



Open Data Research Challenges in the EU

Csaba Csáki^(✉)  and Andrea Kő 

Corvinus University of Budapest, Budapest, Hungary
{Csaki.Csaba, Andrea.Ko}@uni-corvinus.hu

Abstract. Although the concept of ‘open data’ has been around for some time, it has gained more attention in the public sphere over the last decade. This momentum has attracted the interest of ‘mainstream’ e-Government researchers as well, what is well indicated by the increased number of papers addressing the topic. Even though this interest seemed to be declining two-three years ago, the idea of reusing public sector data provided new ammunition to do research into this phenomenon. This paper relies on the latest literature and relevant studies to review the research area of open government data and to offer recommendations regarding potential research directions that could advance our understanding of the challenges facing the reuse of public sector information.

Keywords: Open data · Open government data · Open data reuse
Open data research agenda · Open data research questions

1 Introduction

Providing ‘open data’ has become a key element in the e-Government arsenal in support of transparency and accountability [8, 22]. Its primary purpose is to make available specific sets of data and information produced by various public sector entities or controlled by governmental organizations typically through Internet outlets [29]. One typical example of open government data (OGD) is data related to public expenditures and procurement – including calls for tender, contracts, purchase items and prices as well as general spending data – published both during purchasing processes and retrospectively. But there is a large amount of data related to other areas such as data generated during various legal, official or administrative processes or while executing various government tasks and functions. Beyond their primary purpose (i.e. being used during public sector processes) these data sets are worthy of additional value generation [23]. Reusable public sector data may be related to economic, social, societal, demographic or health matters, but certain legal, judicial, or property/real estate registry data may also have potential for added value [3]. Utilizing public sector data in innovative, marketable services has become a successful practice in several countries over the last few years [29]. Accordingly, scientific interest in this area has increased as well [52]. There have been several models and interpretive frameworks put forward or had been reused – such as the application of the ecology metaphor to address the complexity of the open data arena in general and to describe the relationships among its actors in particular [21]. Although there are various research efforts moving along different directions to explore the questions related to the reuse of OGD

[46, 52], it is still not clear what the important trends and open questions are for the next decade. This paper attempts to propose important and promising research areas and questions within the open data domain.

The paper is organized as follows. After this introduction the paper reviews the history of OGD in an international context with special focus on the European Union. This is followed by a review of the most popular and relevant scientific models and frameworks. After the discussion of our research approach and methods, the fourth section starts with a presentation of research areas based on most recent literature. This is followed by our own proposal of areas concluded from an analysis of key research papers. The main part of the discussion concludes with specific research questions proposed for the areas identified. The paper concludes with a summary of the most relevant theoretical findings and practical recommendations.

2 Open Government Data

2.1 A Brief History of Open Government Data

The idea to push for open government and to make public sector data available is not new. Indeed, in the US it was already raised in the 50s of the last century that the government should be ‘open’ – at least in a legal sense [37]. The primary goal was to achieve better accountability, and according to the argument this required that government data should be more accessible as well. This expectation can be found in the principles of freedom of information or in the legal (and constitutional) requirements of ‘right to information’ [22], which was the dominant motivation to request access to public sector information or data ‘owned’ by state or government entities. Over the last two decades – on the back of the spreading use of the Internet – the idea of open data has gotten new fuel from the technology backed e-Government initiative and during the turn of the millennium governmental (data) portals were created in several countries [42]. Strictly speaking the term ‘open data’ as an expression with special meaning may be dated from the 2006 manifesto of the Open Knowledge Foundation [7, 35], although that call was mainly a generic proposal, as it also concerned scientific and other data. Data may be called ‘open’ if it is freely accessible in machine readable format and it is (legally) free to be used, reused, or redistributed for any purpose [35] – typically assuming that the source is attributed and the results are shared [29]. Open government data as a special area on its own emerged towards the end of the first decade of this century when it was brought into focus as part of the (rejuvenated) open government movement [2, 34]. The core of the open government concept is that citizens have the right to get access to data, information and documents generated by governments as well as to public sector procedures involved [8]. Over the last decade more and more countries have initiated their own open government program and a result the number of accessible datasets has increased considerably [3, 7, 42]. At the same time – and this is especially true for data possessed by governments and their institutions – the real value of open data lies in their further reuse and utilization, mainly because OGD makes several commercial service-oriented endeavours possible, even for businesses that don’t own any data [29]. This implies that commercial utilization gets higher emphasis on top

of strengthened social, societal, and political requests, with the latter aimed at increasing the participation of citizens in democratic and governmental decision making processes. In the context of the European Union, the Public Sector Information (PSI) program of the European Commission started in 2003 [18] initiated a push for OGD publication, while the Digital Agenda initiative of 2010 [15, 16] has moved towards a framework encouraging the socio-economic utilization of data such published.

2.2 Open Data and Its Reuse in the European Union

The first significant appearance of the open data concept at the Union level came as early as the late 90s as the Commission already recognized the secondary value of public sector information and named data to be a key resource [14]. Later, in the so called Public Sector Information (aka PSI) directive (Directive 2003/98/EC) the Commission encouraged member states to make public sector information available for re-use by third parties as much as possible [18]. Prior to this legal statement the question of data openness was left to the member states to regulate. The directive aimed to catalyse the development of new services through providing public sector data at low price with supportive conditions [22]. The European Commission promoted open data initiatives again in its Digital Agenda for Europe program initiated in 2010 [14, 17]. The 2013 amendment (Directive 2013/37/EU) broadened the scope of the directive with the “*open data, unless*” standpoint [19]. The expectation was that the availability of public data stimulate the secondary use of such data, which not only promotes government transparency but supports information industries as well. The potential value that may result from the re-use of open public sector information in Europe is huge: it is estimated to be between €27 billion [9] and €68 billion [40, 47].

However, there are roadblocks to fulfil this potential. For example, van Loenen et al. [47] call the related EU *data protection legislation* a “*very hungry caterpillar*”, which cause problems for the successful execution of the EU digital agenda through obstructing the implementation of open government data policies for mapping data in the EU. Furthermore, Zódi [51] – while reviewing the implementation of the PSI directive in Hungary – identifies additional factors that could hinder the success of the directive. He considers copyright issues, proprietary data formats, and overpricing as serious challenges to overcome, and also adds, that it is difficult to calculate marginal costs. He notes, that public sector entities as data owners has no motivation to share their data or, when they enter the market on their own, they have an unfair advantage.

Several aspects of the Open Government Data area has been addressed during the EGOVIS conference series [24–27]. Martin et al. [30] present open data ecosystem approach implemented in BE-GOOD European program. BE-GOOD is an Interreg VB NWE project aiming to unlock, re-use and extract value from Public Sector Information (PSI) to develop data driven services in the area of infrastructure and environment. The authors developed a new open data ecosystem framework, which is based on the analysis of existing open ecosystem models. They introduced a new role called stimulator and a new stimulating function into the open data ecosystem concept. The main specificity of the stimulator function is that it involves thinking about and influencing the ecosystem. The stimulating function then has a decisive role in risk management

within the ecosystem. The new approach was customized for public procurement context.

Palmirani et al. [36] discussed the Open Government Data legislation framework in force in the Italian legal system. Their paper provide an overview of an empirical research conducted on Italian Municipal web sites (covering 35 portals) to investigate the connection between the Open Government Data legislation and the Italian Transparency Act.

Schmitz et al. [43] presented a pilot project on Linked Open Data (LOD) and e-Participation, promoted by the European Parliament and developed by the Publications Office of the European Union (OP). They detailed the main features of LOD and an e-Participation platform based on open source and semantic web technologies. The main goal of the project was to allow citizens to actively participate in public consultations within the EU decision-making process. Their solution gives citizens the possibility to participate in the preparation of documents throughout the law-making process, for example participants may make comments and amendments on each document fragment, or express their sentiment on them.

Hansen et al. [20] analyzed the background, extent and expected impact of the Danish open government data initiative. Their research focused on the role of open public sector information as a major step towards a digital society. They applied the principles of the Open Government Data initiative as a discussion framework for the Danish approach to open government data. They draw attention to the observation that open government data is just one factor in promoting innovation, while human resources, like skilled specialists and researchers, entrepreneurship, and venture capital are perhaps more important.

According to the literature review, researchers discuss open data and the related issues mainly from policy or technical point of view. A few papers deal with other considerations, like organizational challenges, but a holistic view of related potential research questions and problems is missing.

3 Frameworks and Models Proposed for the Research of Governmental Open Data

It was put forward already in the 90s that the process and activities of managing (organizational) data may be compared to the manufacturing and logistics processes of physical products [49]. This lead to the rise of the ‘data supply chain’ concept that was built around the production-delivery-consumption metaphor of creating, recording, storing, and using data.

The first data-transparency solutions appearing under the e-Government banner handled the issue mainly from a technical point of view and offered ‘platforms’ where (certain) governmental data could be published. This actually meant a one directional approach. The next step, still rooted in the technical approach to e-Gov, created ‘portals’ that typically offered APIs (interfaces) which allowed an avenue to pose queries in a specific language. This was followed by the more interactive 2.0 solutions that allowed for feedback as well [10, 11].

The main goal of the data supply chain approach was to enable the application of the quality assurance models developed for manufacturing processes [50]. This idea has resurfaced in the context of OGD almost two decades later in the work of Groth [13] who considered the important questions of who is responsible for the quality (and problems) of data and how to properly manage the sources of data collection. In this regard the data supply chain starts with the creation of data which then can be passed on and may be combined with other data (or datasets). In addition, data may go through various transformations until it gets to its final user [28].

Combining different types of data and data coming from differing sources forms the basis of the Big Data approach, where this combination of types and sources contributes to the creation of added value [31]. This gives way to the application of the value chain metaphor to data, similarly to the value chain of industrial production [41]. Indeed, the data value chain model fits well with the question of open data reuse, since the final goal of OD reuse is the creation of socio-economic value through the development of OD-based new, innovative services. One potential criticism of both the data supply and data value chain is that they consider the movement of data along a linear model. Consequently, there are arguments for a more life-cycle like approaches to the understanding of the nature of data use – including open data (see for example [38]). The main contribution of the data life-cycle models is that the producers of data are consumers of it at the same time and vice versa – from a different point of view (considering a different source). To simplify, one may say that while the supply chain and life-cycle models of data focus on the connections between suppliers and consumers (supply and demand), the value chain interpretation considers the context as well and its focus is on the process/activities of transforming data elements in order to produce higher level information in support of a given goal.

While the static platform and portal solutions of e-Gov (mentioned earlier) allow for the publication of data primarily from the point of view of public sector actors (as a responsibility or legal-regulatory expectation), the option to reuse open data brings into the picture a few new actors on the ‘consumer’ side in order to generate higher added value [8]. Such roles include data providers (participants who provide better – easier, more organized – access to open data), the data cleansers, and service developers among others [29]. One should also consider that the added value is often not the final goal, the impact achieved is more important, which might manifest itself in the form of economic advantage or social well-being [7]. Trying to understand these more complex roles lead researchers to the application of the ‘ecosystem’ approach to open data, and this model has gained momentum over the last few years [52]. So much so, that now it is a dominant stream in open data research papers [45]. Considering the origins of the ecosystem approach to data, it is rooted in data (and information systems) ‘ecology’, but it must be noted that the original literature using the ecosystem model to describe relationships in the open data area did not provide a clear definition (or none at all) and there is no common, accepted notation how to depict open data ecosystems (roles, processes and relationships in them). This resulted in the diverging set of building blocks (and corresponding notations) used by various authors building on the ecosystem metaphor to explain various aspects of the open data phenomenon. The principles of ecology were already applied during the turn of the millennium to understand and explain the issues of organizational information sharing [12, 33].

Ecology is the “*scientific study of the processes influencing the distribution and abundance of organisms and the transformation and flux of energy and matter*” ([12], p. 74). In this regard information ecology means “*a system of people, practices, values, and technologies in a particular local environment*” ([33], p. 49) and studying information ecologies implies the description and understanding of the elements of such systems and their relationships. Therefore, the goal is to identify and describe roles, tasks, and relationships as well as to show their change and evolution over time in a given environment. In their forward looking book Davenport and Prusak [6] identified the following elements of an information ecology: outside environment, organizational environment, information environment and within them stakeholders, strategy, culture and behaviour, principles and rules, processes, and finally (technological) architecture. This vision is embraced by the ecosystem approach, which for open data first has been used by Parsons et al. [38]. Based on the data lifecycle model they defined an ‘information ecosystem’ as “*the people and technologies collecting, handling, and using the data and the interactions between them*” (p. 557). This metaphor has since been appropriated and applied by others. As a critique it should be noted that neither the roles and actors of an (open data) ecosystem nor its processes has achieved an acceptable level of standardization yet, and there is no widely accepted framework.

Considering the literature of open data research (with its dominant organizational interest and diverging focus) as well as the current state of the art of open data models (including the uncertainty surrounding the various interpretations and models), it appears useful to establish a holistic research framework that offers clear and well defined (sub)areas and allows for the posing of relevant questions worthy of scientific interest.

4 Methodological Considerations

4.1 Research Questions

According to the aim of the research and based on the literature reviewed, three research questions have been formulated: (1) Which frameworks and corresponding holistic dimensions would be relevant in restructuring the main research areas identified by the literature review? (2) What are the decisive research areas in the “open data” research domain? (3) What are the recent, important open research questions in the European “open data” research? The rest of the paper will address these research questions and provide answers for them in Sect. 6 – based on the methodological approach discussed in the next Subsection.

4.2 Research Strategy: Systematic Analysis of Literature

The selected research method is literature review based, as suggested by vom Brocke et al. [48]. Their framework for literature reviewing has five phases: (1) definition of review scope; (2) conceptualization of topic; (3) literature search; (4) literature analysis and synthesis; (5) research agenda. The first step (definition of review scope) is a critical one as it determines the subsequent phases. To clarify the definition of review

scope, Cooper's taxonomy [5] could be applied. It has six dimensions: research focus, goal, organization, perspective, audience and coverage. The vom Brocke research framework was applied in the following way. The review scope is open data domain. In the first step, the Cooper taxonomy was used as well. Research focus is the overview of open government data related research outcomes. The goal is a critical but reconstructive review of the related literature. Issues considered were mainly conceptual, while the perspective followed is a neutral representation. Target audience includes scholars and practitioners, and the coverage is representative (in the sense of the vom Brocke approach, as the whole corpus of literature is represented through a sample selection). The second step of vom Brocke framework is conceptualization of topic, which can be performed by using terminology, taxonomy or ontology. They suggest collecting the key terms in this step and defining them, which was covered in the second and third sections above. The third step is the search of the literature. A search of English language articles and books was executed in both the Scopus database and in Google Scholar using "open government data" (483) in conjunction with each of the following terms (the number inside the parentheses indicates hits in Scopus – double checked in Scholar): "literature review" (89), "research agenda" (98), "taxonomy" (24), "overview" (84), "history" (32) or "research framework" (2). From the resulting pool those articles from the last fifteen years were kept which: a) presented research questions and orientations; b) provided terminology or taxonomy; c) dealt with historical overview; d) included literature review; e) discussed the regulatory environment. The original pool consisted of 127 items which was manually reduced to 10. In the literature analysis and synthesis step two experts processed the articles and structured them according to the aspects and dimensions detailed in Table 1.

5 Research Areas in the Open Data Literature

Scientific papers fitting the above conditions consider open data related research questions from the point of view of several research fields. During the last fifteen years or so there have been numerous attempts to review the history of the open data area and to sketch its potential future.

Arzberger et al. [1] were among the first to address the questions of open data reuse scientifically. Their study was supported by OECD and investigated the opportunities arising from opening up research results financed through public funds. It recommended five areas in connection to the accessibility of public sector data: technological, institutional and managerial, financial and budgetary, legal and policy, and cultural and behavioural. Harrison et al. [21] put the concept of open government into the centre of their research and claimed four areas as important: policies and practices, users (civil society and business), technology and innovation, as well as context (which may include legal, regulatory, and economic environments). Within these areas they proposed seven topics to be relevant in relation to open data: the process of identifying data of interest, setting priorities for data collection, collecting data, publishing data, utilizing data, value creation, and sustainability. Lindman et al. [29] focused their research efforts on understanding the services built on open data and their research proposal listed challenges grouped into seven categories: information, technologies,

Table 1. Open data research areas in relevant literature

Areas	Papers										Areas of this paper (Sect. 6)
	Arzberger et al. [1]	Harrison et al. [21]	Lindman et al. [29]	Munk et al. [32]	Zuiderwijk et al. [52]	Davies and Perini [7]	Charalabidis et al. [4]	Kankanhalli et al. [23]	Styrin et al. [45]	Susha et al. [46]	
History of open data						+					f
Politics				+	+	+					a
Policy, regulatory, and legal (environment)	+	+		+	+	+	+		+	+	a
Financial and budgetary	+										a
Organization, management (e.g. readiness, motivation)	+			+	+	+	+			+	b
Culture and behaviour (preparedness)	+	+				+				+	b
Practice and execution (operations)		+			+	+			+		b
Processes and activities								+			b
Participants and roles (stakeholders, collaboration)		+	+	+					+	+	b, e
Users, consumers (use)			+		+		+				e
Technology and infrastructure (standards)	+	+	+		+		+	+			c
Information			+							+	c
Data management									+	+	c
Products and services			+								d
Innovation					+						d
Impact, value, profit						+	+			+	e
Society					+						e
Theory				+	+			+			f
Domains, applications								+			d, e

processes and activities, products and services, participants, customers, and environment. While investigating open data related innovation Zuiderwijk et al. [52] – while reviewing relevant literature – proposed seven research perspectives: legislative, political, social, economical, institutional, operational, and technical. They identified three main research directions based on these perspectives: theory and development; rules, use, and innovation; as well as infrastructure and technologies. Davies and Perini [7] investigated the impact of open data initiatives and identified four areas on which to concentrate their efforts: the history of open data, evaluating readiness, implementation case studies, and impact analysis. Charalabidis et al. [4] reviewed several research programs (among them four of the above five) and constructed thirty-five topics under four umbrella areas (management and policies, infrastructures, use and value, and interoperability). According to Styrin et al. [45] there are three focal points within the open data domain, namely government policy and practice, data management, and handling stakeholders. One of the latest open data research program has been put together by Kankanhalli et al. [23] where they put forward three research directions: domain-specific studies, investigating the application of tools, and finally theoretical foundation and research methodologies used. The study recently published by Susha et al. [46] considered the collaboration between sectors which lead them to propose a taxonomy for the open data domain. This taxonomy introduces fourteen dimensions (within two groups): data sharing (in it type, content, administrative level, diversity of

data providers, support, and the level of access) and data usage (target audience, user selection, policy problems, incentives, continuity of collaboration, outcome, collaboration among users, and purpose of use). In the context of *Europe* Munk et al. [32] reviewed research challenges of open data in Hungary within the framework of the European Union directives and phrased questions grouped into three areas: conceptual challenges and questions of interpretation; the complex relationships of national and union level regulations; and the analysis of application areas (this latter including questions of semantics and technology, among others).

The range of areas appearing in the above papers has been summarized in Table 1 (augmented with the coding of the areas proposed later in this paper).

6 Research Areas Proposed – and Research Questions Identified

Considering the research areas and directions discussed by the literature and the dominance of technical and organizational issues in research outcomes, we suggest a more holistic approach in research areas discussion, which include the following elements:

- (a) *Context*: considers policy, regulations, legal background, and other environmental elements such as governance (including non-organizational public service issues);
- (b) *Organizational aspects of the public sphere*: participants, roles, decisions, processes, and other organizational issues belong here – as well as specific case studies, country status reports and related analysis;
- (c) *Technology and data*: this area covers platforms, standards, data typology, questions of data quality, frameworks of data quality assessment, availability (scope of data accessible), usability, and linked open data (including issues related to provenance);
- (d) *Reuse*: this addresses the (direct) utilization of open data, questions of innovation, and value added services;
- (e) *End users*: this not only includes users of open data, but covers the investigation of (actual) societal and economic impact;
- (f) *Theory*: discussions over theoretical foundations, questions of terminology, modelling issues, and historical overview belong here – not to mention research agendas.

For each area above this research proposes a few nagging, important or timely questions – considering both the literature as introduced and discussed above as well as the analysis of open data in the EU.

Context: (1) What are the most important elements of the regulatory environment and how are they connected (e.g. what areas are being covered and do they overlap; how the Union/Commission level relates to national laws)? (2) What are the main IT applications identified by the regulatory environment as being key in supporting end-users?

Organizational aspects of the public sphere: (1) What approaches and solutions have been applied in EU countries for open government data management and what

best practices exist? (2) What would be a suitable maturity model to compare these practices?

Technology and data: (1) What are the typical open data architectures, and what would be their advantages and disadvantages? (2) What are the relevant standards and how are they related? (3) How can end-user services and applications support the wider utilization of semantic technologies? (4) How can the quality of open government data be measured? (5) How would it be possible to identify the reasons of open data quality problems and how can those problems be resolved?

Reuse: (1) What reusability models exist and how can they be evaluated? (2) How added value may be captured and measured? (3) How can we interpret ‘innovation’ in this context?

End users: (1) What societal impacts may be identified? (2) How would it be possible to increase the efficiency and effectiveness of the results in order to provide better services using open data (this relates to the technological and organizational areas too)?

Theory: Open data related literature is conceptually rich, but these concepts have different definitions, interpretations in some cases and not always have been applied in the same manner. There are projects aimed at developing a common terminology, taxonomy or ontology of this domain (e.g. [39] or [44] and see also <http://data.europa.eu/euodp/en/linked-data>), but these projects are isolated and don’t take into account national specialities. Related research questions: (1) What is the most suitable research methodology when developing terminology, taxonomy, or ontology for OGD? (2) How can we evaluate the existing terminologies, taxonomies and ontologies and what is their ‘quality’? (3) What are the options to integrate the existing terminologies, taxonomies and ontologies? (4) How can we customize and utilize the existing terminologies, taxonomies and ontologies in local/national environments?

7 Summary and Further Research

The growth in the ‘data industry’ has been explosive over the last few years. Indeed, the majority of digital information stored today has been produced over the last few years. Accordingly, more and more research projects put data related challenges into the centre of their interest – and the question of open government data is no exception. The goal of the study presented here was to provide an overview of the OGD research areas based on most recent publications, structure these areas, and offer recommendations regarding potential future research directions. This paper reviewed the theoretical background of governmental open data and presented the most important interpretive frameworks and models of this field. Analysing key documents and articles related to open data in the European Union (at the Union level) highlighted the dominance of policy and technology related approaches, but also identified the trend towards an end-user perspective (mostly through an organizational focus). However, beyond these frames, there appears to be no holistic handling of this important field end-to-end. Therefore, this paper attempted to provide such a holistic research picture. As a key contribution, Sect. 5 has presented an overall set of criteria how the OGD field is segmented according to literature. This has led to a new, structured set of

research areas, with specific research questions posed in each. These questions were formulated based on the relevant literature with the aim of enriching the area's potential to deliver new insights. Within the open data research areas identified, there is an increasing interest towards the questions of reuse and societal impact. Consequently, this study has expanded on the analysis of the open data field and reframed the related research efforts. Regarding future direction, our own interest is primarily focused on questions of OD theory, the application of the ecosystem approach and the challenges of improving the quality of open data.

References

1. Arzberger, P., et al.: An international framework to promote access to data. *Science* **303** (5665), 1777–1778 (2004)
2. Bates, J.: The strategic importance of information policy for the contemporary neoliberal state: the case of open government data in the United Kingdom. *Gov. Inf. Q.* **31**(3), 388–395 (2014)
3. Center for Open Data Enterprise: Open Data Impact Map Report (2016). <http://opendataenterprise.org/map/reports/May2016Report.pdf>. Accessed 3 Mar 2018
4. Charalabidis, Y., Alexopoulos, C., Loukis, E.: A taxonomy of open government data research areas and topics. *J. Organ. Comput. Electron. Commer.* **26**(1–2), 41–63 (2016)
5. Cooper, H., Hedges, L.V., Valentine, J.C. (eds.): *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation, New York (2009)
6. Davenport, T.H., Prusak, L.: *Information Ecology: Mastering the Information and Knowledge Environment*. Oxford University Press, New York (1997)
7. Davies, T., Perini, F.: Researching the emerging impacts of open data: revisiting the ODDC conceptual framework. *J. Community Inform.* **12**(2), 148–178 (2016)
8. Dawes, S.S., Helbig, N.: Information strategies for open government: challenges and prospects for deriving public value from government transparency. In: Wimmer, M.A., Chappelet, J.-L., Janssen, M., Scholl, H.J. (eds.) *EGOV 2010*. LNCS, vol. 6228, pp. 50–60. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14799-9_5
9. Dekkers, M., Polman, F., te Velde, R., de Vries, M.: *Measuring European public sector information resource. Final report of study on exploitation of public sector information - Benchmarking of EU framework conditions* (2006)
10. DiNucci, D.: Fragmented future. *Print* **53**(4), 32 (1999)
11. De Kool, D., Van Wamelen, J.: Web 2.0: a new basis for e-government? In: *Proceedings of the 3rd International IEEE Conference on ICTTA*, pp. 1–7 (2008)
12. Fedorowicz, J., Gogan, J.L., Ray, A.W.: The ecology of interorganizational information sharing. *J. Int. Inf. Manag.* **13**(2), 73–85 (2004)
13. Groth, P.: Transparency and reliability in the data supply chain. *IEEE Internet Comput.* **17** (2), 69–71 (2013)
14. EC - European Commission: *Public sector information: a key resource for Europe*, COM (98)585, Brussels (1999). <http://aei.pitt.edu/archive/00001168/>. Accessed 4 June 2018
15. European Commission: *A digital agenda for Europe*, Communication to the European Parliament, the Council, the European Economic and Social Committee, and the Committee for the Regions, COM 245 final (2010)
16. European Commission: *Digital Agenda: Turning government data into gold* (2011). <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1524&format=HTML&aged=0&language=EN&guiLanguage=en>. Accessed 20 May 2015

17. European Commission: Open data: an engine for innovation, growth and transparent governance, Communication to the European Parliament, the Council, the European Economic and Social Committee, and the Committee for the Regions, COM 882 final (2011)
18. European Parliament and Council: Directive 2003/98/EC of the European Parliament and of the council of 17 November 2003 on the re-use of public sector information. OJ L, 345, 90 (2003)
19. European Parliament and Council: Directive 2013/37/EU of the European Parliament and of the council of 26 June 2013 amending Directive 2003/98/EC on the reuse of public sector information. OJ L, 175, 1 (2013)
20. Hansen, H.S., Hvingel, L., Schröder, L.: Open government data – a key element in the digital society. In: Kő, A., Leitner, C., Leitold, H., Prosser, A. (eds.) EGOVIS/EDEM 2013. LNCS, vol. 8061, pp. 167–180. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40160-2_14
21. Harrison, T.M., Pardo, T.A., Cook, M.: Creating open government ecosystems: a research and development agenda. *Future Internet* 4(4), 900–928 (2012)
22. Janssen, K.: The influence of the PSI directive on open government data: an overview of recent developments. *Gov. Inf. Q.* 28(4), 446–456 (2011)
23. Kankanhalli, A., Zuiderwijk, A., Tayi, G.K.: Open innovation in the public sector: a research agenda. *Gov. Inf. Q.* 1(34), 84–89 (2017)
24. Kő, A., Leitner, C., Leitold, H., Prosser, A. (eds.): *Technology-Enabled Innovation for Democracy, Government and Governance*, vol. 8061. Springer, Heidelberg (2013). <https://doi.org/10.1007/978-3-642-40160-2>
25. Kő, A., Francesconi, E. (eds.): *Electronic Government and the Information Systems Perspective*, vol. 8650. Springer, Heidelberg (2014). <https://doi.org/10.1007/978-3-319-10178-1>
26. Kő, A., Francesconi, E.: *Electronic Government and the Information Systems Perspective*, vol. 9265. Springer, Heidelberg (2016). <https://doi.org/10.1007/978-3-319-22389-6>
27. Kő, A., Francesconi, E. (eds.): *Electronic Government and the Information Systems Perspective*, vol. 10441. Springer, Heidelberg (2017). <https://doi.org/10.1007/978-3-319-64248-2>
28. Li, P., Wu, T.Y., Li, X.M.: Constructing data supply chain based on layered PROV. *J. Supercomput.* 73(4), 1509–1531 (2017)
29. Lindman J., Rossi, M., Tuunainen, V.: Open data services research agenda. In: *Proceedings of HICSS-46*, pp. 1239–1246 (2013)
30. Martin, S., Turki, S., Renault, S.: Open data ecosystems. In: Kő, A., Francesconi, E. (eds.) EGOVIS 2017. LNCS, vol. 10441, pp. 49–63. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64248-2_5
31. Miller, H.G., Mork, P.: From data to decisions: a value chain for big data. *IT Prof.* 15(1), 57–59 (2013)
32. Munk, S., Fleiner, R., Micsik, A., Sikolya, Zs., Nyáry M.: Kapcsolt nyílt kormányzati adatok és kutatásuk keretei Magyarországon. *Pro Publico Bono – Magyar Közigazgatás*, vol. 4, pp. 144–169 (2014, in Hungarian)
33. Nardi, B.A., O’Day, V.: *Information Ecologies – Using Technology with Heart*, chapter 4, pp. 49–57. MIT Press, Cambridge, London (1999)
34. Obama, B.: (U.S. Executive Office): Open government directive (2009). http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m2010/m10-06.pdf. Accessed 7 Aug 2017
35. Open Knowledge Foundation: Open Knowledge Definition (2006). <http://www.openknowledge.org/>. Accessed 7 Sept 2017

36. Palmirani, M., Martoni, M., Girardi, D.: Open government data beyond transparency. In: Kő, A., Francesconi, E. (eds.) EGOVIS 2014. LNCS, vol. 8650, pp. 275–291. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10178-1_22
37. Parks, W.: The open government principle: applying the right to know under the constitution. *George Wash. Law Rev.* **26**(1), 1–22 (1957)
38. Parsons, M.A., et al.: A conceptual framework for managing very diverse data for complex, interdisciplinary science. *J. Inf. Sci.* **37**(6), 555–569 (2011)
39. Pattuelli, M.C., Provo, A., Thorsen, H.: Ontology building for linked open data: a pragmatic perspective. *J. Libr. Metadata* **15**(3), 265–294 (2015)
40. Pira International Ltd., University of East Anglia, and KnowledgeView Ltd.: Commercial exploitation of Europe’s public sector information. Final report for the European Commission Directorate General for the Information Society (2000)
41. Porter, M.E.: *Competitive Advantage: Creating and Sustaining Superior Performance*. Simon and Schuster, New York (1985)
42. Sáez Martín, A., De Rosario, A.H., Del Carmen Caba, M.: An international analysis of the quality of open government data portals. *Soc. Sci. Comput. Rev.* **34**(3), 298–311 (2016)
43. Schmitz, P., et al.: Linked open data and e-Participation in the EU law-making process. In: Kő, A., Francesconi, E. (eds.) EGOVIS 2016. LNCS, vol. 9831, pp. 79–89. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44159-7_6
44. Southwick, S.B., Lampert, C.K., Southwick, R.: Preparing controlled vocabularies for linked data: benefits and challenges. *J. Libr. Metadata* **15**(3–4), 177–190 (2015)
45. Styrin, E., Luna-Reyes, L.F., Harrison, T.M.: Open data ecosystems: an international comparison. *Transform. Gov.: People Process Policy* **11**(1), 132–156 (2017)
46. Susha, I., Janssen, M., Verhulst, S.: Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: taxonomy development. In: *Proceedings of the 50th Hawaii International Conference on System Sciences*, pp. 2691–2700 (2017)
47. van Loenen, B., Kulk, S., Ploeger, H.: Data protection legislation: a very hungry caterpillar: the case of mapping data in the European Union. *Gov. Inf. Q.* **33**(2), 338–345 (2016)
48. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the giant: on the importance of rigour in documenting the literature search process. In: *Proceedings of ECIS 2009, Paper 161* (2009)
49. Wang, R.Y., Storey, V.C., Firth, C.P.: A framework for analysis of data quality research. *IEEE Trans. Knowl. Data Eng.* **7**(4), 623–640 (1995)
50. Wang, R.Y., Strong, D.M.: Beyond accuracy: what data quality means to data consumers. *J. Manag. Inf. Syst.* **12**(4), 5–33 (1996)
51. Zódi, Zs.: Elvek és valóság Európában: a PSI-irányelv érvényesülésének problémái. In: Sály, P. (Szerk) *Jogtörténeti és jogelméleti tanulmányok*. Miskolc, Miskolci Egyetemi Kiadó, pp. 174–185 (2015 in Hungarian)
52. Zuiderwijk, A., Helbig, B., Gil-García, J.R., Janssen, M.: Innovation through open data - a review of the state-of-the-art and an emerging research agenda. *J. Theor. Appl. Electron. Commer.* **9**(2), 1–8 (2014)