

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339952304>

# Open Data Availability and Suitability for Financial Analyses

Chapter · March 2020

DOI: 10.1007/978-3-030-42416-9\_25

CITATIONS

0

READS

21

4 authors, including:



**Gulcan Baskurt**

Beykent Üniversitesi

1 PUBLICATION 0 CITATIONS

[SEE PROFILE](#)



**Peiman Alipour Sarvari**

Luxembourg Institute of Science and Technology (LIST)

23 PUBLICATIONS 68 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Airline network analysis based on star-shaped hubs [View project](#)



Industry 4.0 [View project](#)



# Open Data Availability and Suitability for Financial Analyses

Gulcan Baskurt<sup>2</sup>, Sébastien-Augustin Martin<sup>1</sup>, Peiman Alipour Sarvari<sup>1</sup> (✉),  
and Djamel Khadraoui<sup>1</sup>

<sup>1</sup> Trusted Distributed and Optimized Systems,  
Luxembourg Institute of Science and Technology, Belval, Luxembourg  
{sebastien.martin, peiman.alipour, djamel.khadraoui}@list.lu  
<sup>2</sup> Management Department, Faculty of Management of Beykent University, Istanbul, Turkey  
gulcan.bskurtt@gmail.com

**Abstract.** The purpose of this study is to explore the benefits of Open Data initiatives for the financial sector. The first approach is to take the financial activities, to consider their data needs (in terms of data kinds and characteristics) and consider if they are available, and in a suitable format for these goals. Even though it has been done for some cases and applied to the case of Luxembourg, we are the first and foremost group to suggest and explain a systematic method, based on an Open Data literature review, to assess the dimensions characterizing the availability (not only the data accessibility, but also the conditions around this accessibility) and the suitability of the data (not only their existence but other dimensions of their accessibility, like licenses). Although the accessibility of financial, economic and commercial information is highly dependent on contextual legal and political frameworks, this framework is adaptable enough to be used for different countries, and even beyond the case of financial data. The last part is showcasing the importance of the title of this study for financial data in Luxembourg. This contribution intends to be useful both for people in charge of the Open Data initiatives in the public and private sectors; for practitioners and researchers aiming at reusing these data; and for Open Data researchers.

**Keywords:** Financial data assessment · Open Data · Open Government Data

## 1 Introduction

The increasing size and diversity of available data is the fundamental part of the digital economy, with the notion that data may be the new oil. Although it is excessive in its formulation, yet consistent with the larger and larger production and availability of data and with the emergence of concepts such as Big Data and the progress of computing architectures, allowing to build models based on large amounts of data. It is also essential for academic financial discipline and financial activities. The field is rapidly evolving since the financial crisis of 2007, with the consecutive evolving regulatory framework (for example to protect the general public from harmful financial products), the emergence

of new business models around data provision, and technological trends such as cloud computing.

At the crossroads of these trends, Open Data may have a central position, as it is aiming to ensure the data availability as well as their publication in a suitable way for different kinds of reusers with different goals.

The main question addressed here is to explore the benefits of adopting the Open Data paradigm for financial data, as well as the conditions required to leverage these benefits and the obstacles to mitigate. This study intends to be useful for financial researchers and practitioners as they need to identify and characterize available data, and to (potential) data handlers aiming at publishing their data following the Open Data principles. If the main audience targeted are the people working with financial data, we are also contributing to the Open Data research through a topic lens, as it has been done for other domains, e.g., on judicial data (Marković and Gostojić 2018).

To address the question of open data availability and suitability for financial analyses, the first part is considering the intersections between Open Data research and financial data or, in other words, how Open Data is leveraged by financial literature and conversely: for which purpose, following which approaches. Consistently with the insights from the Open Data literature review, we propose a method as an evaluation framework, allowing characterization of the data about their availability and their suitability. This method is designed to be useful beyond the sole case of financial data. This framework is then applied to the financial data in Luxembourg, chosen as a case study.

## 2 Financial Data and Open Data Research

### 2.1 Financial Data

Finance at large is divided into personal, corporate & public finances, implying the leverage of different disciplinary perspectives and approaches, different regulation, objectives, and data. When this comes to the definition of the object that represents “financial data,” there is no single and straightforward definition. It is why, even in textbooks, financial data are more often defined according to their properties ((Koop 2006), (Tsay 2010)), as they are constituting time series or cross-sectional data, or on very generic distinctions: quantitative, qualitative, historical or real time. For example, for corporate finance, financial data consist of pieces or sets of information related to the financial health of a business. The pieces of data are used by internal management to analyze business performance and determine whether tactics and strategies need to be altered. People and organizations outside business will also use financial data reported by the business to judge its creditworthiness, decide whether to invest in the business and determine whether the business is complying with government regulations.

On top of this, some data may be considered as belonging directly to the financial domain; finance is also and more using other types of data like corporate data or economic data at large. Moreover, finance research or practice are fruitfully using unstructured data, e.g., through the use of natural language processing technics. The book on financial information coordinated by Victoria Lemieux is well illustrating these issues (Lemieux 2013). It is leading to a deep heterogeneity of kinds and parameters, requiring relying on a broad definition of financial data to envision them from the lens of Open Data.

## 2.2 Financial Information Ecosystem and the Case for Open Data

Assessing the potential of the Open Data paradigm firstly requires considering how the financial information is currently accessed, and what kinds of ecosystems do exist around this information.

For a long time, financial information market was dominated by firms like Bloomberg and Reuters. The rise of Internet has led to the emergence of a large number of websites offering financial data more or less for free, but not always providing detailed information on the intellectual property dimension or the limitations opposed to the data reuse. Another issue of these offers is the absence of statements on the long-term access to the data, nor on the technical means offered to access the data. Even big players like Yahoo Finance and Google Finance can choose to discontinue their services and did it with their Application Program Interfaces (APIs). The discontinuation of these APIs, even if the data may remain accessible through other means, is hindering the automation of financial information processing and analysis. The release of the original raw data is a remedy against this issue, ensuring data availability on the long-term, even when the intermediary services disappear or decide to change their policy.

Another purpose of Open Data is also to increase the value creation from the re-use of the data, the business models (Magalhaes et al. 2014) have been explored by the literature, especially those based on data store (Boyd and Crawford 2012) (Pereira et al. 2018). According to these studies, Open Data should enable the emergence of new players on the market, lowering the entry cost and allowing various niche markets. A good illustration of these possibilities is the case of Quandl, which is showing at the same time the opportunity of Open Data to sustain a business model for a new player, and the associated limitations. This platform requires registration, even to access Open Data from third parties, which is allowed by some Open Data licenses not implementing the share-alike condition, the permanence which is not an issue as these data are remaining theoretically accessible from the original source. Another issue is the guarantee of permanence of the access to the data, an important concern for data citation - e.g., the data cited by (Ott 2014) - more for research than industry stakeholders. Open Data platforms, through policy and technical means such as (U.R.I) URIs, tend to better insure the data availability in the long term.

## 2.3 Open Data

There is a need to explain first what are Open Data and what are Open Data principles. Against a naive approach, Open Data is not equal to any kind of data that may be accessed for free at any moment on the Internet, without considering the potential restrictions. The actual definition is more binding and is based on few but strong pillars.

Even if the sources of Open Data are much older, rooted both in the concept of freedom of information and transparency – required component for a functional market - and in the principles of scientific inquiry (e.g. reproducibility), Open Data itself got some attention and application after the launch of an Open Data initiative by the United States government in 2009. According to the definition suggested by the Open Knowledge Foundation, “Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share-alike” (Open

Knowledge Foundation 2013). Open corporates (Gkatziaki et al. 2017) and Open Banking (Brodsky and Oakes 2017) may also be considered as subsets or cognates of Open Financial Data. Echoing the common distinction between three kinds of finance, (Davies and Perini 2016) identifies three kinds of data at large: “Data about governments; Data about companies and markets; and Data about citizens.”

The purposes of Open Data are to impact all the sectors for political, social, and economic impacts. For the latter, Open Data is a mean to foster innovation and to ensure a trusted environment leading to risk mitigation where one may re-use the data without any legal concern and so to ensure fair use. One purpose of the Open Data initiatives is also to strengthen the trust (Omole 2017) of the re-users in the data provided on the platforms, on the ways used to deliver these data and in the relationship between the data provider and data re-user;

Open Data, although it is for the moment mainly applied to public bodies, especially governmental ones, is not limited to the public sector: Open Government Data is only one subset. Indeed, there are many other applications: scientific data, data held by private bodies but related to or coming from public interest duties or activities, and private Open Data at large, where (mostly) companies commit to open their data, either to just comply with a legal binding, or to be engaged in an Open Innovation (Chesbrough et al. 2006) strategy.

#### 2.4 Use of Open Data to Get Financial Insights

A large part of the interest brought to Open Data is related to different research topics than the usual ones in financial activities. Because of its roots, Open Data literature also tends to focus on e-government (Pina et al. 2010) and on political and democratic impacts. In this perspective, financial data shall be released as a tool to ensure the transparency (Lemieux et al. 2014) and then the accountability of government (Pina et al. 2010), or its agencies (Tunney and Thomas 2015). To ensure the transparency of governments and their accountability, research focuses on the analysis of data about government contracting (Parkhimovich and Minina 2017) and government spending (Hartog and Mulder 2017) at large. Some portals, especially in the Anglo-Saxon world, are dedicated to accountability, aiming at providing citizens the means to check how their taxes are spent, this is, for example, the case described by (Paige 2017).

This trend is not only aiming at government bodies, but also at other sectors, either they are benefiting from public spending (e.g., non-profit organizations), they are regulatory asked to do so, or following their own will (Gkatziaki et al. 2017). For the case of Non-Profit Organizations (NPOs), (Marshall et al. 2016) are showing the limits of the currently public opened datasets in the UK, which structure is not suitable to get a deep understanding of how this sector is functioning, and also implying interoperability issues.

Another important research stream is related to the Linked Open Data approaches (Vafopoulos et al. 2017), which are at the crossroads of Open Data and semantic technologies, especially to suggest alignment methodologies through the use of ontologies (Najdenov et al. 2014), (Li et al. 2011). These approaches are dealing with the structures and semantics concerns about the available data and the ways to make them interoperable (Gkatziaki et al. 2017). More generally, these studies intend to enable the standardization

of financial data (Cavanillas et al. 2016) (Lemieux 2013), especially around the XBRL model, a prerequisite for the emergence of big and comparable datasets.

From the financial standpoint itself, the potential of Open Data has been examined as input for a broad range of topics. At the global scale, financial Open Data have been considered for commodities markets (Giles 2017), to analyze price volatility (Ott 2014), to make stock price prediction (Leung et al. 2014), or to characterize systemic risk (Song 2016) and predict economic crises (Palasca and Jaba 2014).

Open Data is also used to make analyses at the scale of companies' finances, for example, for monitoring companies development (Kapkaev and Sorokin 2018), or assess the credit risk (Mei 2009). A large data availability as well as the development of concepts and tools to use it (Flood et al. 2016) may lead to the emergence of new methodologies, the application of existing concepts at different scales and the development of new questions, similar to what has happened in the Sciences with the Fourth Paradigm (Hey et al. 2009). As an example, it may be illustrated by the increasing importance of the so-called macroprudential supervision (Flood et al. 2013), and Open Data is one component required to develop this trend: data availability and associated tools are creating "a natural role for a macroprudential supervisor to monitor the evolution of the counterparty network as a whole" beyond the traditional micro-prudential approach. Open financial data have also been used to assess banking performance (Akhisar and Karpak 2010) and bank failures (Canbas et al. 2005). For the latter, the authors use Open Data and show, as it is leading to the provision of more information more conveniently, its contribution to decreasing the data gathering costs – and more generally the cost of banks' monitoring, to increase the number and the quality of the risk analyses, and finally to reinforce trust.

What is showing these insightful albeit limited studies is that Open Data impacts and motivations are not limited to the principles, but that the mere application of these principles can increase the financial analysis efficiency. Still, these insights are somewhat scattered: the purpose of the next part is to design a global framework based on the current Open Data literature with the purpose to analyze the available data relevant for financial analysis.

### 3 Assessment Framework

The framework below is designed after a literature review of Open Data research. It is based on the works assessing the consistency with Open Data definition and principles (Kučera et al. 2015), on the assessment of data openness, on data quality, especially those synthesized in (Vetrò et al. 2016) and (Máková et al. 2018), and on data release prioritization, considering academic publications and reports. For most of them, these frameworks are sharing a common focus on the data publisher. What is also need is to adopt a re-user perspective, to assess the suitability of these data for downstream activities. Almost all the points are already covered by merging the insights from literature and the various frameworks already published. The main contribution of this framework is to deeper introduce the needs of the re-users and to take into account the nature and the needs related to financial data and analyses, that is why some indicators are completed and some are added. For each point, a short definition is provided as well as, when there is one, the best reference from the literature.

It may be interesting to apply metrics to get a scoring of each dataset against each criterion, for example, through Likert scales or as percentages of data matching a criterion against the whole data set. However, these metrics depend on the context, they shall have a different weight depending on the perspective of each reuser: a limited license is forbidding by design a commercial reuse; for other needs, a low number of rows or a low temporal coverage availabilities may prevent the use of certain technics, such as neural networks (Table 1).

**Table 1.** Framework to analyze financial data availability and suitability

1. Data availability	1.1. Availability; Possibility to access the relevant data; (OpenDataMonitor 2015)
	1.2. High-level usability; Kinds of processing required to use the data; (Oviedo et al. 2015)
	1.3. Cost; Data are provided for free, or at the marginal cost; (Share-PSI 2016)
	1.4. Size; Size of the data, the number of rows and dimensions as well as the is documented and fitting the requirements of a given purpose
	1.5. Raw data; The raw data are accessible
	1.6. Completeness; Data is analyzable without accessing additional information; (Charalabidis et al. 2018)
	1.7. Availability in the long-term; Data provider is ensuring the long-term availability (and accessibility) of the data through policy or technical means
2. Semantics	2.1. Vocabularies; Data are respecting a vocabulary, especially a standardized one; (W3C 2017)
	2.2. Model; Data are structured following a documented model, preferably a standardized one; (Abella et al. 2014)
	2.3. Semantic coverage; Data semantics are fitting the requirements of a given purpose; (Oviedo et al. 2015)
	2.4. Semantic granularity; Data are provided at the scale convenient for a given purpose; (Oviedo et al. 2015)
	2.5. Temporal coverage; Data are provided at the convenient time-windows for a given purpose; (W3C 2017)
	2.6. Temporal granularity; Data are provided at the convenient granularity for a given purpose; (W3C 2017)
	2.7. Spatial coverage; Data are provided at the convenient geographic area for a given purpose; (Abella et al. 2014)
	2.8. Spatial granularity; Data are provided at the convenient granularity for a given purpose; (Abella et al. 2014)

(continued)

**Table 1.** (continued)

3.	Data access and identifiers	<p>3.1. Data discoverability/findability; Data are exposed on a platform a website and indexed (see also data documentation); (Share-PSI 2016)</p> <p>3.2. Direct access to the data; Data are provided immediately and without the obligation to register; (Abella et al. 2014)</p> <p>3.3. Bulk download; It is possible to download the whole dataset in a request; (Oviedo et al. 2015)</p> <p>3.4. Single URLs; Single consistent URL for downloading data; Open Data Institute 2013)</p> <p>3.5. URI; Each data set has a unique identifier; (W3C 2017)</p> <p>3.6. (Documented) API; Data may be accessed through API; (W3C 2017)</p>
4.	Data documentation/metadata	<p>4.1. Metadata; The data are documented by a consistent and relevant set of information; (Charalabidis et al. 2018)</p> <p>4.2. Machine-readable metadata; Metadata are machine-readable; (OpenDataMonitor 2015)</p> <p>4.3. Metadata vocabularies and standards; Metadata is respecting a documented vocabulary, preferably a standardized one; (W3C 2017)</p> <p>4.4. Data provenance; The processing stages from raw data as well as the relevant stakeholders are documented at least at a high-level; (Oviedo et al. 2015)</p> <p>4.5. Data versioning; The versions are documented and remain accessible and referenceable; (W3C 2017)</p> <p>4.6. Skills; Statement on the skills (technical or conceptual) required to understand and process the data</p> <p>4.7. Feedback mechanism; Data documentation is including a mean to collect feedbacks; (Open Data Institute 2013)</p>
5.	Data formats	<p>5.1. Machine-readable formats; Data formats are machine-readable, e.g., CSV instead of a table in a PDF document; (OpenDataMonitor 2015)</p> <p>5.2. Open formats; Data formats are respecting the Open Source principles, and are documented (Abella et al. 2014)</p> <p>5.3. Standardized formats; Data format is respecting a standard; (W3C 2017)</p>

(continued)



**Table 1.** (continued)

6.	Data quality	6.1. Timeliness; Data are provided at a convenient time for a given purpose (including real-time); (Charalabidis et al. 2018)
		6.2. Accuracy; Data is accurate enough for a given purpose; (Oviedo et al. 2015)
		6.3. Consistency; Data fields and data sets of same kinds are displayed consistently; (Charalabidis et al. 2018)
		6.4. Updates; Data is up-to-date; (Oviedo et al. 2015)
		6.5. Quality documentation; Quality issues are evaluated and documented; (Open Data Institute 2013)
7.	Legal	7.1. Ownership; Owners of the data are clearly identified; (Charalabidis et al. 2018)
		7.2. Legal statements; Clear rights statement are expressed, detailing rights issues; (Open Data Institute 2013)
		7.3. Constraints; The constraints are consistent with Open Data principles, allowing re-use with the lowest barriers; (Charalabidis et al. 2018)
		7.4. Privacy; Privacy issues are clearly expressed, with the ways to overcome them; (Charalabidis et al. 2018)

#### 4 Application to the Financial Data in Luxembourg

The framework has been used for Luxembourg, considering the requirements for different financial activities from the different branches of this discipline. This framework helped to discriminate the strong and weak points for each kind of Data.

Because of the legal commitments, there is good coverage for a large part of public finance. The related data are published on the national Open Data platform, <https://data.public.lu>, which is providing most government data, that is data collected by the government during its activities, for example, the government budget is published on the national Open Data platform since 2016. However, there are no data concerning the contracts signed by the government, but only aggregated data from the budget perspective.

The legal framework – the European Directive transposed in Luxembourg - and the implementation of the best practices lead to a good quality for the indicators directly related to Open Data: the centralized platform is ensuring a convenient discovery of the data, the metadata are standardized and consistent, the formats are machine-readable and open for most of them, and there are no legal obstacles to the reuse of data. If the data are following a consistent model, they are displayed at a very aggregated level, convenient for the work of government, but not providing detailed and massive data that would be required to use tools such as those based on neural networks.

Personal finance data are showing an intermediate case. Personal finance data are, in essence, very fine-scaled data about individuals or households, and therefore constitute personal data, ruled by the privacy regulation. One of the ways to use these data can be

compared to the principles of Open Banking (Brodsky and Oakes 2017), whose objective is both to determine mechanisms that will allow individuals or households to directly benefit from the benefits of Fintechs. It is not clear however how Open Banking, which is at its very early stage in Luxembourg, will match at least partially the requirements of Open Data, or if it will be limited to the banking system stakeholders. Until now, none of these data is currently available. At the statistical scale, following its duties to monitor the economic and social conditions of the country, the national statistical agency, STATEC, is publishing data based on personal finance, but aggregated at the scale of the population. It is leading to figures based at the finest one month, making these data closer to what is considered as economic data. Like public finance data, these data are published suitably from the Open Data principles perspective. However, they are raising an issue concerning their semantic granularity, limiting both the kinds of research questions or financial services that could be built, as well as the applicable tools.

Corporate and market data are the case the least favorable case concerning Open Data. Like personal finance, some data are available at the statistical scale. Excepted these data, the information is highly fragmented, depending on the organizations in charge of their collect. The business registry is providing some amount of data for free, but these data do not fulfill the other criteria, for example, the machine-readability criterion. Stock exchange data are showing the same limitations. The platform is not directly providing high-frequency data, and if the website is displaying data that might be easily scraped, the website owner is forbidding it. It means that for most of the financial analyses in this domain, and even if the data may be accessed under different rules, the approaches may not be implemented nor reproduced while satisfying the conditions of the Open Data paradigm.

## 5 Conclusion

This study is firstly providing a methodological contribution to analyze financial data based on the Open Data paradigm and taking into account the needs of financial discipline and practice.

The case of Luxembourg is showing that there is a contrast between an industry relying on a large consumption of data and information and making it the basis of its business, and the low uptake of Open Data in this domain. Some issues are more hindering depending on the kind of user: business reusers are more affected by license restrictions than researchers who are benefiting from exceptions, some are affecting all the reusers like a bad semantic granularity. This framework intends to be useful to get a quick statement on the applicability of a methodology or a tool based on the characteristics of the available data.

Available data, by far, are not coinciding with the definition of Open Data. It calls for a better explanation of Open Data outside government, first and foremost to institutions whose duty is to provide these data, but also to companies who would like to engage in an Open Data policy, for example aiming at benefiting from Open Innovation externalities.

Unsurprisingly, the best level of data availability and quality is reached by public finance data. However, even for these data, we showed some limits, some in terms of availability, more in terms of suitability, preventing at least partially their reuse. This

may be explained from the semantic and temporal granularity perspectives: data released by the government, mostly because of their statistical nature, are at the scale of the year, whereas a lot of financial questions consider a day or intraday scale. This dissonance is preventing the use of innovative concepts, mostly because the statistical view is providing rare data from the perspective of neural network approaches, hence the need to turn to solutions suitable for this scarcity. This framework may help stakeholders in charge of the data release to better align their publications with the needs of re-users when it is possible. The identified issues are showing the interest of applying an Open Data paradigm to financial data indirectly, that may be a lever to increase the availability of data, their discoverability, and their quality.

The lack of data availability and suitability is preventing a part of the commercial uptake. The current financial ecosystem is nevertheless showing the potential of increased data availability, where the data brokers could continue and develop their business models while the long-term availability of a set of raw data would be guaranteed.

Assessing data availability and suitability is of paramount importance for Open Data re-users, who have to assess the data in a fast and reliable way, and for decision-makers for whom automation is a key concept to lower the costs induced by Open Data policies. Building on state-of-the-art works on automatic data quality assessment, further work will implement a tool allowing to score datasets against this framework, taking into account not only Open Data strictly speaking but Open Data candidates' datasets, considering both metadata and data contents.

## References

- Abella A, Ortiz-de-Urbina-Criado M, De-Pablos-Heredero C (2014) Meloda, a metric to assess open data reuse. *El profesional de la información* 23:582–588
- Akhisar I, Karpak B (2010) AHP as an early warning system: an application in commercial banks in Turkey. In: Ehrgott M, Naujoks B, Stewart TJ, Wallenius J (eds) *Multiple criteria decision making for sustainable energy and transportation systems*. Springer, Heidelberg, pp 223–233
- Boyd D, Crawford K (2012) Critical questions for big data. *Inf Commun Soc* 15:662–679
- Brodsky L, Oakes L (2017) *Data sharing and open banking*. McKinsey & Company, New York
- Canbas S, Cabuk A, Kilic SB (2005) Prediction of commercial bank failure via multivariate statistical analysis of financial structures: the Turkish case. *Eur J Oper Res* 166:528–546. <https://doi.org/10.1016/j.ejor.2004.03.023>
- Cavanillas JM, Curry E, Wahlster W (eds) (2016) *New horizons for a data-driven economy: a roadmap for usage and exploitation of big data in Europe*. Springer, Cham
- Charalabidis Y, Zuiderwijk A, Alexopoulos C, Janssen M, Lampoltshammer T, Ferro E (2018) Organizational issues: how to open up government data? In: Charalabidis Y, Zuiderwijk A, Alexopoulos C, Janssen M, Lampoltshammer T, Ferro E (eds) *The world of open data: concepts, methods, tools and experiences*. Springer, Cham, pp 57–73
- Chesbrough H, Vanhaverbeke W, West J (2006) *Open innovation: researching a new paradigm*. Oxford University Press, Oxford
- Davies T, Perini F (2016) Researching the emerging impacts of open data: revisiting the ODDC conceptual framework. *J Commun Inf* 12:148–178
- Flood MD, Lemieux VL, Varga M, Wong BW (2016) The application of visual analytics to financial stability monitoring. *J Finan Stab* 27:180–197

- Flood MD, Mendelowitz AI, Nichols W (2013) Monitoring financial stability in a complex world. In: Lemieux V (ed) *Financial analysis and risk management: data governance, analytics and life cycle management*. Springer, Heidelberg, pp 15–45
- Chen Q, Giles DE (2017) Risk analysis for three precious metals: an application of extreme value theory. Department of Economics, University of Victoria
- Gkatziazi V, Papadopoulos S, Diplaris S, Kompatsiaris I (2017) Large-scale open corporate data collection and analysis as an enabler of corporate social responsibility research. In: Kompatsiaris I, Cave J, Satsiou A, Carle G, Passani A, Kontopoulos E, Diplaris S, McMillan D (eds) *Internet science*. Springer, Cham, pp 266–280
- Hartog M, Mulder B (2017) Accountable and responsible disclosure of financial open government data: open spending initiatives enhancing civic engagement. In: *International conference ICT, society and human beings*, pp 127–133
- Hey AJG, Tansley S, Tolle K (2009) *The fourth paradigm: data-intensive scientific discovery*. Microsoft Research, Redmond
- Kapkaev I, Sorokin D (2018) The implementation of the comfortable investment environment on the basis of the effectiveness assessment of the corporation. *Procedia Soc Behav Sci* 238:697–703. <https://doi.org/10.1016/j.sbspro.2018.04.052>
- Koop G (2006) *Analysis of financial data*. Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, Chichester
- Kučera J, Chlapek D, Klímek J, Necaský M (2015) Methodologies and best practices for open data publication. In: *DATESO*
- Lemieux VL (ed) (2013) *Financial analysis and risk management: data governance, analytics and life cycle management*. Springer, Heidelberg
- Lemieux VL, Phillips P, Bajwa HS, Li C (2014) Applying visual analytics to the global legal entity identifier system to enhance financial transparency. In: *Conference presentation, conference on data standards, information, and financial stability*, Loughborough University, pp 11–14
- Leung CK-S, MacKinnon RK, Wang Y (2014) A machine learning approach for stock price prediction. In: *Proceedings of the 18th international database engineering and applications symposium*. ACM, New York, pp 274–277
- Li X, Bao J, Hendler JA (2011) Fundamental analysis powered by semantic web. In: *2011 IEEE symposium on computational intelligence for financial engineering and economics (CIFER)*, pp 1–8
- Máchová R, Hub M, Lnenicka M (2018) Usability evaluation of open data portals: evaluating data discoverability, accessibility, and reusability from a stakeholders' perspective. *Aslib J Inf Manage* 70:252–268. <https://doi.org/10.1108/AJIM-02-2018-0026>
- Magalhaes G, Roseira C, Manley L (2014) Business models for open government data. In: *Proceedings of the 8th international conference on theory and practice of electronic governance*. ACM, New York, pp 365–370
- Marković M, Gostojić S (2018) Open judicial data: a comparative analysis. *Soc Sci Comput Rev* <https://doi.org/10.1177/0894439318770744>
- Marshall M, Kirk DS, Vines J (2016) Accountable: exploring the inadequacies of transparent financial practice in the non-profit sector. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, New York, pp 1620–1631
- Mei H (2009) Research on credit risk recognizing model in enterprises of credit-sale. In: *2009 international conference on electronic commerce and business intelligence*, pp 332–336
- Najdenov B, Pejčinovski H, Cieva K, Jovanovik M, Trajanov D (2014) Open financial data from the macedonian stock exchange. In: *International conference on ICT innovations*. Springer, pp 115–124
- Omole DW (2017) *Trust in International Digital Content: Open Access and Access to Information at the World Bank*

- Open Data Institute (2013) ODI Open Data Certificate. <https://certificates.theodi.org/en/about/badgelevels>. Accessed 13 June 2019
- Open Knowledge Foundation (2013) What is Open Data? <http://opendatahandbook.org/guide/en/what-is-open-data/>. Accessed 15 June 2019
- OpenDataMonitor (2015) Methodology of OpenDataMonitor. <https://opendatamonitor.eu/frontend/web/index.php?r=site%2Fmethodology>. Accessed 8 June 2019
- Ott H (2014) Volatility in cereal prices: intra-versus inter-annual volatility. *J Agric Econ* 65:557–578. <https://doi.org/10.1111/1477-9552.12073>
- Oviedo E, Mazón JN, Zubcoff JJ (2015) Quality and maturity model for open data portals. In: 2015 Latin American computing conference (CLEI), pp 1–7
- Paige BE (2017) Open data portals in northern New England states. PhD thesis, University of British Columbia
- Palasçă S, Jaba E (2014) Leading and lagging indicators of the economic crisis. *Revista Română de Statistică* nr 62:31
- Parkhimovich O, Minina E (2017) Insight provoking multi-criterial visualization of Russian open finance data: from data acquisition to processing and analyzing. In: Proceedings of the 10th international conference on theory and practice of electronic governance. ACM, New York, pp 552–553
- Pereira GV, Virkar S, Vignoli M (2018) Exploring the political, social and cultural challenges and possibilities associated with trading data: the case of data market Austria (DMA). ACM, p 128
- Pina V, Torres L, Royo S (2010) Is E-Government promoting convergence towards more accountable local governments? *Int Public Manage J* 13:350–380. <https://doi.org/10.1080/10967494.2010.524834>
- Share-PSI (2016) Best Practices for Sharing Public Sector Information. <https://www.w3.org/2013/share-psi/bp/>. Accessed 6 June 2019
- Song JW (2016) Systemic risk measures based on open financial data. PhD thesis, Seoul National University
- Tsay RS (2010) *Analysis of financial time series*. Wiley, Hoboken
- Tunney S, Thomas J (2015) Public access to NHS financial information: from a freedom of information regime to full open-book governance? *Soc Theor Health* 13:116–140. <https://doi.org/10.1057/sth.2014.19>
- Vafopoulos M, Koukourikos A, Vafeiadis G, Negkas D, Skaros I, Tzani A (2017) A data-driven model for linking open economic information. In: Kompatsiaris I et al (eds) *Internet science*. Springer, Cham, pp 329–343
- Vetrò A, Canova L, Torchiano M, Minotas CO, Iemma R, Morando F (2016) Open data quality measurement framework: definition and application to open government data. *Gov Inf Q* 33:325–337. <https://doi.org/10.1016/j.giq.2016.02.001>
- W3C (2017) Data on the Web Best Practices. <https://www.w3.org/TR/dwbp/>. Accessed 7 June 2019