Big Data in the Public Sector. Linking Cities to Sensors

Marianne Fraefel[(✉)], Stephan Haller [iD], and Adrian Gschwend

Bern University of Applied Sciences, Bern, Switzerland
{marianne.fraefel,stephan.haller,adrian.gschwend}@bfh.ch

Abstract. In the public sector, big data holds many promises for improving policy outcomes in terms of service delivery and decision-making and is starting to gain increased attention by governments. Cities are collecting large amounts of data from traditional sources such as registries and surveys and from non-traditional sources such as the Internet of Things, and are considered an important field of experimentation to generate public value with big data. The establishment of a city data infrastructure can drive such a development. This paper describes two key challenges for such an infrastructure: platform federation and data quality, and how these challenges are addressed in the ongoing research project CPaaS.io.

Keywords: Big data · Internet of things · Open government data · Linked data · Public sector · Smart city · Data quality · Platform federation

1 Introduction

The digitization of the economy and society becomes apparent in the many applications and devices that use and produce data. As businesses and individuals use available technological innovations to improve business and facilitate the demands of everyday life, governments around the world are struggling with how to best put these advancements to use in the public sector and create public value. The European Commission for example has acknowledged that "data has become an essential resource for economic growth, job creation and societal progress" [1] and is working on a policy and a framework for the free flow of data to reap the potential benefits and address challenges both in the technical as well as in the societal and legal fields. One of the primary difficulties lies in the diversity and the speed of the technological developments: The deployment of sensors delivers a multitude of new data sets, but with sometimes unreliable data, big data and machine learning is deployed for data analysis, and linked data and open government data approaches are used to make data more accessible to a wider clientele. The societal challenges that digitization will bring about manifest themselves first in the metropolitan, urban environment; hence the so-called "smart city" is an ideal field for experimentation to better understand and learn about the opportunities and potential pitfalls. While the term "smart city" is certainly hyped and many different activities are carried out under this label, two points are interesting when looking at cities that are generally regarded as pioneers in the field, e.g., Amsterdam, Barcelona, or Vienna.

They all use a private-public-partnership (PPP) model in order to bring together actors from different sectors and with different expertise and interests [2, 3], and the establishment of a platform for information exchange and data access is seen as a key enabler for an effective implementation of a smart city programme [4].

With the main goal of developing such a platform for smart city innovation, we launched in 2016 the CPaaS.io project.[1] In this 30-month research and innovation action between Europe and Japan, data from various sources are made accessible via a cloud-based platform to application developers and service providers. Data sets from open government data portals and other administrative or publicly available data can be linked with Internet of Things (IoT) data, e.g., data from sensors deployed in the communal infrastructure or from sensors worn by participants in a city event. The project is also developing several application use cases in the domains of event management, water management, and public transportation and pilots these in partnering cities like Amsterdam, Sapporo and Tokyo. The aim of this paper is to discuss the particular challenges of implementing a smart city platform related to data management and in particular data quality management of various data sources, including IoT data in particular. The use of this type of data is relatively new to governments and irrespective of the usage context, the question of how to validate IoT data is still quite open. In order to be adopted, a city data infrastructure needs to provide information on data quality in form of metadata and, as will be shown, linked data provides several advantages in that respect. At the current stage of the project, we provide generic considerations on the named challenges, based on selected research in the field. Thus, we do not account for potentially differing requirements depending on e.g. the size or smart city maturity (cf. [2]) of a city.

The remainder of this paper is structured as follows: in the following section we summarize the state of adoption of big data in the public sector, with a special focus on the context of smart cities and the usage of the Internet of Things. Section 3 provides an overview of two challenges that need to be addressed for the successful deployment of such a platform: improving data quality and facilitating platform federation. Section 4 then describes linked data as a solution approach to tackle these issues and describes the state of the art in the field, while Sect. 5 goes into more detail how the solution is implemented in the context of the CPaaS.io project. Section 6 finally contains the conclusions and outlines future work.

2 Big Data in the Public Sector

2.1 Big Data Opportunities for the Public Sector and State of Adoption

Big data is about generating value through collecting and analysing information to extract knowledge and insight (cf. [5, 6]). Governments are increasingly aware that big data offers value potentials for the public sector [7]. As scholars point out however, implementation by now tends to be limited [7, 8] or as Desouza and Jacob put it, there

---

[1] The abbreviation stands for "City Platform as a Service – Integrated and Open". See http://www.cpaas.io for further details about this Horizon 2020 project.

is "some tension between the promise of big data and reality" [9]. Accordingly, big data in the public sector has only recently started to raise academic interest (cf. [10]), but is expected to gain more attention within big data research [11]. A first set of studies and reports rather looks at the public sector as a *data producer* for big data applications in other sectors (e.g. [12, 13]). Governments generate and collect large amounts of data through their everyday operations and the public sector is thus one of the most data-intensive sectors. Since public sector sources comply with high quality standards, they are considered an essential resource for the data-driven economy, which is reflected in the many open government data (OGD) initiatives that seek to make this data available for re-use [14]. A second stream of research focuses explicitly on governments as *big data user*. This work includes cross-case studies on existing cases of big data implementation [15, 16], general considerations on the opportunities and challenges of big data adoption by the public sector (e.g. [7, 9, 17, 18]), considerations on the preconditions for using big data [8] and/or specific fields of application, such as policy-making (e.g. [19] also [15, 16]).

Based on available research, *potential benefits* of big data adoption in the public sector can be categorized as follows [7]: A first set of opportunities relates to *improving the knowledge base*. As in other sectors, data analysis is used for generating new insights. Big data analytics can be applied to various domains of public administration (cf. [15–18]) and holds promises to improve all stages of policy-making [16]: Better and faster insights derived from big data analysis (e.g. through machine learning) may help to better react to unintended effects of a policy decision [19]. It may help to earlier detect mistakes, frauds or security threats [20]. Policy-makers can also use big data technologies to conduct policy impact assessments or gain a better understanding of citizen interests and opinions through the analysis of new data sources, e.g. social media, helping them to prioritize policy issues [7, 15].

A second set of opportunities relates to *improvements in effectiveness*. Data analysis may be used to tailor service provisioning towards the needs of different citizen groups, increasing their satisfaction [18]. Better insights can also contribute to solving social problems related to public transportation, healthcare provision or energy production [8]. Provided as open data, the public sector may facilitate the innovation of products and services by third parties.

A third set of opportunities relates to *improvements in efficiency*. Big data can be used to achieve greater internal transparency and to improve data sharing across administrative organizations. Available estimates suggest that the public sector could generate considerable revenues through better exploitation of data [12, 13]. Leveraging new data sources may also positively impact data generation by public administrations, e.g., when producing official statistics [15].

2.2 Smart Cities as Big Data Application Domain in the Public Sector

Depending on the application domain and the type of data generated, big data analysis in the public sector is closely related to the concept of smart cities [18] (for a definition see [21]). Cities are considered as distinct domain, in which the use of ICT in general [22] and big data in particular are expected to generate impact [7, 17, 23–25]. Thought

leaders in the field expect that innovative examples of big data usage are more likely to be found at the city or regional level, since it is easier to get policy makers involved in small-scale initiatives, in new forms of collaboration and data usage [16]. Also, a mapping of smart cities in Europe reveals that there are more smaller than large smart cities, while larger cities have more resources and tend to be more ambitious in scope and more mature regarding implementation [2]. As several authors stress, the Internet of Things is an important data source in the smart city context: "The public sector is increasingly characterized by applications that rely on sensor measurements of physical phenomena such as traffic volumes, environmental pollution, usage levels of waste containers, location of municipal vehicles, or detection of abnormal behaviour" [7]. The analysis of such IoT data sources in combination with other data has the potential to improve urban management and the quality of life of city inhabitants: "Data from different sources need to be integrated and analyzed for smart urban planning, smart transportation, smart sanitation, smart crime prevention, etc." [17]. As Scuotto et al. point out however, "the relationship between IoT and smart cities is still largely unexplored" [26], which requires more research, e.g. on typical technological challenges to be tackled (cf. [23]).

3 Technological Challenges of Implementing a Smart City Platform

While governments are considered as catalysts for boosting a data-driven economy and growth through opening up their data, big data adoption in the public sector is also confronted with a range of constraints and challenges [7]. These are related to *governance* (e.g. agreements for integrating data sources across organizations, data-driven culture), *implementation* (e.g. organizational maturity in terms of IT facilities and data systems, required skills) and *risk management* (privacy, security) (cf. [7, 8, 17–19]).

As Munné points out [7], it is important that the public sector gains "adoption momentum", moving from marketing around big data to real experience, to derive lessons learned on which applications are valuable and how to deploy them: "This requires the development of a standard set of big data solutions for the sector." The CPaaS.io project provides such a solution for the smart city context and supports experimentation and capability building. One of several challenges to be addressed relates to the federation of existing platforms. Another typical challenge relates to data management. A linked data approach is suited to address both the challenge of ensuring system interoperability as well as data interoperability (cf. [27]).

3.1 Federation of Smart City Platforms

To exploit the full potential of a big data strategy, it is not enough that cities just implement a big data platform on their own. Unfortunately, this approach is still common today, leading to data silo solutions lacking interoperability (cf. [17]). Cities though are not standalone entities, they are embedded in a region, in a country, and they often cooperate with other cities – today also on a global scale. Cities thus need to strive for

interoperability of their platforms and the possibility to federate instances: This will enable data analysis across regions from which all participating cities can profit, for example by better understanding traffic patterns or in order to provide better services to an increasingly mobile population. Standards can help to achieve this, but often are not enough, as the adoption especially of data standards on a global scale is slow due to historic, legal and cultural differences.

3.2 Governance and Management of Data

In the age of big data, datasets become increasingly "complex", which requires adequate capabilities for managing the data [9]. With the growing need to integrate data from multiple sources, data quality management becomes both more important but also challenging [28]. In the context of developing city data infrastructures, data management and in particular the management of data quality, i.e., ensuring that data is fit for use and free of defects [29] are important aspects (cf. [28, 30]) and part of an organization's overall data governance [31] (see Fig. 1).
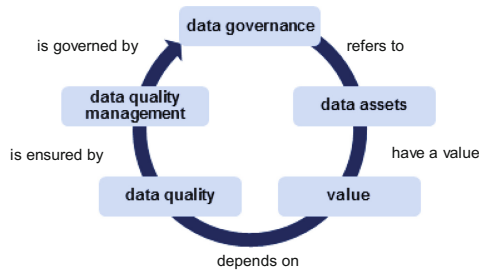

Fig. 1. Data governance and related concepts (adapted from [31])

Smart city platforms are used for making decisions and providing services based on the results of querying various datasets, which entails that applications need to be trusted and accepted and data quality plays a major role in that respect (cf. [32–34]). For a city data infrastructure aimed at integrating IoT data, managing data quality is particularly crucial, as sensors are an inherently unreliable data source. Sensors can become decalibrated, delivering inaccurate data readings, or they can fail or lose connectivity completely. Resolution, sensitivity, timeliness and provenance are other factors affecting the validity of IoT data.

As a requirement, data quality is well understood [35] and there are many methodologies to conduct data quality management (cf. [36]) as well as models and frameworks for assessing the quality of specific types of data, such as linked data [32], IoT data [33], open government data [37] or more generic big data [28]. What constitutes "good data quality" is however depending on the context of its use and thus very much application-dependent. A city data infrastructure aiming to support a multitude of possible applications must provide sufficient metadata about the data quality, while it is left to the application to decide if the data is good enough to be used. This requirement is also grasped by the emerging "Smart Data" paradigm, according to

which successful big data implementation has "a clear meaning (semantics), measurable data quality, and security (including data privacy standards)" [35]. This entails making data more accessible through adding metadata for structuring and integration across separate data silos and for storing information on data quality as well as benefitting from already available open and linked data.

4   Linked Data as Solution Mechanism

In the context of big data, *linked data* is both a specific type of *data source* and an *approach* for facilitating data integration and re-usage through providing clear meaning. This is essential, since only through understanding the context sensitive meaning of data can one assess whether data can be combined to generate value [27]. As Shiri points out: "the formalized, structured and organized nature of linked data and its specific applications, such as linked controlled vocabularies and knowledge organization systems, have the potential to provide a solid semantic foundation for the classification, representation, visualization and the organized presentation of big data" [38].

As the cross-case study on big data adoption in policy-making shows [16], public administrations use a variety of data sources from administrative data, official statistics, surveys, sensors and social media. The data used may be either open or restricted. These siloed data sources are typically accessed over platform-proprietary APIs. To gain new insight about the data it is vital to fuse it from these different sources. This can be done by transforming the data into a more generic form, which is more accessible and provides standardized APIs on top of it. Such a generic API needs to provide a common way to exchange information between these sources and help the API consumer to understand the semantics and the meaning of the information. The W3C *semantic web and linked data technology stack* [24] aims at solving these problems. The *RDF data model* provides well-known schemas and ontologies as lingua franca, HTTP as transport layer, URIs as decentralized identifiers and multilingualism in its core. This makes it the data model of choice for bridging between data silos (cf. [27]).

In the past few years a lot of effort went into publishing best practices. In 2016, W3C released the "*Data on the Web Best Practices*" recommendation [38]: After roughly 10 years of open data movement [39], the document summarizes best practices and recommendations about how to publish open data, especially in the context of what needs to be taken into consideration to ensure that the published data is of maximum value for the public. Most of the recommendations are related to machine readability and discoverability of open data.

In the domain of *schema* and *ontologies*, several search engine giants launched schema.org [39], an initiative to "create and support a common set of schemas for structured data markup on web pages" Meanwhile schema.org seems to use "a simple RDF-like graph data model" and exposes its schema as embedded RDF. Over the past years schema.org had a huge impact; many sites started to include structured information

within their websites and the support of first RDFa[2] and later JSON-LD[3] made people use semantic web technologies without being really aware of it. This increases visibility and perception of the semantic web as a whole.

Developments that are still work in progress revolve around *constraint languages*; examples of that are Shapes Constraint Language (SHACL) and Shape Expressions (ShEx). RDF is a graph data model and by design it is possible to express any relationship between a subject and an object. In real world applications, it is often necessary to define structural constraints and validate RDF instance data against those. This can be done with both of the languages.

In the domain of *IoT related ontologies* there are even more options available and under discussion. Several groups are pushing their own concepts and ontologies, among others: Spatial Data on the Web Working Group's SSN Ontology, IoT + schema.org, Web of Things (WoT) Interest Group and the EU H2020's FIESTA-IoT project. It is too early to tell yet which of the proposed constraint languages and IoT ontologies will see wider adoption in the next years.

5   Implementation in the CPaaS.io Project

The smart city platform as developed by the CPaaS.io project is based on a common reference architecture, but is for pragmatic reasons – mainly in order to have instances up and running quickly in the two main regions of the project – implemented on top of existing frameworks (see Fig. 2).
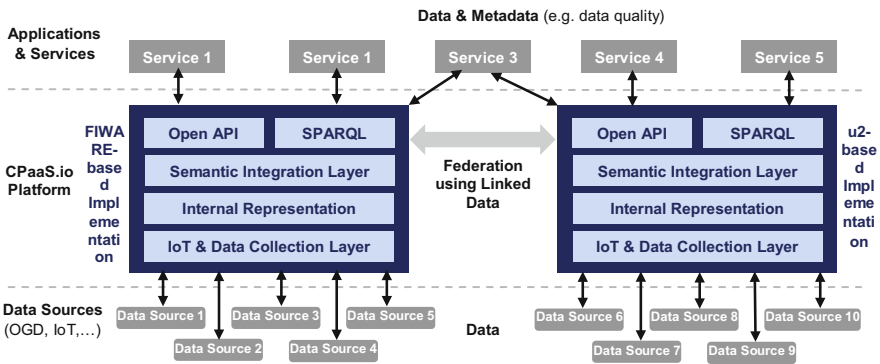

Fig. 2. Simplified CPaaS.io implementation architecture

The implementation in Europe is based on FIWARE [41], and the one in Japan on the u2 architecture [42]. The disadvantage of having two different platform implementations within the project is that data federation across instances becomes more challenging. However, in real life it cannot be expected that all cities will use the same

---

2   RDF in Attributes, see RDFa Primer: http://www.w3.org/TR/xhtml-rdfa-primer.
3   JSON for Linking Data (JSON-LD) is a JSON based RDF serialization. See json-ld.org.

platform implementation anyway, so the two implementations within the project serve as a test regarding the real-world viability of the platform. For example in the domain of public transportation, it is important that innovations developed in one city can easily be transferred to another city.

The CPaaS.io Platform aims at fusing data from different sources, in particular the FIWARE platform in Europe and the u2 platform in Japan. From a data consumer perspective, it should not matter where the data is stored, CPaaS.io will facilitate discovery and access to information in these data storages via generic APIs.

For that reason, CPaaS.io will use linked data and RDF to facilitate integration of data from platforms like FIWARE and u2. Neither of the two platforms is currently supporting RDF and linked data out of the box. CPaaS.io will integrate a semantic layer that enables mapping existing data to RDF. This semantic integration layer can be implemented in different phases and levels.

Initially the semantic layer will simply expose metadata as linked data, using common vocabularies and best practices as described in [38]. This enables users to query information about available data within the FIWARE and u2 platform as linked data. Access to this metadata layer will be done by providing a SPARQL endpoint that can be queried.

In a second step, data residing in FIWARE or u2 is mapped to RDF by extending the respective data model of each platform. In the case of FIWARE, this can be done by using the new NGSIv2 data model that supports JSON-LD representations. By providing appropriate tools and user interfaces, FIWARE users can thus map existing data to RDF representations. The ucode data model of u2 is close to RDF as it stores information in a triple-like data model. The semantic integration layer only needs to map internal ucode IDs to publicly used and dereferencable URIs, preferably as HTTP URIs to allow linked data usage like it has already been done for the Tokyo Metro real time data system.

To be able to query this kind of data, a SPARQL endpoint will proxy requests to the platform. CPaaS.io will provide a virtual-graph feature similar to what RDF graph databases provide to access relational data, using W3C standards like R2RML. Users will thus be able to run SPARQL queries on data residing in FIWARE or u2. In a final step FIWARE and u2 will implement its own SPARQL endpoint.

6 Conclusions and Outlook

While the public sector can be considered one of the most data-intensive sectors, actual use of big data in this sector is still rather limited. With the increased deployment of Internet of Things technologies and the international competition for cities to become smart, however, this is likely going to change. A lot of experimentation is still going on in this area to understand both the technologies as well as the applications that create real public value. To reap the potential benefits, cities will need an open city data infrastructure, where third parties can access the relevant city data, including data coming from the Internet of Things, and provide additional services on top. The platform that the CPaaS.io project is developing could serve as the basis for such an infrastructure if

the two crucial issues that we highlighted in this paper are addressed: the ability to federate platform instances, and data quality. Linked data can serve as a possible mechanism to address both. The semantics behind linked data allow combining differently structured data from technically different platforms. And linked data can be used to annotate data sets with quality parameters so that an application using that data can decide if the data quality is good enough for the intended purpose. It is thus a fruitful approach for reaching the "smart data" paradigm.

Using linked data requires adequate vocabularies both for data integration into CPaaS.io and for re-usage by use case applications as well as for data dimensions and measures, accounting for the different types of data used in the project. Standards for such vocabularies and for validating data are still emerging; at this point in time none of these is well accepted yet. In the further course of the project, we will have to define which of the emerging vocabulary standards are suitable for the project and its use cases, and where we need to define our own. Furthermore, we plan to validate the applicability and the value of linked data, as well as the platform as a whole, in real-world use case implementations in European and Japanese cities. Both, the relationship between IoT and smart cities and the adoption of big data in the public sector in general require more research based on real applications. The CPaaS.io project will contribute to gaining new insights in these emerging research fields.

References

1. European Commission: Building a European Data Economy (2017). http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=41205
2. Manville, C., et al.: Mapping Smart Cities in the EU (2014). http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET(2014)507480_EN.pdf
3. Walser, K., Haller, S.: Smart governance in smart cities. In: Meier, A., Portmann, E. (eds.) Smart City. EH, pp. 19–46. Springer, Wiesbaden (2016). doi:10.1007/978-3-658-15617-6_2
4. Vega-Gorgojo, et al.: Case Study Reports on Positive and Negative Externalities (2015). http://byte-project.eu/wp-content/uploads/2015/06/FINAL_BYTE-D3-2-Case-studies-report-1-1.pdf
5. NIST Big Data Public Working Group: NIST Big Data Interoperability Framework: Volume 1, Definitions (2015). http://dx.doi.org/10.6028/NIST.SP.1500-1
6. Curry, E.: The big data value chain: definitions, concepts, and theoretical approaches. In: Cavanillas, J.M., Curry, E., Wahlster, W. (eds.) New Horizons for a Data-Driven Economy, pp. 29–37. Springer, Cham (2016). doi:10.1007/978-3-319-21569-3_3
7. Munné, R.: Big data in the public sector. In: Cavanillas, J.M., Curry, E., Wahlster, W. (eds.) New Horizons for a Data-Driven Economy, pp. 195–208. Springer, Cham (2016). doi:10.1007/978-3-319-21569-3_11
8. Klievink, B., Romijn, B.-J., Cunningham, S., de Bruijn, H.: Big data in the public sector. Uncertainties and readiness. Inf. Syst. Front. 1–17 (2016)
9. Desouza, K.C., Jacob, B.: Big data in the public sector: lessons for practitioners and scholars. Adm. Soc. 1–22 (2014)

10. Gaardboe, R., Svarre, T., Kanstrup, A.M.: Characteristics of business intelligence and big data in e-government: preliminary findings. In: Tambouris, E., et al. (eds.) Electronic Government and Electronic Participation, pp. 109–115. IOS Press, Amsterdam (2015)

11. Akoka, J., Comyn-Wattiau, I., Laoufi, N.: Research on big data—a systematic mapping study. Comput. Stand. Interfaces (2017). doi:10.1016/j.csi.2017.01.004

12. OECD: Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data". OECD, Paris (2013). doi:10.1787/5k47zw3fcp43-en

13. Manyika, J., et al.: Big Data: The Next Frontier for Innovation, Competition, and Productivity. McKinsey Global Institute, New York City (2011)

14. Carrara, W., et al.: Analytical Report 1: Digital Transformation and Open Data (2015). https://www.europeandataportal.eu/sites/default/files/edp_analytical_report_n1_-_digital_transformation.pdf

15. Barbero, M., et al.: Big Data Analytics for Policy Making (2016). https://joinup.ec.europa.eu/sites/default/files/dg_digit_study_big_data_analytics_for_policy_making.pdf

16. Poel, M., Schroeder, et al.: Data for policy: a study of big data and other innovative data-driven approaches for evidence-informed policymaking. Report about the State-of-the-Art (2015). http://media.wix.com/ugd/c04ef4_cee7fd39eed342beb059526226e6a86e.pdf

17. Aggarwal, A.: Opportunities and challenges of big data in public sector. In: Aggarwal, A. (ed.) Managing Big Data Integration in the Public Sector, pp. 289–301. IGI Global, Hershey (2016)

18. Chen, Y.-C., Hsieh, T.-C.: Big data for digital government: opportunities, challenges, and strategies. Int. J. Public Adm. Digit. Age 1, 1–14 (2014)

19. Höchtl, J., Parycek, P., Schöllhammer, R.: Big data in the policy cycle. Policy decision making in the digital era. J. Organ. Comput. Electron. Commer. 26(1–2), 147–169 (2015)

20. Eckert, K.-P., Henckel, L., Hoepner, P.: Big data—Ungehobene Schätze oder digitaler Alptraum (2014). https://www.oeffentliche-it.de/documents/10181/14412/Big+Data+ungehobene+Schätze+oder+digitaler+Albtraum

21. ITU-T: Shaping Smarter and More Sustainable Cities (2016). http://wftp3.itu.int/pub/epub_shared/TSB/ITUT-Tech-Report-Specs/2016/en/flipviewerxpress.html

22. Castelnovo, W., Misuraca, G., Savoldelli, A.: Smart cities governance. The need for a holistic approach to assessing urban participatory policy making. Soc. Sci. Comput. Rev. 34(6), 724–739 (2016)

23. Hashem, I.A.T., et al.: The role of big data in smart city. Int. J. Inf. Manag. 36(5), 748–758 (2016)

24. Domingue, J., Lasierra, N., Fensel, A., Kasteren, T., Strohbach, M., Thalhammer, A.: Big data analysis. In: Cavanillas, J.M., Curry, E., Wahlster, W. (eds.) New Horizons for a Data-Driven Economy, pp. 63–86. Springer, Cham (2016). doi:10.1007/978-3-319-21569-3_5

25. OECD: Data-Driven Innovation: Big Data for Growth and Well-Being. OECD Publishing, Paris (2015)

26. Scuotto, V., Ferraris, A., Bresciani, S.: Internet of Things: Applications and challenges in smart cities. Bus. Process Manag. J. 22(2), 357–367 (2016)

27. Janssen, M., Estevez, E., Janowski, T.: Interoperability in big, open, and linked data—organizational maturity, capabilities, and data portfolios. Computer 47(10), 44–49 (2014)

28. Merino, J., et al.: A data quality in use model for big data. Future Gener. Comput. Syst. 63, 123–130 (2016)

29. Fürber, C.: Data Quality Management with Semantic Technologies. Springer, Wiesbaden (2016)

30. Bertot, J.C., Choi, H.: Big data and e-Government: issues, policies, and recommendations. In: Proceedings of the 14th Annual International Conference on Digital Government Research (dg.o 2013), pp. 1–10. ACM, New York (2013)
31. Otto, B.: Organizing data governance: findings from the telecommunications industry and consequences for large service providers. Commun. Assoc. Inf. Syst. **29**(1), 3 (2011)
32. Zaveri, A., et al.: Quality assessment for linked data: a survey. Semant. Web J (2015). http://www.semantic-web-journal.net/content/quality-assessment-linked-data-survey
33. Karkouch, A., et al.: Data quality in Internet of Things. A state-of-the-art survey. J. Netw. Comput. Appl. **73**, 57–81 (2016)
34. European Commission: Commission Staff Working Document. Advancing the Internet of Things in Europe (2016). http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15276
35. Becker, T.: Big data usage. In: Cavanillas, J.M., Curry, E., Wahlster, W. (eds.) New Horizons for a Data-Driven Economy, pp. 143–165. Springer, Cham (2016). doi:10.1007/978-3-319-21569-3_8
36. Batini, C., Scannapieco, M.: Methodologies for information quality assessment and improvement. In: Batini, C., Scannapieco, M. (eds.) Data and Information Quality. DSA, pp. 353–402. Springer, Cham (2016). doi:10.1007/978-3-319-24106-7_12
37. Vetrò, A., et al.: Open data quality measurement framework: definition and application to Open Government Data. Gov. Inf. Q. **33**(2), 325–337 (2016)
38. W3C: Data on the Web Best Practices. W3C Recommendation 31 January 2017 (2017). http://www.w3.org/TR/dwbp/
39. Shiri, A.: Linked data meets big data: a knowledge organization systems perspective. Adv. Classif. Res. Online **24**(1), 16–20 (2014)
40. Guha, R.V., Brickley, D., Macbeth, S.: Schema.org: evolution of structured data on the web. Databases **13**(9), 1–28 (2015)
41. Hierro, J.: FIWARE: An Open Standard Platform for Smart Cities (2014). http://www.slideshare.net/JuanjoHierro/fiware-a-standard-platform-for-smart-cities
42. Sakamura, K.: Open IoT Platform & IoT-Engine (2016). https://www.tron.org/wp-content/uploads