

Big Data

An Interdisciplinary Opportunity for Information Systems Research

“Big data” describes technologies that promise to fulfill a fundamental tenet of research in information systems, which is to provide the right information to the right receiver in the right volume and quality at the right time. The paper illustrates the technical, economical, and legal preconditions of this promise using the phenomenon of “big data hubris”. The authors argue that information systems research is ideally positioned to explore big data from a critical stance and to develop innovative perspectives on the design and use of information, regardless of whether big data is perceived as a disruptive technology or yet another buzzword.

DOI 10.1007/s12599-014-0345-1

The Authors

Dr. Michael Schermann (✉)
Prof. Dr. Helmut Krcmar
 Chair for Information Systems (I17)
 Technische Universität München
 Boltzmannstraße 3
 85748 Garching
 Germany
michael.schermann@in.tum.de
krcmar@in.tum.de

Dr. Holmer Hensen
Prof. Dr. Volker Markl
 Database Systems and Information
 Management Group (DIMA)
 Technische Universität Berlin
 Einsteinufer 17
 10587 Berlin
 Germany
holmer.hensen@tu-berlin.de
volker.markl@tu-berlin.de

Christoph Buchmüller
Till Bitter
Prof. Dr. Thomas Hoeren
 Institut für Informations-,
 Telekommunikations-
 und Medienrecht
 Westfälische Wilhelms-Universität
 Münster
 Leonardo-Campus 9
 48149 Münster
 Germany
christophbuchmueller@uni-muenster.de
till.bitter@uni-muenster.de
hoeren@uni-muenster.de

Received: 2014-04-01
 Accepted: 2014-05-20
 Accepted by the editors of the special
 focus.
 Published online: 2014-09-04

This article is also available in German in print and via <http://www.wirtschaftsinformatik.de>: Schermann M, Hensen H, Buchmüller C, Bitter T, Krcmar H, Markl V, Hoeren T (2014) Big Data. Eine interdisziplinäre Chance für die Wirtschaftsinformatik. WIRTSCHAFTSINFORMATIK. doi: 10.1007/s11576-014-0434-2.

© Springer Fachmedien Wiesbaden 2014

1 Introduction

The term “big data” summarizes technological developments in the area of data storage and data processing that provide the possibility to handle exponential increases in data volume presented in any type of format in steadily decreasing periods of time (Chen et al. 2012; Lycett 2013). Big data provides the opportunity to not only handle but also use and add value to large amounts of data coming from social networks, images and other information, and communication technologies (McAfee and Brynjolfsson 2012). If one believes software

companies, research projects, and first-show cases, big data promises no less than providing the right information to the right receiver in the right volume and quality at the right time (Lycett 2013; Krcmar 2009).

It comes as no surprise that big data has become one of the hyped buzzwords of the very recent past and is being discussed as the next disruptive technology (Eberspächer and Wohlmuth 2013). Nevertheless, critical voices question the novelty of big data and see it as the effort of technology providers to sell new products and services (Buhl et al. 2013).

In line with Steininger et al. (2009), we argue that exploring big data is an opportunity for information systems research. We will not attempt to give a full overview of the current debate on big data because an overview would be outdated by the time this paper is published. Instead, we explore big data using the phenomenon of big data hubris, which is defined as organizational self-esteem caused by the use of big data (Lazer et al. 2014). Based on the example of big data hubris, we discuss how big data can be an opportunity for information systems research, regardless of whether big data is thought of as a disruptive technology or yet another buzzword.

The remainder of the paper is organized as follows: we first explain the phenomenon of big data hubris and then explore the phenomenon from a technological, legal, and economic perspective. We conclude with a set of propositions

intended to support information systems research to glean the opportunities presented by big data.

2 The Phenomenon of the Big Data Hubris

Big data is commonly discussed in the context of completely new products and services as well as innovative business application systems (Eberspächer and Wohlmuth 2013; Schroeck et al. 2012). The core of these examples is quite often the significant opportunities related to processing huge data volumes in varying formats in decreasing periods of time (The Economist Intelligence Unit 2014). The underlying assumption is that as more data are processed, improved decision-making in business, administration, politics, and even in private life become possible. As an example, McAfee (2013) argues that previous decision-making processes are faulty due to the human component (i.e. decisions based on intuition or a “gut feeling”). Big data, on the other hand, provides the chance to establish data-driven decision-making processes without human interpretations.

The project Google Flu Trends, a platform for predicting flu outbreaks, provides an illustrative example of big data (Ginsberg et al. 2009). Google correlated historical data on reports of flu cases from the U.S. Center for Disease Control with its own data on frequency, time, and place of searches. Google identified about 45 keywords related to the forecasting of waves of flu with the goal of being able to predict the occurrence and spread of influenza epidemics in real time (Ginsberg et al. 2009).

In contrast, recent studies show that confidence in the ability to quickly process large quantities of heterogeneous data is premature. Again, Google Flu Trends serves as an illustrative example. Lazer et al. (2014) show that the Google flu prediction is at least partially inaccurate: Google Flu Trends overestimated the extent of influenza outbreaks by 50 percent and did not recognize non-seasonal influenza outbreaks (Lazer et al. 2014).

The attractiveness of big data tempts researchers to use big data, often implicitly, as a replacement for traditional methods of data analysis. If traditional

decision-making is replaced by big data without taking into account inherent challenges such as construct validity, reliability of data, and the context of data, an unjustified reliance on big data with a subsequent risk of incurring severe mistakes might be the end result (“big data hubris”).

Lazer et al. (2014) argue that the phenomenon of big data hubris typically arises from the well-discussed prime examples of the benefits of big data, e.g., the analysis of content from social networks whose quality is difficult to access. These examples typically come with a focus on processing large amounts of data stemming from a homogeneous data set, as was the case in the Google Flu Trends search requests. The much more difficult and complex semantic integration of different data sources is, however, often not undertaken due to the extensive effort required.

At the same time, many pilot examples are based on publicly available data from social networks (e.g. Google, Facebook or Twitter). It should be remembered that these companies continuously develop their algorithms. This raises the question to what extent data on the various versions of the algorithms can be accurately compared if at all (Lazer et al. 2014).

We argue that the effective use of big data and thus the avoidance of big data hubris requires an interdisciplinary approach that combines technological, legal, and economic perspectives. An interdisciplinary approach is required to use big data for the design of innovative information systems in business and management.

3 Big Data: An Interdisciplinary Analysis¹

At the heart of the discussion on big data is its promise to provide the right information to the right receiver in the right volume and quality at the right time (The Economist Intelligence Unit 2014; Krcmar 2009). In the following, we discuss technological, legal, and economic preconditions required to realize this promise.

3.1 Technological Preconditions: User-Friendly Transparency Through Declarative Systems for Big Data

A critical technological precondition is to provide effective systems for big data that enable reliable and effective data analysis to a wide range of users.

Big data dramatically increases the complexity of methods and procedures for analyzing data. This means we need new methods for the collection, integration, and analysis of big data. The development and use of complex statistical algorithms, machine learning methods, linear algebra, signal processing, data mining, text mining, graph mining, video mining, and visual analysis are the focus of interest. These processes go beyond the classical operations of relational algebra, which are realized in today’s database systems (Klein et al. 2013).

The main objective is to ensure that big data methods and procedures can be used not just by a few specialists but also by a wide range of users. Current big data technologies, however, require extensive knowledge of system programming or parallel programming which effectively limits the user group to specialists.

Therefore, researchers and practitioners are currently working on the development of declarative languages for the specification, optimization, and parallelization of complex data analyses that go beyond the traditional data languages such as SQL, XPath, and XQuery for XML data or SPARQL for RDF data (American National Standards Institute 1992; Berglund et al. 2010; Boag et al. 2011; Prud’hommeaux and Seaborne 2008). A declarative specification of data analysis means that the description of the analysis problem is decoupled from its execution.

For big data no declarative specification of data analysis programs exists that scales to any computer system and automatically optimize and parallelize data-independently with respect to storage and statistical properties (Agrawal et al. 2012). The specification of such a declarative language for big data requires, in addition to classic operations of relational algebra, the use of complex user-defined functions as well as the specification of iterative algorithms. With these specifications in place, methods of linear algebra, machine learning, language-, video-, and

¹The results and findings in this section are based on a study that was carried out by the authors on behalf of the Federal Ministry of Economy and Energy in 2013. This study can be accessed at: http://www.dima.tu-berlin.de/menue/forschung/big_data_management_report/.

signal processing can be integrated into the declarative description of big data analytics.

The map/reduce paradigm for big data was introduced for the specification and processing of user-defined functions on datasets or groups of datasets (Dean and Ghemawat 2004). This functional programming model consists of two second-order functions in which any first-order function could be used for transformation or selection of datasets (map) as well as for aggregation of datasets (reduce). The popular big data system Hadoop implements this programming model and automatically parallelizes analyses on large computing clusters (White 2012). These approaches, however, are just a first step towards a declarative language for big data, as they do not necessarily support automatic optimization and parallelization of complex iterative data analysis algorithms. Thus, data analysis remains restricted to a group of expert users who simultaneously possess programming knowledge as well as skills in data analysis and machine learning.

Moreover, big data is available with a wide variety of quality levels. For example, geo-positioning data have a certain imprecision due to the number of available satellites. Therefore, users must know the origin of the data and its quality in order to be able to estimate the correctness of analysis results, draw appropriate sampling, and make appropriate model assumptions. Similarly, probabilistic models are used more frequently whose limitations and boundaries need to be incorporated into the decision support.

In the application of complex analysis procedures, an iterative procedure is productive to determine the correct parameter settings. Hence, it is necessary to make the various steps of the data analysis comprehensible to the user. This could be realized, for example, with the help of interactive, often visually supported, tools (“visual analytics”). These tools support an interactive modeling of assumptions, basic conditions, and interests of analysis that are iteratively supported through appropriate partial analyses. An effective visualization can assist in performing a quick and accurate assessment of the results quality as a function of the parameter settings (Thomas and Cook 2005).

3.2 The Legal Requirement: Responsibility and Legal Certainty for Value Creation with Big Data²

The legal requirement in terms of the creation of value stemming from the use of big data is to specify effective frameworks that channel the handling of big data from a social perspective. In principle, big data is mainly discussed from the perspective of data privacy protection. Although this is highly relevant from a societal perspective, the phenomenon of big data hubris shows that copyright, ownership, and legal liability considerations must also play a central role in the use of big data.

Copyright protects intellectual creations that have an adequate level of originality (§ 2 para. 2 UrhG). However, data do not often meet the requirements specified in § 2 para. 2 of the German Copyright Act because the content is either purely factual or because the content lacks the required degree of individuality (Dreier and Schulze 2013, § 20 Rn. 130).

Nevertheless, copyright issues must be respected and taken into consideration when handling data. In addition, the so-called *sui generis* protection for databases (§§ 87a UrhG) may be applicable. Because there are no formal rules or regulations, *sui generis* protection is assumed as a means of providing protection for the investment risk of creating a database (Wandtke and Bullinger 2009, 87a Rn. 55 ff.; Dreier and Schulze 2013, § 87a Rn. 14; Gaster 1999, Rn. 476). Copyrights could conflict with the necessary data handling in the analysis of user-generated content obtained from third parties in social media. Data obtained from these types of sources could, for example, be protected under copyright applicable to literary (§ 2 para. 1 No. 1 UrhG), or photographic works (§ 2 para. 1 No. 5 UrhG, § 72 UrhG; Solmecke and Wahlers 2012).

Obtaining the appropriate rights of use in accordance with §§ 31 ff. of the German Copyright Act is subject to a number of practical problems. In digital ecosystems, the potential to violate numerous foreign laws and regulations is almost inevitable, making it necessary to establish user agreement contracts (Klass 2013). In practice, however, this is hardly possible. Here, “fair use clauses”, common in the U.S. but not yet recognized in the German legal system, might protect

against copyright infringement and other breaches of data privacy protection.

In addition, given the current state of law in Germany, it is still unclear whether a property right for data even exists. The possibility to claim property rights assumes that data can be assigned to a legal entity under the legal system. In general, jurisprudence assumes that a disk with stored data is protected through property rights (Karlsruhe 1996; Konstanz 1996). Due to networked databases, however, in the real world this legal recourse is of little help for the ownership of media (Meier and Wehlau 1998).

The recourse to criminal law seems promising with respect to the clarification of ownership questions. For example, § 303a StGB explicitly protects data. The assignment of the protected property to a legal entity is carried out on the basis of an “act of scripture”, that is, through the technical production process. A transfer of this concept to civil law allows an unambiguous assignment and thus the possibility of establishing “data property”.

Although such an assignment of ownership of data may appear finalized or settled at first glance, data increasingly provide a value that is gaining in importance: for example, in assessing the value of companies and in bankruptcy issues. If absolute ownership of data is established, the owner could assert property right claims that would affect any subsequent handling of the dataset.

The faulty or incorrect analysis of data can have serious consequences. Liability can be incurred should the provider of big data transmit erroneous data. In the case of incorrect data transmission between a supplier and customer, a claim for damages for breach of duty of an obligation according to § 280 para. 1 BGB comes into consideration. The victim may also take a tortious claim for damages into account if erroneous data transmission lead to a violation of legal interests in the meaning of § 823 para. 1 BGB. A lawsuit may ensue regardless of any contractual relationship established between the involved parties. However, this situation generally only occurs when legal duties to ensure data safety are violated on the part of the provider.

The requirements on obligations to protect data in the area of big data are not resolved yet. In regards to obtain data, if

²These issues are examined from the perspective of German regulations on data privacy protection; with globalized information processing, international regulations are obviously necessary.

the data is used as the basis for subsequent decisions which might directly impact “life and limb” of a person, particular attention should be placed on any relevant legal responsibilities to ensure the safety of the data transmission. If the potential legal issues are of minor relevance, the implementation of strict safety rules might present an insurmountable obstacle to conduct business.

Any risk of liability depends on the circumstances of the individual case. It can usually be assumed that ensuring data safety requires a minimal amount of legal effort (Reese 1994). Even if property or ownership rights were infringed upon, the victim would have to prove a culpable violation in order to establish a legal claim. In such a case, the victim could benefit from the principles of producer liability according to § 823 para. 1 BGB. This is a fault-based liability for which the burden of proof lies with the claimant (Lehmann 1992). Because the producer’s liability was originally developed for industrial products, it is unclear whether these principles can also be applied to big data (Meyer 1997).

With heterogeneous data of varying quality and the increasing use of probabilistic analysis procedures, the question arises, when are erroneous data present? The first indications of criteria developed to answer this question were discussed in the U.S. in the context of the “Federal Data Quality Act” (Office of Information and Regulatory Affairs 2002). In these guidelines, criteria such as usefulness, completeness, objectivity, accountability, and clarity are mentioned (Gasser 2003). While these criteria may be applicable in the U.S., it might be difficult, if not impossible, to apply them globally. Whether the quality of data was sufficient or not is thus still in the judgment and rulings of individual cases.

3.3 The Economic Condition: New Value Networks for Big Data

Although the book value of Facebook at the time of the IPO was only about \$6 billion, Facebook had reached a market value of approximately \$104 billion on the day of the first IPO. This makes it clear that regardless of the specific market conditions, data have become a central component in new business models (Mandel 2012; Mayer-Schönberger and Cukier 2013). Therefore, new value-added networks have established themselves around big data.

Previously, data were considered by-products of different application systems in the context of business processes. In the context of big data, data have increasingly acquired a self-contained value. Using big data technologies, data can be used to establish competitive advantages (Manyika et al. 2011; Mayer-Schönberger and Cukier 2013). In particular, through the integration of different data sources, data are no longer earmarked. Similarly, the innovative use of data determines the value that can be achieved from them. For example, competing market participants may also exploit data (e.g. from social networks) (Mandel 2012; Mayer-Schönberger and Cukier 2013).

Organizations deal differently with data and its potential value proposition (Manyika et al. 2011). Several market players, such as banks, telecommunication service providers or health insurance companies, collect data for other purposes but with the help of big data technologies, the collected data can be harnessed for new business models (Mayer-Schönberger and Cukier 2013).

Three central competencies provide the basis for establishing new business models with big data. First, the actors must be able to combine different data sources with regard to their syntax and semantic and pragmatic characteristics. Second, cooperation with other actors is necessary in order to process the needed data. Third, it is important to collect data in a way that it can be used in a variety of contexts. Within different context, data can have a different value. For example, Google Maps and Google Street serve as the basis for the geographic contextualization of other data sources (Mayer-Schönberger and Cukier 2013; Manyika et al. 2011).

While so far the processes of data preparation, analysis, and use have been mainly corporate activities, big data favors a value network with specialized actors (Gustafson and Fink 2013). Organizations as part of this value network establish competitive advantages by specializing on central core competencies and building value chains with other service providers.

The following central roles can be identified within the big data value network (Mayer-Schönberger and Cukier 2013; Gerhardt et al. 2012):

- Data collectors are actors who focus on the collection of data and can establish appropriate access controls. Often data

collectors use data for their own purposes but can make data accessible for other actors in the value network.

- Technology manufacturers are actors who provide necessary hardware and software. Some manufacturers also provide appropriate methods and procedures for the preparation, integration, and analysis of data. An essential feature of technology manufacturers is a standardized product and service offer (Manyika et al. 2011).
- Specialists are service providers who have expanded their expertise in customized data preparation, integration, and analysis, and offer their expertise as service. In general, specialists do not possess their own data but offer their expertise to other market actors, such as data collectors.
- Data aggregators are actors who focus on the provision of specific data analysis for the mass market. An example are providers of real time traffic information who aggregate data from different sources (smartphones, data cards, official traffic information) and offer them as a service (Gustafson and Fink 2013).
- Data users are actors who have recognized the potential value of specific data and want to use it for the development of new services or products. In their role as data user, actors have an idea how to use data without necessarily owning the data or possessing the skills to use the data.
- Brokers are intermediaries who have specialized on mediating between different actors and on establishing temporary value chains. A prominent example of this is kaggle.com, a platform that announces public competitions for data analysis and brings different actors together (Carpenter 2011).
- Regulators are actors that specialize in checking and certifying regulatory assumptions such as data privacy protection, data security, and quality on behalf of data users. Regulators can thereby be public institutions at national and international levels as well as private organizations (Gerhardt et al. 2012).

Against the background of the phenomenon of “big data hubris”, it becomes apparent that especially regulators and brokers take a central role in the context of big data. Regulators create necessary qualitative conditions whereas brokers mainly build expertise in the establishment of effective value chains with big data.

Table 1 Evaluation of big data as disruptive technology (based on Christensen 1997)

Properties of disruptive technologies	Big data evaluation
The technology performs worse than established technologies according to traditional evaluation criteria.	Traditional technologies of data analysis can be operated by means of declarative languages from a wide range of users. However, big data leads to probabilistic results that can currently only be created and interpreted by experts.
The technology introduces new rating criteria.	Big data enables the analysis of large data sets relying on different formats with decreasing response times.
The technology is being observed by mainstream users with caution.	While there are already a number of examples and pilot applications for big data, the core processes of organizations are not yet in the focus.
Initially, the technology will be offered and used by start-ups and persons changing careers/businesses changing their focus.	Users of big data technology are especially small businesses that see the technology as an opportunity for new business models. Among the providers of big data technology are start-ups as well as established technology providers.
The technology leads to changes in the value creation processes.	In the field of big data a variety of new roles are being created.

4 Three Theses for the Scientific Assistance and Design of a Disruptive Technology

Following Christensen (1997), it is typical that disruptive innovations pose challenges that cannot be solved directly. The previous discussion shows that big data potentially represents a disruptive technology. Significant shifts are indicated not only from a technical but also from a legal and business perspective while at the same time significant future value creation potentials can be expected. **Table 1** summarizes previously considered aspects of big data and evaluates them in terms of postulated properties of disruptive technologies (Christensen 1997).

On the one hand, the technological, legal, and economic development around big data offers an opportunity for information systems research, with its interdisciplinary research tradition, to develop valuable knowledge gains for the explanation and prediction of information systems in business and administration. On the other hand, the traditionally distinct focus of information systems research regarding methods of designing innovative information systems opens considerable potential in terms of the relevance of information systems research in business and administration (Brinkkemper 1996).

Our proposition for the methodological and interdisciplinary approach to big data in information systems research comprises the following three theses.

- (1) Education and training concepts for the effective and responsible use of big data in research and practice

The example of “big data hubris” shows that appropriate training and education concepts are essential for the effective use of big data in research as well as in practice. Independent from big data, the ubiquity of modern information technology generates potentially valuable data sources, which can be used for the acquisition of knowledge in information systems research. The reliable use of modern methods for data analysis (e.g., machine learning, data mining, network analysis) and broad knowledge regarding their frameworks and weaknesses are the prerequisite for the use of big data as an opportunity for conducting research in information systems.

- (2) Development of modeling tools for integrated analysis and design of business processes in consideration of big data

The example of “big data hubris” shows that by simply linking different data sources, no added value can be realized. Information systems research has a long tradition in developing and testing modeling tools and methods for implementing information systems and business processes. For value creation of big data, it is necessary to consider data analysis and the integration and design of business processes. The consideration of interactions between big data as an initiator of new business processes and as a requirement framework for the corresponding data analysis is central to the development of appropriate modeling tools.

- (3) Development of resilient reference models for the responsible use of big data

As the legal perspective on big data has shown, big data requires newly developed

or adapted guidelines regarding the use of data in light of privacy and data security issues. Similarly, new processes for dealing with risks in terms of quality, use, and responsibilities are necessary. Big data will establish itself as a trusted basis for decision-making with the help of corresponding processes. There is, however, a lack of reliable reference models that describe and propose the necessary mechanisms, structures, and processes for introducing and using big data in organizations.

Overall, it should be noted that these three core theses are not related to specific technological developments but are merely based on the assumption that both data with traditional roots and big data will equally play a more important role for the explanation and design of information systems, business processes, and decision-making processes from this point in time onward.

References

- Agrawal D, Bernstein P, Bertino E, Davidson S, Dayal U, Franklin M, Gehrke J, Haas L, Halevy A, Han J, Jagadish HV, Labrinidis A, Madden S, Papakonstantinou Y, Patel JM, Ramakrishnan R, Ross K, Shahabi C, Suciu D, Vaityanathan S, Widom J (2012) Challenges and opportunities with big data: a community white paper developed by leading researchers across the United States. Computing Research Association, Washington
- American National Standards Institute (1992) American national standard for information systems: database language SQL (ANSI X3.135-1992). American National Standards Institute, Washington
- Berglund A, Boag S, Chamberlin D, Fernández MF, Kay M, Robie J, Siméon J (2010) XML path language (XPath) 2.0. World Wide Web Consortium, Cambridge

Abstract

Michael Schermann, Holmer Hemsén,
Christoph Buchmüller, Till Bitter,
Helmut Krcmar, Volker Markl,
Thomas Hoeren

Big Data

An Interdisciplinary Opportunity for Information Systems Research

“Big data” describes technologies that promise to fulfill a fundamental tenet of research in information systems, which is to provide the right information to the right receiver in the right volume and quality at the right time. For information systems research as an application-oriented research discipline, opportunities, and risks arise from using big data. Risks arise primarily from the considerable number of resources used for the explanation and design of fads. Opportunities arise because these resources lead to substantial knowledge gains, which support scientific progress within the discipline and are of relevance to practice as well.

From the authors’ perspective, information systems research is ideally positioned to support big data critically and use the knowledge gained to explain and design innovative information systems in business and administration – regardless of whether big data is in reality a disruptive technology or a cursory fad. The continuing development and adoption of big data will ultimately provide clarity on whether big data is a fad or if it represents substantial progress in information systems research. Three theses also show how future technological developments can be used to advance the discipline of information systems. Technological progress should be used for a cumulative supplement of existing models, tools, and methods. By contrast, scientific revolutions are independent of technological progress.

Keywords: Big data, Business information systems, Systems engineering, Big data hubris, Disruptive technology, Information management

- Boag S, Chamberlin D, Fernández MF, Florescu D, Robie J, Siméon J (2011) XQuery 1.0: an XML query language. World Wide Web Consortium, Cambridge
- Brinkkemper S (1996) Method engineering: engineering of information systems development methods and tools. *Information and Software Technology* 38(4):275–280
- Buhl HU, Röglinger M, Moser F, Heidemann J (2013) Big data: a fashionable topic with(out) sustainable relevance for research and practice? *Business & Information Systems Engineering* 55(2):65–69
- Carpenter J (2011) May the best analyst win. *Science* 331(6018):698–699
- Chen H, Chiang RH, Storey VC (2012) Business intelligence and analytics: from big data to big impact. *MIS Quarterly* 36(4):1165–1188
- Christensen C (1997) The innovator’s dilemma: when new technologies cause great firms to fail. Harvard Business Review, Boston
- Dean J, Ghemawat S (2004) MapReduce: simplified data processing on large clusters. In: Proc 6th symposium on operating system design and implementation, San Francisco
- Dreier T, Schulze G (2013) Urheberrechtsgesetz: Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Beck, München
- Eberspächer J, Wohlmuth O (eds) (2013) Big Data wird neues Wissen. Münchner Kreis, München
- Gasser U (2003) Information quality and the law or how to catch a difficult horse. The Berkman Center for Internet & Society Research, Harvard Law School, Cambridge
- Gaster JL (1999) Der Rechtsschutz von Datenbanken: Kommentar zur Richtlinie 96/9/EG mit Erläuterungen zur Umsetzung in das deutsche und österreichische Recht. Heymanns, Köln
- Gerhardt B, Griffin K, Klemann R (2012) Unlocking value in the fragmented world of big data analytics: how information intermediaries will create a new data ecosystem. Cisco Internet Business Solutions Group, San Jose
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L (2009) Detecting influenza epidemics using search engine query data. *Nature* 457(7232):1012–1014
- Gustafson T, Fink D (2013) Winning within the data value chain. Innosight. <http://www.innosight.com/innovation-resources/strategy-innovation/winning-within-the-data-value-chain.cfm>. Accessed 2014-02-28
- Klass N (2013) Neue Internettechnologien und das Urheberrecht: Die schlichte Einwilligung als Rettungsanker? *ZUM – Zeitschrift für Urheber- und Medienrecht* 57(1):1–9
- Klein D, Tran-Gia P, Hartmann M (2013) Big data. *Informatik Spektrum* 36(3):319–323
- Krcmar H (2009) Informationsmanagement. Springer, Heidelberg
- Lazer D, Kennedy R, King G, Vespignani A (2014) The parable of Google Flu: traps in big data analysis. *Science* 343(3):1203–1205
- Lehmann M (1992) Produkt- und Produzentenhaftung für Software. *NJW – Neue Juristische Wochenschrift* 1992(28):1721–1725
- Konstanz LG (1996) Kein Schadensersatz bei Datenverlust durch Stromausfall aufgrund Leitungsbeschädigung bei Baggerarbeiten. *NJW – Neue Juristische Wochenschrift* 1996(40):2662
- Lycett M (2013) Datafication: making sense of (big) data in a complex world. *European Journal of Information Systems* 22(4):381–386
- Mandel M (2012) Beyond goods and services: the (unmeasured) rise of the data-driven economy. Progressive Policy Institute, Washington
- Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH (2011) Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute, Atlanta
- Mayer-Schönberger V, Cukier K (2013) Big data: a revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, New York
- McAfee A (2013) Big data’s biggest challenge? Convincing people NOT to trust their judgment. <http://blogs.hbr.org/2013/12/big-datas-biggest-challenge-convincing-people-not-to-trust-their-judgment/>. Accessed 2014-02-28
- McAfee A, Brynjolfsson E (2012) Big data: the management revolution. *Harvard Business Review* 2012(October):1–9
- Meyer A (1997) Die Haftung für fehlerhafte Aussagen in wissenschaftlichen Werken. *ZUM – Zeitschrift für Urheber- und Medienrecht* 41(1):26–33
- Meier K, Wehla A (1998) Die zivilrechtliche Haftung für Datenlöschung, Datenverlust und Datenzerstörung. *NJW – Neue Juristische Wochenschrift* 1998(22):1585–1591
- Office of Information and Regulatory Affairs (2002) Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by federal agencies. Office of Management and Budget, Executive Office of the President, Washington
- Karlsruhe OLG (1996) Haftung für Zerstörung von Computerdaten. *NJW – Neue Juristische Wochenschrift* 1996(3):200–201
- Prud’hommeaux E, Seaborne A (2008) SPARQL query language for RDF. World Wide Web Consortium, Cambridge
- Reese J (1994) Produkthaftung und Produzentenhaftung für Hard- und Software. *DStR – Deutsches Steuerrecht* 1994(31):1121–1126
- Schroek M, Shockley R, Smart J, Romero-Morales D, Tufano P (2012) Analytics: the real-world use of big data. IBM Global Business Services, Somers
- Solmecke C, Wahlers J (2012) Rechtliche Situation von Social Media Monitoring-Diensten – Rechtskonforme Lösungen nach dem Datenschutz- und dem Urheberrecht. *ZD – Zeitschrift für Datenschutz* 2012(12):550–554
- Steininger K, Riedl R, Roithmayr F, Mertens P (2009) Fads and trends in business and information systems engineering and information systems research – a comparative literature analysis. *Business & Information Systems Engineering* 51(6):411–428
- The Economist Intelligence Unit (2014) The data storm: retail and the big data revolution. The Economist, London
- Thomas JJ, Cook KA (2005) Illuminating the path: the research and development agenda for visual analytics. IEEE Computer Society Press, Washington
- Wandtke A-A, Bullinger W (2009) Praxiskommentar zum Urheberrecht. Beck, München
- White T (2012) Hadoop – the definitive guide: storage and analysis at internet scale. O’Reilly, Sebastopol