay of raw materials (data), actors, processes, integration and complex business structures.

Data in Supply Chain Management

Nowadays, firms from multiple industries invest on data-driven decision-making and business analytics solutions, for improving their performance and operations (Akter et al., 2016; Feki et al., 2016; S F Wamba et al., 2015). The use of data in supply chain management as well as raw materials in production and service improvement is presented the last decade as a solution for enterprises to create value by using their analytical skills in visualising, optimising and transforming their supply chains (Hazen et al., 2016; Waller and Fawcett, 2013). Although data can be used along with the core business focus in different industries, we can observe that the recent data evolution expanded the business scope and disrupted the operating models providing opportunities to work solely on data as the main “raw material”, processing this data, creating new products and services and also reselling and exchanging data (Spanaki et al., 2017, 2016).

The topic of management within logistics and supply chain literature has a strong focus on the Supply Chain Analytics (SCA) as a strategic way in developing supply chain efficiencies at tactical and operational levels (Wang et al., 2016). Research studies following this direction focus on how the analytics can be applied to strategic decisions related to SCM (Gunasekaran et al., 2017; Wang et al., 2016; Zhong et al., 2015), how efficiency and effectiveness of supply chains can be improved through the use of data (Ilie-Zudor et al., 2015; Waller and Fawcett, 2013) as well as the data strategies and servitization around supply chains (Baines et al., 2017; Opresnik and Taisch, 2015). The current direction of Supply Chain Management (SCM) reflects that the research analyzes mostly the use of data in supply chains; nevertheless our research focus will be on the supply chains of data, setting data as a raw material/resource triggering a manufacturing process, analogous to the product/service manufacturing processes (Spanaki et al., 2017, 2016).

Data Management and Quality

Literature reviews and frameworks referring to data/ information as raw materials and the data manufacturing analogy are presented in various studies of data management in of previous decades (1980-2000). The data manufacturing analogy presents that data raw materials analogous to physical raw materials are moved through a manufacturing process which reshapes/reconfigures them in information/data products (Table 1).
Data as a "raw material" was initially introduced by Brodie (1980) through the analogy between product manufacturing and data manufacturing process when data quality was a major concern in transforming data to valid information and knowledge (Arnold, 1992; Fox et al., 1994; Redman and Blanton, 1997). Some of the most indicative studies around these areas developed the concept of data manufacturing analogy in order to find out the path for better data quality (Huh et al., 1990; March and Hevner, 2007; Redman, 1998; Wang, 1998) and they designed frameworks that describe and track data manufacturing processes (Ballou et al., 1998; Ronen and Spiegler, 1991; Wang, 1998; Wang et al., 1995). A simple framework of *input-process- output* describing the similarities between the two manufacturing processes is proposed by Wang et al. (1995) and calls for continuously defining, measuring, analysing, and improving data quality. Mostly, the data manufacturing analogy was focusing on data quality and the ways to ensure that we can trust the data we use in manufacturing processes.

Product vs. Data/Information Manufacturing		
	Product Manufacturing	**Data/Information Manufacturing**
Input	Raw Materials	Raw Data (or Data Products)
Process	Materials Processing (Assembly Line)	Data processing (via Information Systems and Analytical Processes)
Output	Physical Products	Data/ Information Products

Table 1 –Data Manufacturing Analogy (Karafili et al., 2018; Wang, 1998; Wang et al., 1995)

Data Access and Usage Control

The research focus of data sharing studies usually is around the protection of the data itself (for used and shared data), as well as the databases where they are stored (Gertz and Jajodia, 2008) or the network used for the transfer of this data (Kaufman, 2009). Due to the increasing connectivity between users, there is a parallel increase of the associated security breaches and attacks. Protecting and securing the environment where the data can transferred/stored/used or even re-used was always and is still remaining a major challenge for all the interested parties (Karafili et al., 2015). Data-centric security solutions have dominant position in the literature (Bayuk, 2009; Kim et al., 2010; Pearson and Casassa-Mont, 2011; Wang et al., 2010; Zhou et al., 2010) and specifically the data protection of the data transfers and transactions. Data-centric security solutions present two main challenges associated with the access and the usage control of the data. Both of them, have been widely studied and the research has developed multiple solutions for solving such problems (Ferraiolo et al., 1992; Lazouski et al., 2012). The existing solutions will be presented briefly at this section to distinguish the proposed approach from previous research paths.

Role-based access control according to Ferraiolo and Kuhn (Ferraiolo et al., 1992) can be presented based on specifying the user roles in the data access controls. Motivated by this direction, we expand this context with a data representation technique for specified user roles with defined usage access and policies for the data. Usage control (UCON) as defined by Park and Sandhu (Park and Sandhu, 2004) is a widely studied concept following different approaches, some of them represent the UCON for controlling the access and usage of digital information emphasizing on the problem of rights delegation (Park and Sandhu, 2002).

An additional approach for sharing and accessing data is the use of sticky policies (Pearson and Casassa-Mont, 2011). Sticky polices are machine readable policies that contain conditions and constraints attached to data that describe how the data should be treated while shared among multiple parties. The sticky policy paradigm, technologies for enterprise privacy enforcement and the exchange of customer data are represented through a privacy control language for specified privacy rights and obligations (Karjoth and Schunter, 2002). The privacy control language presents authorization management and access control for user consents, obligations and distributed administration, with extension of the sticky policy paradigm also for the cloud environment (Trabelsi and Sendor, 2012).

Before creating, sharing and using the data, the data subject, controller and processor should agree regarding the different rules that describe how the data should be treated, that are called data sharing agreements (Swarup et al., 2006). The DSAs describe not only the agreements between the data subject, controller, and processor, but also the compliance of the different business and regulatory contexts for data sharing. A language representation of different rules for data sharing agreements (DSAs) as presented by Matteucci et al. (2010) fails to provide expressivity. This language cannot permit the representation of complex DSAs, as well as analysis for the DSAs and leaves unsolved the problem of deciding which rules to apply to the DSAs.

All the above represented approaches, from the data access and usage control, to the sticky policies and finally the DSAs representation, seem incomplete to provide a decision background for the rules that should apply to the shared data. Following this motivation, we propose a combinatory analysis of the rules with a conflict resolution technique. The proposed analysis is based on abductive (Kakas et al., 1992) and argumentation based reasoning (Bondarenko et al., 1997; Dung, 1995), as this technique can facilitate decision making mechanisms under conflicting knowledge (Bandara et al., 2009; Kakas and Moraitis, 2003).

Methodological Approach

The approach presented in this study assumes that data are processed by different entities. Data Sharing Agreements (DSAs) are established between the entities for data processing composed of various constraints and rules. An expressive policy analysis language is used in this case for representing the DSAs. Data quality is the main focus of the data processing mechanism; therefore the policy language is enriched to capture various data quality properties like accessibility, timeliness and accuracy. The used policy language permits the analysis of the various policies and the detection of the rising conflicts, redundancies or the missing cases. We follow argumentation and abductive reasoning to build our proposed method to capture and solve conflicts between context dependent rules. The introduced analysis permits the construction of precise DSAs that can be applied in various contexts during data processing phase.

General Representation Context

We introduce here the general representation context, where we put a background case of collecting, processing and sharing data. We show in the coming section how our methodology for Data Sharing Agreements is used in this scenario for naturally representing the needs of the various actors involved, and effectively deciding who can access to what part of the data.

The main actor of our scenario has a data-related activity, and decides to start collecting data around this specific activity. The collection of the data can be made in different ways: manually, through IoT devices, drones etc. The way the data is collected can influence directly the accuracy aspect. The collection of data can occur in different intervals within the day/time, e.g., every hour/day/month, always. The collection time/period influences the freshness of the data. This freshness is strictly related also to the type of data.

- The owner of the data-related activity is called the **data owner**, and in this case s/he is also the **data subject**, as part of his data are collected as well, e.g., name of the activity, his personal data, and the **data controller** as s/he determine the purpose and means of processing the data.
- The **data recipients** are all actors that want to access/use/share the data owner's data, the stakeholders involved in the activity or similar activities. Data recipients need to comply with the data controller rules.
- Sometimes the owner of the activity relies on *third parties* that provide the technical support for collecting the data. In this case owner is the **data owner/subject** but he can delegate the control of his/her activity data to a third party that is now the **data controller**.
- The **data processor** is an entity (public authority, agency, legal person) that is processing the data on behalf of the controller. In our use case, the collected/processed/shared data can be stored in the cloud. Thus the cloud provider is considered the data processor as far as it respects the instructions of the controller. The controller rules that should be respected by the processor can also have a legal nature, e.g., if the controller is in an EU country, the cloud provider should as well be in an EU country and cannot send the data to countries outside the EU and EEA.
- A **third party** is an entity (public authority, agency, legal person) that is not the data subject, data controller or processor, and that under the direct authority of the controller or processor is authorized to process the data. In our case, a company that is outside of the immediate boundaries of the data owner which is granted access can be considered a third party. Once access is obtained, the third party becomes a data controller and has to comply with the data protection principals.

The Policy Language

The proposed model is based on the policy analysis language (Craven et al., 2009) that is constructed using the Event Calculus. This language represents the required rules and constraints for accessing, using and sharing the data. The policy regulation rules are composed of predicates and domain descriptions, and represent the authorization and obligation rules. The authorization rules have in their structure a specific **subject**, as well as specified **targets**, and **actions**. A brief introduction to the language with some of the main predicate is given below.

$$req(Sub, Tar, Act, T) \qquad obl(Sub, Tar, Act, T_s, T_e, T)$$
$$permitted(Sub, Tar, Act, T) \qquad denied(Sub, Tar, Act, T)$$

The above predicate represents correspondingly: a request made from the subject *Sub*, at the instant of time *T*, to perform a certain action *Act* at the target *Tar*; the obligation for a given subject to perform an action during the period of time from T_s to T_e; that a given subject is permitted/denied at time *T*, to perform a certain action to the target. A domain description predicate is *holdsAt*, which means that a given property/predicate is true in a given instance of time.
The used policy language can represent the permission, denial and obligation concepts for the DSAs. In the following examples, we introduce the representation of some DSAs rules from our use case.
Example 1. For example if Mary that is the owner of the data and she wants to access them. In this case, she is allowed to access her own data. We use the owner property for stating that a given actor is the owner of the data.

$$Permitted(Mary, Data, access, T) \leftarrow holdsAt(owner(Mary, Data), T).$$

Example 2. A third party TP that needs access to Mary's data is required to delete them after 5 years of access to this data. We express the time in moths in this case (60 months). The obligation starts from the moment the access is granted, and they have time T_0 until 60 months after the access is granted to delete them.

$$\begin{aligned} obl(TP, Data, delete, T, T_0, T) \leftarrow\ & permitted(TP, Data, access, T), \\ & holdsAt(owner(Mary, Data), T), \\ & T < T_0 < T + 60. \end{aligned}$$

Representation of the Quality and Sharing Aspects
The sharing and usage of data raises issues around the description of other properties related to the quality of the collected data. When working with data quality the entities that are using, sharing, storing the data are called **data consumers**. In our case, the data consumers are considered the data collectors and the data processors. The data quality is an important factor when we are working with data consumers, where **data quality** is defined as data that fit the requirements for being used by data consumers (Wang and Strong, 1996).
Accessibility: Also another important aspect of data quality is **accessibility**. Our DSA representation together with the used policies set data accessibility permissions. The used methodology ensures the data accessibility complies to data owner's constraints, and other legal internal and external frameworks and regulations.
Accuracy: An important data quality attribute which should be met is the **accuracy** of the collected data. When data are collected by an human actor, we can put in act a deontic obligation for the actor to use a particular accuracy level when collecting and storing the data. This case can suffer from human errors. On the other hand, when a device is collecting the data, e.g., an IoT device, we can be more specific and ensure the data accuracy by checking the parameters of the various devices.

$$\begin{aligned} accuracy(Data, T) \leftarrow\ & holdsAt(meas(Device, Data), T), \\ & holdsAt(acceptP(Device), T). \end{aligned}$$

In the above case, we state that the data collected are accurate, as the device that collected them is able to measure the data with acceptable parameters, (*acceptP*), for ensuring data accuracy.
Freshness: Another important data quality characteristic is data **freshness** which is the degree data represent reality in the required point in time. The notion of data freshness

is part of the **timeliness** as a data quality aspect. For our scenario we will work only on the data freshness predicate, as it represents the data timeliness depending on the different contexts. Data freshness is a predicate that expresses the last time when a given piece of data has been updated:

$$freshness(Tar, T)$$

where Tar is the targeted data, and T is the last instant of time when the data were updated. Let us see how we can apply the above predicate to our use case.

Example 3. In our scenario, the data are fresh if the measurements (eas) are made in particular time difference from one to the other, e.g., X, and the previous measurement respected the freshness as well. The value X depends by the type of data and the data consumer.

$$\begin{aligned} freshness(Data, T) \leftarrow\ & holdsAt(meas(Device, Data), T), \\ & holdsAt(meas(Device, Data), T_0), \\ & freshness(Data, T_0), \\ & T = T_0 + X. \end{aligned}$$

Conclusion

The purpose of this research is to extend the ‘data manufacturing’ concept of previous decades, to a data-intensive environment across organisational, individual and country boundaries, where data are accessible by different entities. By using the data processing approach, we unfold the potential areas of interest of data manufacturing for multiple entities. Within this context, we argue that data can be processed and can create value through tailoring techniques across organisational boundaries with the help of DSAs and usage control rules. Value creation could also be realized while developing data products/services as well as disrupting the manufacturing landscape to facilitate such a change.

Acknowledgements

Supported by FP7 EU-funded project Coco Cloud grant no.: 610853, and EPSRC Project CIPART grant no. EP/L022729/1. Erisa Karafili was supported by the EU’s H2020 under the Marie Skłodowska-Curie grant agreement No 746667.

References

Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R., Childe, S.J., 2016. How to improve firm performance using big data analytics capability and business strategy alignment? International Journal of Production Economics 182, 113–131. https://doi.org/10.1016/j.ijpe.2016.08.018

Arnold, S.E., 1992. Information manufacturing: the road to database quality. Database 15, 32–39.

Baines, T., Ziaee Bigdeli, A., Bustinza, O.F., Shi, V.G., Baldwin, J., Ridgway, K., 2017. Servitization: revisiting the state-of-the-art and research priorities. International Journal of Operations & Production Management 37, 256–278. https://doi.org/10.1108/IJOPM-06-2015-0312

Ballou, D., Wang, R., Pazer, H., Tayi, G.K., 1998. Modeling information manufacturing systems to determine information product quality. Management Science 44, 462–484.

Bandara, A.K., Kakas, A.C., Lupu, E.C., Russo, A., 2009. Using argumentation logic for firewall configuration management, in: 2009 IFIP/IEEE International Symposium on Integrated Network Management, IM 2009. pp. 180–187. https://doi.org/10.1109/INM.2009.5188808

Bayuk, J., 2009. Data-centric security. Computer Fraud and Security 2009, 7–11. https://doi.org/10.1016/S1361-3723(09)70032-6

Bondarenko, A., Dung, P.M., Kowalski, R.A., Toni, F., 1997. An abstract, argumentation-theoretic approach to default reasoning. Artificial Intelligence 93, 63–101. https://doi.org/10.1016/S0004-3702(97)00015-5

Brodie, M.L., 1980. Data quality in information systems. Information & Management 3, 245–258.

Craven, R., Lobo, J., Ma, J., Russo, A., Lupu, E., Bandara, A., 2009. Expressive policy analysis with enhanced system dynamicity, in: Proceedings of the 4th International Symposium on Information, Computer, and Communications Security - ASIACCS ’09. p. 239. https://doi.org/10.1145/1533057.1533091

Delen, D., Demirkan, H., 2013. Data, information and analytics as services. Decision Support Systems 55, 359–363.
Dung, P.M., 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence 77, 321–357. https://doi.org/10.1016/0004-3702(94)00041-X
Feki, M., Boughzala, I., Wamba, S.F., 2016. Big Data Analytics-Enabled Supply Chain Transformation: A Literature Review, in: Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS). IEEE Computer Society, pp. 1123–1132.
Ferraiolo, D., Barkley, J., Kuhn, R., 1992. Role-Based Access Controls. ACM Transactions on Information and System Security 2, 34–64. https://doi.org/10.1145/300830.300834
Fox, C., Levitin, A., Redman, T., 1994. The notion of data and its quality dimensions. Information processing & management 30, 9–19.
George, G., Haas, M.R., Pentland, A., 2014. Big data and management. Academy of Management Journal 57, 321–326.
Gertz, M., Jajodia, S., 2008. Handbook of database security: Applications and trends, Handbook of Database Security: Applications and Trends. https://doi.org/10.1007/978-0-387-48533-1
Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S.F., Childe, S.J., Hazen, B., Akter, S., 2017. Big data and predictive analytics for supply chain and organizational performance. Journal of Business Research 70, 308–317. https://doi.org/10.1016/j.jbusres.2016.08.004
Hazen, B.T., Boone, C.A., Ezell, J.D., Jones-Farmer, L.A., 2014. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. International Journal of Production Economics 154, 72–80. https://doi.org/10.1016/j.ijpe.2014.04.018
Hazen, B.T., Skipper, J.B., Ezell, J.D., Boone, C.A., 2016. Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. Computers & Industrial Engineering 101, 592–598. https://doi.org/http://dx.doi.org/10.1016/j.cie.2016.06.030
Huh, Y., Keller, F., Redman, T., Watkins, A., 1990. Data quality. Information and Software Technology 32, 559–565. https://doi.org/10.1016/0950-5849(90)90146-I
Ilie-Zudor, E., Ekárt, A., Kemeny, Z., Buckingham, C., Welch, P., Monostori, L., 2015. Advanced predictive-analysis-based decision support for collaborative logistics networks. Supply Chain Management 20, 369–388. https://doi.org/10.1108/SCM-10-2014-0323
Jones-Farmer, L., Ezell, J.D., Hazen, B.T., 2014. Applying Control Chart Methods to Enhance Data Quality. Technometrics 56, 29–41. https://doi.org/10.1080/00401706.2013.804437
Kakas, A., Moraitis, P., 2003. Argumentation based decision making for autonomous agents. Proceedings of the second international joint conference on Autonomous Agents and Multiagent Systems 883–890. https://doi.org/10.1145/860575.860717
Kakas, A.C., Kowalski, R.A., Toni, F., 1992. Abductive logic programming. Journal of Logic and Computation 2, 719–770. https://doi.org/10.1093/logcom/2.6.719
Karafili, E., Nielson, H.R., Nielson, F., 2015. How to trust the re-use of data, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 72–88. https://doi.org/10.1007/978-3-319-24858-5_5
Karafili, E., Spanaki, K., Lupu, E.C., 2018. An argumentation reasoning approach for data processing. Computers in Industry 94, 52–61. https://doi.org/10.1016/j.compind.2017.09.002
Karjoth, G., Schunter, M., 2002. A privacy policy model for enterprises, in: Proceedings 15th IEEE Computer Security Foundations Workshop. CSFW-15. pp. 271–281. https://doi.org/10.1109/CSFW.2002.1021821
Kaufman, L.M., 2009. Data security in the world of cloud computing. IEEE Security and Privacy 7, 61–64. https://doi.org/10.1109/MSP.2009.87
Kim, Y.J., Thottan, M., Kolesnikov, V., Lee, W., 2010. A secure decentralized data-centric information infrastructure for smart grid. IEEE Communications Magazine 48, 58–65. https://doi.org/10.1109/MCOM.2010.5621968
Lazouski, A., Martinelli, F., Mori, P., 2012. A prototype for enforcing usage control policies based on XACML, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 79–92. https://doi.org/10.1007/978-3-642-32287-7_7
Lycett, M., 2013. "Datafication": making sense of (big) data in a complex world. European Journal of Information Systems 22, 381–386. https://doi.org/10.1057/ejis.2013.10
March, S.T., Hevner, A.R., 2007. Integrated decision support systems: A data warehousing perspective. Decision Support Systems 43, 1031–1043. https://doi.org/10.1016/j.dss.2005.05.029
Matteucci, I., Petrocchi, M., Sbodio, M.L., 2010. CNL4DSA: A controlled natural language for data sharing agreements, in: Proceedings of the ACM Symposium on Applied Computing. https://doi.org/10.1145/1774088.1774218
Mishra, D., Gunasekaran, A., Papadopoulos, T., Childe, S.J., 2016. Big Data and supply chain management: a review and bibliometric analysis. Annals of Operations Research 1–24.
Opresnik, D., Taisch, M., 2015. The value of big data in servitization. International Journal of Production Economics 165, 174–184. https://doi.org/10.1016/j.ijpe.2014.12.036
Park, J., Sandhu, R., 2004. The UCONABC Usage Control Model. ACM Trans. Inf. Syst. Secur. 7, 128–174. https://doi.org/10.1145/984334.984339
Park, J., Sandhu, R., 2002. Towards usage control models: beyond traditional access control, in: Proceedings of the

Seventh ACM Symposium on Access Control Models and Technologies - SACMAT ’02. p. 57. https://doi.org/10.1145/507721.507722
Pearson, S., Casassa-Mont, M., 2011. Sticky policies: An approach for managing privacy across multiple parties. Computer 44, 60–68. https://doi.org/10.1109/MC.2011.225
Redman, T.C., 1998. The impact of poor data quality on the typical enterprise. Communications of the ACM 41, 79–82.
Redman, T.C., Blanton, A., 1997. Data quality for the information age. Artech House, Inc.
Ronen, B., Spiegler, I., 1991. Information as inventory. A new conceptual view. Information and Management 21, 239–247. https://doi.org/10.1016/0378-7206(91)90069-E
Spanaki, K., Adams, R., Mulligan, C., Lupu, E., 2016. Data Supply Chain (DSC): development and validation of a measurement instrument, in: 23rd EurOMA Conference. Trondheim Norway.
Spanaki, K., Gürgüç, Z., Adams, R., Mulligan, C., 2017. Data supply chain (DSC): research synthesis and future directions. International Journal of Production Research. https://doi.org/10.1080/00207543.2017.1399222
Swarup, V., Seligman, L., Rosenthal, A., 2006. Specifying data sharing agreements, in: Proceedings - Seventh IEEE International Workshop on Policies for Distributed Systems and Networks, Policy 2006. pp. 157–160. https://doi.org/10.1109/POLICY.2006.34
Trabelsi, S., Sendor, J., 2012. Sticky policies for data control in the cloud, in: 2012 10th Annual International Conference on Privacy, Security and Trust, PST 2012. pp. 75–80. https://doi.org/10.1109/PST.2012.6297922
Waller, M.A., Fawcett, S.E., 2013. Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain. Journal of Business Logistics 34, 249–252. https://doi.org/10.1111/jbl.12024
Wamba, S.F., Akter, S., Coltman, T., Ngai, E.W.T., 2015. Guest editorial: Information technology-enabled supply chain management. Production Planning and Control 26, 933–944. https://doi.org/10.1080/09537287.2014.1002025
Wamba, S.F., Akter, S., Edwards, A., Chopin, G., Gnanzou, D., 2015. How “big data” can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics 165, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031
Wang, C., Wang, Q., Ren, K., Lou, W., 2010. Privacy-preserving public auditing for data storage security in cloud computing, in: Proceedings - IEEE INFOCOM. https://doi.org/10.1109/INFCOM.2010.5462173
Wang, G., Gunasekaran, A., Ngai, E.W.T., Papadopoulos, T., 2016. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. International Journal of Production Economics 176, 98–110. https://doi.org/10.1016/j.ijpe.2016.03.014
Wang, R.Y., 1998. A product perspective on total data quality management. Communications of the ACM 41, 58–65.
Wang, R.Y., Reddy, M.P., Kon, H.B., 1995. Toward quality data: An attribute-based approach. Decision Support Systems 13, 349–372.
Wang, R.Y., Strong, D.M., 1996. Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems 12, 5–33.
Yeganeh, N.K., Sadiq, S., Sharaf, M.A., 2014. A framework for data quality aware query systems. Information Systems 46, 24–44. https://doi.org/http://dx.doi.org/10.1016/j.is.2014.05.005
Zhong, R.Y., Xu, C., Chen, C., Huang, G.Q., 2015. Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors. International Journal of Production Research. https://doi.org/10.1080/00207543.2015.1086037
Zhou, W., Sherr, M., Marczak, W.R., Zhang, Z., Tao, T., Loo, B.T., Lee, I., 2010. Towards a data-centric view of cloud security, in: Proceedings of the Second International Workshop on Cloud Data Management - CloudDB ’10. p. 25. https://doi.org/10.1145/1871929.1871934