

Chapter 12

Big Data in the Finance and Insurance Sectors

Kazim Hussain and Elsa Prieto

12.1 Introduction

The finance and insurance sector by nature has been an intensively data-driven industry for many years, with financial institutes having managed large quantities of customer data and using data analytics in areas such as capital market trading. The business of insurance is based on the analysis of data to understand and effectively evaluate risk. Actuaries and underwriting professionals depend upon the analysis of data to be able to perform their core roles; thus it is safe to state that this data is a dominant force in the sector.

There is however an increase in prevalence of data which falls into the domain of big data, i.e. high volume, high velocity, and high variety of information assets born out of the advent of new customer, market, and regulatory data surging from multiple sources. To add to the complexity is the co-existence of structured and un-structured data. Unstructured data in the financial services and insurance industry can be identified as an area where there is a vast amount of un-exploited business value. For example, there is much commercial value to be derived from the large volumes of insurance claim documentation which would predominately be in text form and contains descriptions entered by call centre operators, notes associated with individual claims and cases. With the help of big data technologies not only can value be more efficiently extracted from such a data source, but the analysis of this form of unstructured data in conjunction with a wide variety of datasets to extract faster, targeted commercial value. An important characteristic of big data in this industry is value—how can a business not only collect and manage big data, but how can the data which holds value be identified and how can organizations forward-engineer (as opposed to retrospectively evaluate) commercial value from the data.

K. Hussain • E. Prieto (✉)
Atos Spain, S.A., Albarracín, 25 28037 Madrid, Spain
e-mail: kazim.hussain@atos.net; elsa.prieto@atos.net

12.1.1 Market Impact of Big Data

The market for big data technology in the financial and insurance domains is one of the most promising. According to TechNavio's forecast (Technavio 2013), the global big data market in the financial services sector will grow at a CAGR of 56.7 % over the period 2012–2016. One of the key factors contributing to this market growth is the need to meet financial regulations, but the lack of skilled resources to manage big data could pose a challenge.

The key vendors dominating this space include Hewlett-Packard, IBM, Microsoft, and Oracle that are global well-established players with a generalist profile. However, the appeal of the market will be a pull factor on new entrants in the coming years.

With data being the most important asset, this technology is especially favourable and differentiating for financial services organizations, as said by the IBM Institute for Business Value's report "Analytics: The real-world use of big data in financial services" (IBM 2013). By leveraging this asset, banks and financial markets firms can gain a comprehensive understanding of markets, customers, channels, products, regulations, competitors, suppliers, and employees that will let them better compete. Therefore, this is a positive trend in the market and is expected to drive the growth of the global big data market in the financial services sector.

In terms of data strategy, financial services organizations are taking a business-driven approach to big data: business requirements are identified in the first place and then existing internal resources and capacities are aligned to support the business opportunity, before investing in the sources of data and infrastructures. However, not all financial organizations are keeping the same pace. According to the IBM report, while 26 % are focused on understanding the principal notions (compared with 24 % of global organizations), the majority are either defining a roadmap related to big data (47 %) or already conducting big data pilots and implementations (27 %).

Where they lag behind their cross-industry peers is in using more varied data types within their big data implementations. Slightly more than 21 % of these firms are analysing audio data (often produced in abundance in retail banks' call centres), while slightly more than 27 % report analysing social data (compared to 38 % and 43 %, respectively, of their cross-industry peers). This lack of focus on unstructured data is attributed to the on-going struggle to integrate the organizations' massive structured data.

12.2 Analysis of Industrial Needs in the Finance and Insurance Sectors

The advent of big data in financial services can bring numerous advantages to financial institutions. Benefits that come with the greatest commercial impact are highlighted as follows:

Enhanced Levels of Customer Insight, Engagement, and Experience With the digitization of financial products and services and the increasing trend of customers interacting with brands or organizations in the digital space, there is an opportunity for financial services organizations to enhance their level of customer engagement and proactively improve the customer experience. Many argue that this is the most crucial area for financial institutes to start leveraging big data technology to stay ahead, or even just keep up with competition. To help achieve this, big data technologies and analytical techniques can help derive insight from newer unstructured sources such as social media.

Enhanced Fraud Detection and Prevention Capabilities Financial services institutions have always been vulnerable to fraud. There are individuals and criminal organizations working to defraud financial institutions and the sophistication and complexity of these schemes is evolving with time. In the past, banks analysed just a small sample of transactions in an attempt to detect fraud. This could lead to some fraudulent activities slipping through the net and other “false positives” being highlighted. Utilization of big data has meant these organizations are now able to use larger datasets to identify trends that indicate fraud to help minimize exposure to such a risk.

Enhanced Market Trading Analysis Trading the financial markets started becoming a digitized space many years ago, driven by the growing demand for the faster execution of trades. Trading strategies that make use of sophisticated algorithms to rapidly trade financial markets are a major benefactor of big data.

Market data can be considered itself, as big data. It is high in volume, it is generated from a variety of sources, and it is generated at a phenomenal velocity. However, this big data does not necessarily translate into actionable information. The real benefit from big data lies in effectively extracting actionable information and integrating this information with other sources. Market data from multiple markets and geographies as well as a variety of asset classes can be integrated with other structured and unstructured sources to create enriched, hybrid datasets (a combination of structured and unstructured data). This provides a comprehensive and integrated view of the market state and can be used for a variety of activities such as signal generation, trade execution, profit and loss (P&L) reporting, and risk measurement, all in real-time hence enabling more effective trading.

12.3 Potential Big Data Applications in Finance and Insurance

Three potential applications for the finance and insurance sector were described and developed in Zillner et al. (2013, 2014) as representatives of the application of big data technologies in the sector (Table 12.1).

Table 12.1 Summary of big data application scenarios for the finance and insurance sector

Name	Market manipulation detection.
Summary	Detection of false rumours that try to manipulate the market.
Synopsis	Financial markets are often influenced by rumours. Sometimes false rumours are intentionally placed in order to distract and mislead other market participants. These behaviours differ based on the intended outcome of the manipulation. Examples of market abuse are market sounding (the illegal dissemination of untrue information about a company whose stock is traded on exchanges) and pump and dump (false positive reports are published about a company whose shares are tradable with the goal of encouraging other market participants to buy stock in the corresponding company; an increase in demand would cause the price of the stock to rise to an artificial level).
Business objectives	Identifying hoaxes and assessing the consistency of new information with other reliable sources.
Name	Reputational risk management.
Summary	Assessment of exposure to reputational risk connected to consulting services offered by banks to their customers.
Synopsis	A negative perception can adversely affect a bank's ability to maintain existing, establish new business relationships, or continued access to sources of funding. The increase in the probability of default (issuer credit risk), the price volatility, and the difficulties to exchange specific financial products on restricted markets have all contributed to the increase of the reputational and operational risk associated with brokerage and advisory services. Banks and financial institutions usually offer third party financial products. This implies that a lack of performance of a third party product could have real impacts on the relationship between the bank and its customers.
Business objectives	To monitor third parties' reputation and the effects of reputation disruption on the direct relationship between banks and customers.
Name	Retail brokerage.
Summary	Discover topic trends, detect events, or support the portfolio optimization/asset allocation.
Synopsis	A general trend in the whole industry of retail brokerage and market data is to come up with functionalities that offer actionable information. The focus is no longer on figures based on quantitative historical data, e.g. key figures or performance data. Instead, investors look for signals that have some kind of predictive element yet are easy to understand. In that sense, the extraction of sentiments and topics from textual sources is a perfect add-on for the conventional data and functionalities that are already offered by retail brokerage companies.
Business objectives	Collecting and reviewing various sources of financial information (on markets, companies, or financial institutions) repeatedly by automation of this task.

12.4 Drivers and Constraints for Big Data in the Finance and Insurance Sectors

The successful realization of big data in finance and insurance has several drivers and constraints.

12.4.1 Drivers

The following **drivers** were identified for big data in the finance and insurance sector:

- **Data Growth:** Financial transaction volumes are increasing, leading to data growth in financial services firms. In capital markets, the presence of electronic trading has led to an increase in the number of trades. Data growth is not limited to capital markets businesses. The Capgemini/RBS Global Payments study for 2012 (Capgemini 2012) estimates that the global number of electronic payment transactions is about 260 billion and growing between 15 and 22 % for developing countries.
- **Increasing scrutiny from regulators:** Regulators of the industry now require a more transparent and accurate view of financial and insurance businesses, this means that they no longer want reports; they need raw data. Therefore financial institutions need to ensure that they are able to analyse their raw data at the same level of granularity as the regulators.
- **Advancements in technology mean increased activity:** Thanks largely to the digitization of financial products and services, the ease and affordability of executing financial transactions online has led to ever-increasing activity and expansion into new markets. Individuals can make more trades, more often, across more types of accounts, because they can do so with the click of a button in the comfort of their own homes.
- **Changing business models:** Driven by the aforementioned factors, financial institutions find themselves in a market that is fundamentally different from the market of even a few years ago. Adoption of big data analytics is necessary to help build business models for financial institutions geared towards retention of market share from the increasing competition coming from other sectors.
- **Customer insight:** Today the relationship between banks and consumers has been reversed: consumers now have transient relationships with multiple banks. Banks no longer have a complete view of their customer's preferences, buying patterns, and behaviours. Big data technologies therefore play a focal role in enabling customer centricity in this new paradigm.

12.4.2 Constraints

The **constraints** for big data in the finance and insurance sector can be summarized as follows:

- **Old culture and infrastructures:** Many banks still depend on old rigid IT infrastructure, with data siloes and a great many legacy systems. Big data, therefore, is an add-on, rather than a completely new standalone initiative. The culture is an even bigger barrier to big data deployment. Many financial organizations fail to implement big data programs because they are unable to appreciate how data analytics can improve their core business.
- **A lack of skills:** Some organizations have recognized the data and the opportunities the data presents; however they lack human capital with the right level of skills to be able to bridge the gap between data and potential opportunity. The skills that are “missing” are those of a data scientist.
- **Data “Actionability”:** The next main challenge can be seen in making big data actionable. Big data technology and analytical techniques enable financial services institutions to get deep insight into customer behaviour and patterns, but the challenge still lies in organizations being able to take specific action based on this data.
- **Data privacy and security:** Customer data is a continuing cause for concern. Regulation remains a big unknown: what is and is not legally permissible in the ownership and use of customer data remains ill-defined, and that is an inhibiting factor to rapid and large-scale adoption.

12.5 Available Finance and Insurance Data Resources

The financial service system has several major pools of data that are held by different stakeholders/parties. Data are classified into three major categories:

Structured Data This refers to information with a high degree of organization, such that inclusion in a relational database is seamless and readily searchable by simple, straightforward search engine algorithms, or other search operations. Examples of financial structured data sources are:

- Trading systems (transaction data)
- Account systems (data on account holdings and movements)
- Market data from external providers
- Securities reference data
- Price information
- Technical indicators

Unstructured Data Although the financial industry has previously focused on high velocity market data, it is now moving towards unstructured data to changing trading dynamics. Examples of financial unstructured data are:

- Daily stock feeds
- Company announcements (ad-hoc news)
- Online news media
- Articles/blogs
- Customers' feedback/experiences

Semi-structured Data A form of structured data that does not conform to the formal structure of data models associated with relational databases or other forms of data tables, but even so contains tags or markers to separate semantic elements and enforce hierarchies of records and fields within the data. Examples of semi-structured data are expressed in meta-languages (mostly XML-based) such as:

- Financial products Markup Language (FpML)
- Financial Information eXchange (FIX)
- Interactive Financial eXchange (IFX)
- Market Data Definition Language (MDDL)
- Financial Electronic Data Interchange (FEDI)
- Open Financial eXchange (OFX)
- eXtensible Business Reporting Language (XBRL)
- SWIFTStandards

Nowadays the amount of unstructured information in enterprises is around 80–85 %. The financial and insurance industry has vast repositories of structured data in comparison to other industries, with a large amount of this information having its origin inside the organization.

12.6 Finance and Insurance Sector Requirements

12.6.1 *Non-technical Requirements*

Data Protection and Privacy Particularly in the EU, there are numerous data protection and privacy issues to consider when undertaking big data analytics. Regulatory requirements dictate that personal data must be processed for specified and lawful purposes and that the processing must be adequate, relevant, and not excessive. The impact of these principles for financial services organizations is significant, with individuals being able to ask financial services organizations to remove or refrain from processing their personal data in certain circumstances.

This requirement could lead to increased costs for financial services organizations, as they deal with individuals' requests. This removal of data may also

lead to the dataset being skewed, as certain groups of people will be more active and aware of their rights than others.

Confidentiality and Regulatory Requirements Any information related by a third party that is subject to big data analytics is likely to be confidential information. Therefore, financial services organizations will need to ensure that they comply with their obligations and that any use of such data does not give rise to a breach of their confidentiality or regulatory obligations.

Liability Issues Just because big data contains an enormous amount of information, it does not mean that it reflects a representative sample of the population. Therefore there is a risk of misinterpreting the information produced and liability may arise where reliance is placed on that information. This is a factor that financial services organizations have to take into account when looking at using big data in analytical models and ensuring that any reliance placed upon the output comes with relevant disclaimers attached.

12.6.2 Technical Requirements

Data Extraction and Sentiment Classification Though the definition of sentiment is vague, in general, a sentiment on an object is a positive or negative view, attitude, emotion, or appraisal on or from a document author or actor.

Sentiment is often expressed in a domain-specific way, and using non-domain-specific vocabulary may lead to misclassifications. The goal is to extract facts and sentiments concerning the financial use cases: financial instruments, situations, conditions, indicators, and experts' assessments regarding these instruments, as well as investors' sentiment, etc. The classification of sentiment can be done at several levels: words, phrases, sentences, paragraphs, documents, and even multiple documents, and then aggregate.

Data extraction needs to cope with noise, misinformation, irony, bias, or uncertainty. In addition, with sentiment it is important not only to determine the sentiment of a piece of information, but how words affect the semantic orientation and how sentiment changes.

Data Quality The more timely, accurate, and relevant the data (along with good analytics), the better the assessment of the current financial state is. This requires better processes of identifying and maintaining the data sources of interest, verifying, cleaning, transforming, integrating, and deduplicating data. Due to the large amount of available data, there is a need for automation and scalability processes. Language detection methods also need to be refined to improve precision and reliability.

Data Acquisition For banks and financial services providers, the volume of data they generate, consume, store, and access will increase exponentially year over year. The applications depend on acquiring and accessing massive amounts of

historical heterogeneous information and live feeds of unstructured, semi structured, and structured information. A significant amount of data comes from internal structured data, though there is a growing trend towards external unstructured data (from news, blogs, articles, social networks, and websites). Even when there can be a wide variety of data sources to access, the actual ones that are required depend on the design for a specific application.

Data Integration/Sharing This describes the task to overcome the heterogeneity of disparate data sources in terms of hardware, software, syntax, and/or semantics by providing access tools that enable interoperability.

The data is usually scattered among different heterogeneous sources with differing conceptual representations (different structures and data semantics) but it is encapsulated into a single, homogeneous data source to the end user.

The motivation for integration may be based on strategic or operational considerations. Regarding strategic considerations and analysis, it may not be required to constantly integrate the data but to integrate data snapshots at a certain point in time. For operational analysis a real-time integration of the most up-to-date information may be required.

Typically data integration is not a once-off conversion but an on-going task, therefore poses the additional constraint that the chosen solution needs to be robust in terms of adaptability, extensibility, and scalability. Approaches leveraging standards such as eXtensible Business Reporting Language (XBRL) and Linked Data show promise (O’Riáin et al. 2012).

This rapid generation of continuous streams of information has challenged the storage, computation, and communication capabilities in computing systems, as they impose high resource requirements on data stream processing systems.

Decision Support Systems (DSS) Model-driven DSS emphasises access to and manipulation of statistical, financial, optimization, and/or simulation models. Models use data and parameters to aid decision-makers in analysing a situation, for instance, assessing and evaluating decision alternatives and examining the effect of changes. This requires integrating information from the knowledge base into financial event detection models, visualization models, decision-models, and for scalable execution of these models.

For some application scenarios, the response of the system should support real-time or near-real-time insights. The velocity of the response is subject to the end user requirements.

In DSS, visualization is an extremely useful tool for providing overviews and insights into overwhelming amounts of data to support the decision-making process.

Data Privacy and Security Top priorities for the financial sector today include on-going regulatory compliance [e.g. Sarbanes-Oxley (SOX) Act, U.S. Government (2002); EU data protection directive, Parliament (1995); cyber security directive, Parliament (2013)] and risk mitigation, continued adaptation to the expectations of

consumers for anywhere/anytime service, reducing operational costs, and increasing efficiencies through use of cloud-based services.

Banking and financial institutions need to secure the storage, transit, and use of corporate and personal data across business applications, including online banking and electronic communications of sensitive information and documents.

The increasingly global nature and high-interconnectivity of the industry makes it necessary to comprehensively address international data security and privacy regulations, from the front to the back-end, and along the full supply chain, including third parties. Data is not always stored in-house but with third parties. Using commercial “cloud” services as data storage locations poses potential privacy and security problems since the terms of service for these products are often poorly understood.

12.7 Technology Roadmap for Big Data in the Finance and Insurance Sectors

For each requirement in the sector, this section presents applicable technologies and the research questions to be developed (Fig. 12.1; Table 12.2). All references presented here are from Curry et al. (2014).

12.7.1 Data Acquisition

- Acquisition pipeline technology.
 - Research Question: Data stream management. Current data analysis in the stored-data domain shall need to move to management of data in the data stream itself.
- Proprietary APIs technology.
 - Research Question: Privacy and anonymization at collection time. The data collection process shall require intrinsic data anonymization and/or decoupling of personal data from data emanating from business processes or otherwise.
 - Research Question: Social APIs. Moving ahead of existing proprietary (or even open) APIs, social APIs into financial services datasets need to be investigated.

Technical Requirement	Technology	Research Question
Data Acquisition	Acquisition pipeline	Data stream management
	APIs technology	Privacy and anonymization at collection time Social APIs
Data quality	Manual processing and validation	Scalable data curation and validation New methods to improve precision and reliability
Data extraction	Language modelling	Statistical language models
	Machine Learning	Required inference functionality
	Scalability in real-time	Processing of large datasets
Data integration / sharing	Wrappers/mediators to encapsulate distributed & automatic data and schema mapping	User-specific integration Data variety: sentiments, quantitative information Scaling methods for large data volumes and near-real time processing.
Decision support	Multi-attribute decision models	Stream-based data mining Machine learning adaptation to evolving content
	Resource allocation in mining data streams	Improved storage, computation and communication capabilities
Data privacy & security	Roles-based IdM and access control	Privacy by design Security by design Data Security for public-private hybrid environments
		Enhanced Compliance management
	Database encryption NoSQL	Apply external encryption and authentication controls

Fig. 12.1 Mapping requirements to research questions in the finance and insurance sectors

12.7.2 Data Quality

- Manual processing and validation technology.
 - Research Question: Scalable data curation and validation.
 - Research Question: New methods to improve precision and reliability.

12.7.3 Data Extraction

- Language modelling technology.
 - Research Question: Obtaining keywords and key-phrases by using statistical language models.

Table 12.2 Timeframe of the major expected outcomes of the big data roadmap for the finance and insurance sector

Technical requirement	Year 1	Year 2	Year 3	Year 4	Year 5
Data acquisition			Social APIs		Data stream management. Privacy and anonymization at collection time
Data quality			Scalable data curation and validation		New methods to improve precision and reliability
Data extraction		Statistical language models	New machine learning techniques to satisfy the newly required inference functionality		Processing of large datasets
Data integration/sharing					User-specific integration Data variety: sentiments, quantitative information Scaling methods for large data volumes and near-real-time processing
Decision support		Stream-based data mining			Machine learning adaptation to evolving content Improved storage, computation, and communication capabilities
Data privacy and security		Apply external encryption and authentication controls	Privacy by design Security by design		Data security for public-private hybrid environments Enhanced compliance management

- Machine Learning technology.
 - Research Question: The size of datasets in financial services makes it necessary for new machine learning techniques to satisfy the newly required inference functionality.
- Scalability in real-time technology: Real-time information is of interest in some application scenarios of financial services.

- Research Question: The challenge of processing large datasets represents a requirement for research in the scalability of data processing in real-time as datasets grow in size and number.

12.7.4 Data Integration/Sharing

- Wrappers/mediators to encapsulate distributed data and automatic data and schema mapping technology: Sources of data in the financial services industry can be distributed across organizations, or across time and space.
 - Research Question: User-specific integration. Integration of data for the benefit of specific users (namely, business processes, or target end user organizations).
 - Research Question: Data variety: sentiments, quantitative information.
 - Research Question: Scaling methods for large data volumes and near-real-time processing. This research challenge is in relation to the “scalability in real time” described earlier, under “data extraction”.

12.7.5 Decision Support

- Multi-attribute decision-models technology: The availability of information from multiple sources will provide multiple attribute types that become available to include in decision-models.
 - Research Question: Stream-based data mining.
 - Research Question: Machine learning adaptation to evolving content.
- Resource allocation in mining data streams technology: Elastic computing today allows for dynamic resource allocation as required. Improvements may be required in resource allocation for near real-time support to decision-making.
 - Research Question: Improved storage, computation, and communication capabilities.

12.7.6 Data Privacy and Security

- Roles-based identity management and access control technology: access control in the context of large datasets will pose a problem when sensitive data (business process related) begins to be exploited in large datasets and integrated with other data, and accessed by third parties.

- Research Question: Privacy by design | Security by design.
- Advances in “privacy by design” to link analytics needs with protective controls in processing and storage.
- Research Question: Data Security for public-private hybrid environments.
- The advent of cloud storage and computation services, however, comes at the expense of data security and user privacy.
- Research Question: Enhanced Compliance management (data protection, others). Research has already been initiated, but needs to continue in providing methodologies and infrastructures that facilitate the monitoring, enforcement, and audit of quantifiable indicators on the security of a business process.
- Database encryption technology: The security concept of NoSQL databases generally relies on external enforcing mechanisms.
- Research Question: Review the security architecture and policies of the overall system and apply external encryption and authentication controls to safeguard NoSQL databases. Data security for public-private hybrid environments.

12.8 Conclusion and Recommendations for the Finance and Insurance Sectors

The Finance and insurance sector analysis for the roadmap is based on four major application scenarios based on exploiting banks and insurance companies’ own data to create new business value. The findings of this analysis show that there are still research challenges to develop the technologies to their full potential in order to provide competitive and effective solutions. These challenges appear at all levels of the big data value chain and involve a wide set of different technologies, which would make necessary a prioritization of the investments in R&D. In broad terms there seems to be a general agreement on real-time aspects, better data quality techniques, scalability of data management and processing, better sentiment classification methods, and compliance with security requirements along the supply chain. However, it is worth mentioning the importance of the application scenario and the real needs of the end user in order to determine these priorities. At the same time, apart from the technological aspects, there are organizational, cultural, and legal factors that will play a key role in how the financial services market takes on big data for its operations and business development.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (<http://creativecommons.org/licenses/by-nc/2.5/>) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this book are included in the work’s Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in

the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt, or reproduce the material.

References

- Capgemini. (2012). *World payments report 2012*. Available at http://www.capgemini.com/resource-file-access/resource/pdf/The_8th_Annual_World_Payments_Report_2012.pdf. Accessed 2014.
- Curry, E., Ngonga, A., Domingue, J., Freitas, A., Strohbach, M., Becker, T. et al. (2014). D2.2.2. Final version of the technical white paper. Public deliverable of the EU-Project BIG (318062; ICT-2011.4.4).
- IBM. (2013). *Analytics: The real-world use of big data in financial services*. IBM Institute for Business Value. Available at http://www-935.ibm.com/services/multimedia/Analytics_The_real_world_use_of_big_data_in_Financial_services_Mai_2013.pdf. Accessed 2014.
- O'Riáin, S., Curry, E., & Harth, A. (2012). XBRL and open data for global financial ecosystems: A linked data approach. *International Journal of Accounting Information Systems*, 13, 141–162. doi:10.1016/j.accinf.2012.02.002.
- Parliament, E. (1995). *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*. Available at <http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31995L0046>. Accessed 2014.
- Parliament, E. (2013). *Proposal for a Directive of the European Parliament and of the Council concerning measures to ensure a high common level of network and information security across the Union, from European Commission*. Available at http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1666. Accessed 2014.
- Technavio. (2013). *Global big data market in the financial services sector 2012-2016*. Retrieved from <http://www.technavio.com>. Accessed 2014.
- U.S. Government. (2002). *Sarbanes-Oxley Act of 2002*. Available at 107th Congress Public Law 204. <http://www.gpo.gov/fdsys/pkg/PLAW-107publ204/html/PLAW-107publ204.htm>
- Zillner, S., Rusitschka, S., Munné, R., Lippell, H., Lobillo, F., Hussain, K. et al. (2013). D2.3.1. *First draft of sector's roadmap*. Public Deliverable of the EU-Project BIG (318062; ICT-2011.4.4).
- Zillner, S., Neururer, S., Munné, R., Prieto, E., Strohbach, M., van Kasteren, T. et al. (2014). D2.3.2. *Final version of the sectorial requisites*. Public Deliverable of the EU-Project BIG (318062; ICT-2011.4.4).