A Comparative Analysis of Smart Cities Frameworks based on Data Lifecycle Requirements

Claudia Roessing[1,2] and Markus Helfert[1,2]
[1]Innovation Value Institute, Maynooth University, Maynooth, Ireland
[2]Lero, The Science Foundation Ireland Research Centre for Software, Ireland

Keywords: Data Lifecycle, Data Lifecycle Requirements, Smart City Framework.

Abstract: Citizens migrate from rural areas to urban centres in search of better living conditions. The rural-urban migration combined with rapid population growth lead to overpopulation, which consequently creates challenges to cities in the use and reallocation of their resources. Smart cities have emerged as an opportunity to assist cities to overcome these difficulties with the usage of information and communication technology (ICT) to improve the lifestyle of their citizens. However, maintenance of a smart city is a difficult task. In this multi-stakeholder system, services from different domains are offered to citizens, which collect data from different sources with different formats that need to be in compliance with regulations, privacy, and security requirements. Therefore, a data lifecycle plays a vital role as a data management framework as a means of reducing the complexity of their ecosystems to assist align their objectives and services offered to the citizens. Prior researches have stated a need for improvement in this framework modelling. The aim of this paper is to address this gap and define data lifecycle requirements which will be used to analyse a selection of smart cities architecture frameworks.

1 INTRODUCTION

Over the years, the population has been growing and moving from rural areas to cities in search of improvement on their living standards, thus leading to several challenges for governments to manage cities (Albino et al., 2015). There are several factors that motivate the migration of people from rural areas to cities, for example, the opportunity to find a better livelihood, climate variability, access to basic services and infrastructure (Manzi, 2016; Haoyang et al., 2019). Rural migration combined with growth population cause overpopulation of cities, impact on urban development, sustainability, pollution and cause a reduction in agricultural production (Fernandez-Anez et al., 2018; Manzi, 2016). It also has an impact on health, social infrastructures, and housing sector that cannot keep up with high demand and often resulting in informal growth of urban settlements (Mahbubur Rahman et al., 2019). The concept of a smart city has appeared as an opportunity to improve the quality of life of its citizens using information and communication technology by offering better quality services and at the same time transforming cities into more sustainable ones (Lim et al., 2018; Pérez-Delhoyo et al., 2016; Rabelo et al., 2017). Digital transformation has brought several opportunities for services and infrastructure management, however, these opportunities bring challenges in several aspects (Lnenicka and Komarkova, 2019). The implementation and maintenance of a smart city is a complex task due to its specific characteristics. A smart city is made of heterogeneous technologies and data, several domains which are composed of multi-stakeholders, which in the end needs to achieve goals and objectives having a focus on citizens (Albino et al., 2015; Siddiqa et al., 2016). And in order to provide all services and products to citizens, a city must be in compliance with regulations, security, and privacy requirements (Liu et al., 2017). Therefore, the use of a data lifecycle is necessary to assist to integrate processes, people, and systems in a smart city, as well as the use of enterprise architectures.

A data lifecycle is a framework that contains phases and activities that data has to go through from its creation, processing, archival, and/or disposal in order to prepare data for relevant users meeting specific requirements for quality and security (Arass et al., 2017; Sinaeepourfard et al., 2016).

Roessing, C. and Helfert, M.
A Comparative Analysis of Smart Cities Frameworks based on Data Lifecycle Requirements.
DOI: 10.5220/0010479302120219
In *Proceedings of the 10th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS 2021)*, pages 212-219
ISBN: 978-989-758-512-8

Enterprise architecture is a conceptual blueprint that captures the essence of business, IT, and its evolution. It is used to align IT infrastructure with the business goals of organizations (Lankhorst, 2017). Enterprise architecture has been used to model smart cities, in order to reduce complexity and to tackle challenges faced by this ecosystem which has to integrate different components, moreover to assist in the communication between stakeholders (Guo and Gao, 2020). Several enterprise architecture frameworks have been developed over the years. Some have been developed for specific domains and others in a more generalized way, causing similarities and disparities between them (Urbaczewski and Mrdalj, 2006).

This paper explores the limitations of data lifecycle modelling and defines a set of requirements in order to bridge the gap regarding a lack of formal specification of this framework in the smart city domain. Moreover, this study provides some comparison on a selected number of smart cities frameworks based on data lifecycle requirements defined by the authors.

This paper is structured as follows: Section 2 provides the research approach followed by this study. In section 3, a background of data lifecycle modelling is presented. Section 4 defines the requirements of data lifecycle models. The selection and comparative analysis of smart cities frameworks are conducted in Section 5. Section 6 presents an illustrative use case followed by a discussion and conclusions in sections 7 and 8 respectively.

2 RESEARCH APPROACH

One of the goals of this article is to analyse a selected number of smart city frameworks based on data lifecycle requirements. With this in mind, a literary review was conducted to define necessary requirements to model a data lifecycle.

For this review, the authors adopted a methodology proposed by Webster and Watson (2002). Firstly, data sources were defined from where relevant studies were going to be collected (Springer Link, Google Scholar, IEEE Xplore, Web of Science). Secondly, we defined keywords to be used as search strings in each library database provided. The keywords used were: data life cycle and data lifecycle requirements. Thirdly, a screening phase was conducted, where duplicate articles were excluded, and abstracts of remained articles were revised to remove ones that were not relevant to this study. In the final step, 27 articles were selected out of a total of 97.

In the next section, this study provides a background on data lifecycles modelling.

3 LIMITATIONS OF DATA LIFECYCLE MODELING

Despite advances in the area of data management, representation of life cycles is still being made in a generic way. Even with the changing role of data in organizations, the framework is modelled from a high abstract point of view, and not representing reality, but an ideal situation, showing data as unproblematic (Carlson, 2014; Cox and Tam, 2018; Pouchard, 2015).

Due to the complexity related to data management in a smart city, it is paramount that models show data transformation throughout the process, from its collection, processing, and service delivery to end-user, also showing various stakeholders involved in a process (Ball, 2012). Another drawback in the representation of a process occurs in the acquisition of data since it only allows collection at beginning of a process, thus not allowing a new acquisition in case of any error in data previously collected (Pouchard, 2015). Thus, there are only a few models that make it possible to return to an earlier stage if necessary. To take better advantage of an immense amount of data to which organizations have access, it is essential that they create value from this data, so that they can offer better services and products to end-users.

The purpose of lifecycles is to provide information to interested parties for those who can make decisions, and because of that, it is a relevant tool to use it. And in order to provide information to stakeholders, data lifecycles need to be updated to assist stakeholders to make the best decisions (Plale and Kouper, 2017).

As stated previously, prior studies recognize models' limitations, identifying that they provide an unrealistic point of view when managing data and only a few models recognize this flaw and try to circumvent it (Cox and Tam, 2018). Overall, the studies provide valuable insights into data lifecycle’s limitations (see Table 1) and strengths but also bring attention to gaps and a necessity for change in the data management field.

Table 1: Data lifecycle limitations.

References	Limitations of data lifecycles						
	Lack of necessary phases	Lack of flexibility to adapt	Lack of user feedback	Do not consider data quality	Lack of reality representation	Do not consider data privacy	Need to contribute to politic, social and economic values
El Arass et al. 2017	X	X					
Pouchard, L. 2015					X		
Plale, B., & Kouper, I. 2017					X		
Cox et al. 2018					X		
Ball, A. 2012					X		
Carlson, J. 2014		X			X		
van Veenstra et al. 2015		X			X		
Sinaeepourfard, A. et al. 2015		X		X			
Elmekki, H. et al. 2019			X	X			X
Faundeen, J. et al. 2017		X					
Sinaeepourfard,A. et al. 2016		X					
Paskaleva, K. et al. 2017						X	
Bohli,J. et al. 2015						X	
Liu, X. et al 2017						X	
Alshammari, M. et al. 2018						X	
Attard, J. et al. 2016							X

4 DATA LIFECYCLE REQUIREMENTS

This section defines requirements that a data lifecycle should have, in order to enhance the way data is modelled and meet researchers and practitioner's needs. Modelling requirements were identified during a literature review (see Appendix A), which showed a lack of formal specification for the framework modelling in the smart city domain. Some problems related to this lack of standardization were also identified in the literature (Cox and Tam, 2018).

1. Phase - represents all steps that data needs to go through to achieve a specific outcome.
2. Activities - processes that are conducted in each stage to prepare data for the next stage or to a final objective.
3. Data input - data used in a stage to be transformed.
4. Data output - it represents data that has been transformed from a previous stage and is going to be used in the next one or it is the final output if a life cycle has reached its end.
5. Role - actor responsible to conduct a phase or activity.
6. Pre and post requirement (phase quality) - it is used to know if activities of a phase have been performed with success, in other words, if data has achieved the goal of a phase, therefore it can proceed to the next one, otherwise, it has to be processed again. These requirements are related to the quality of each phase or activity.
7. Relationship between phases – in order to process data for a specific purpose it is necessary to know the order and relationship between phases of a cycle.
8. Variation driver - it is composed of relevant aspects related to data such as regulations, lifespan, category, and sensitivity. Data can be classified into different categories based on their type of sensitivity. In order to process data, it is necessary that phases and activities be in compliance with regulations. It is also necessary to take into account the lifespan of data because it specifies how long data can be used and stored. These aspects influence the choice of lifecycle activities.

5 ANALYSIS OF SMART CITIES FRAMEWORKS

As stated previously, enterprise architecture is a strategic instrument to organizations, and it can be used to guide organizations to go from a current state to a future one (Lankhorst, 2017).

This section provides a high-level analysis of five smart city frameworks which selection criteria were to be composed of at least three layers including information or data layer and their respective descriptions. The frameworks are analysed using concept centric approach proposed by Webster and Watson (2002) based on data lifecycle requirements defined in the previous section. The selected frameworks are: open geospatial consortium, smart city reference architecture meta-model, Nora, ICT architecture, and government enterprise architecture for Big and Open Linked Data analytics. The frameworks are presented and analysed below.

The Open Geospatial Consortium (Open Geospatial Consortium, 2015) developed a smart city spatial information framework. The framework uses viewpoints based on ISO/IEC 10746, information technology – open distributed processing reference model.

The work emphasizes the importance of location in order to organize smart city services. The framework provides a high-level view of components and it is composed of four layers, application, business, data, and sensing layers. It also contains a security system, cloud-hosted resources, and a list of stakeholders. As it is a high-level structure, it does not detail the applications involved, only the application domain. The structure also does not show relationships between entities, as well as there is a lack of goals and objectives. Data entities are divided into domains and are stored in an urban/municipal database. The business layer shows that analytics and models are used for visualization and decision support.

The Smart City Reference Architecture Meta-Model (smartCityRA) developed by Abu-Matar (2016) emerged to supply the need for heterogeneous ecosystem design. It provides a new approach to design heterogeneous ecosystems like smart cities. The framework consists of building blocks that highlight intra and inter views relationships. SmartCityRA follows ISO / IEC / IEEE 42010.2011 to describe terms of models, views, and viewpoints. The reference architecture was developed in a modular way, thus allowing its extension according to domain experts' needs. The meta-model of the framework consists of eight views that are unified by capability view, which represents business requirements provided by a smart city project. The views are capability, participation, place, services, data, application, infrastructure, and business process. The model provides relationships between views. The capability view can represent the goals of a city, however is modelled in higher abstraction. The framework does not provide objectives either.

NORA is the Dutch Government Reference Architecture (SmartCities, 2011). The framework gets requirements from Europe, Dutch Government, companies, and citizens, which are used to build architecture. Company, information, and technical are the architecture domains that describe who, what, and how for each domain. Maintenance/control and security are also described as domains. Company architecture defines an organization, services/products, and processes. Information defines employees/software, message/data, and information exchange. The technical architecture identifies technical components, data storage, and network. The framework is used to create an initial step in the e-government that is to create reusable e-government assets. The model is used as a primary reference architecture used in new ICT projects. It also provides design principles at different levels of organization, process, information, and technology. Further on, the framework focus on domain-specific reference architectures for various aspects of the Netherlands (municipalities, provinces, and water control boards).

ICT Architecture (SmartCities, 2011) provides a simplified architecture metamodel, which is based on TOGAF. The architecture metamodel is presented in two parts. The framework contains seven domains, governance, business, information systems, and technology are the layers of the architecture. The model also defines characteristics for architecture domains related to interoperability, service orientation, and information security. The governance domain at the top defines business goals, strategic drivers, business principles and guidelines, management models, compliance to laws, regulations, and standards. This information is the basis for developing an organization's architecture.
Service orientation is considered as one domain and it flows from top to bottom, which emphasizes enterprise vision for service orientation. Thus allowing reusability and the ability to exchange architecture components without causing a disruption to service. Dependencies between domains are represented too. The metamodel states an alignment between the scope of requirements and implementation of an enterprise architecture. The metamodel is presented in order to provide guidance to e-government stakeholders regarding recommendations to design ICT architectures.

Government Enterprise Architecture for Big and Open Linked Data Analytics (Lnenicka et al., 2017) developed a conceptual framework focusing on big and open linked data analytics requirements in order to guide developers and designers to create government enterprise architectures in smart cities. Before defining a framework, the study presents requirements and their relationships in a smart city ecosystem. The framework consists of four layers 'business, application, data, and technology architectures. Security and privacy + interoperability + evaluation and monitoring occur in all architecture layers. Data flows from bottom to top while service provisionary occurs from top to bottom. Business architecture defines e-government and governance architecture and open government processes. The application identifies smart application services. Data architecture is composed of programming models for analytics that contain batch and stream processing layers. Followed by data API and other interfaces. The last component of this layer is related to data storage, distributed and scalable databases that contain historical and real time data. Last, the

Table 2: Analysis of smart city frameworks.

Framework	Data Lifecycle Requirements						
	Phase	Activity	Input	Output	Role	Pre/Post Requirement	Phases Relationship
The Open Geospatial Consortium	x		x		x		
smartCityRA		x	x		x		
NORA			x		x		
ICT Architecture		x	x	x	x	x	
Government Enterprise Architecture for Big and Open Linked Data Analytics	x	x			x	x	x

technology layer describes smart ICT infrastructure and a smart environment that contains a network of data sources.

6 ILLUSTRATIVE USE CASE

In order to facilitate the proposed approach, this study will use Footfall counter as a use case. This service is offered in some cities and it aims to collect pedestrian counting in certain locations through the use of sensors. It is used for purpose of knowing the traffic pattern of pedestrians and data is mainly used for tourism, retail development, events, just to name a few. This use case is based on real information.

Data lifecycle requirements applied to this use case can be seen below.

Table 3: Use case Data lifecycle requirements.

Phases	Activities	Input	Output
Plan	Specification	Service description	- Processing plan - Access specification - Definition of roles
Collect	Collection	Set of data values	Collected data values (Location, date, IN, OUT)
Storage	Storage Archive Backup	- Storage plan - Data values	Storage data
Use	Use Manipulation	Retrieved data values	Reports (csv, pdf)
Share	- Preparation - Access system Presentation	Retrieved data values	Disclosed data (location, date, IN,OUT)
Delete	Destruction	Processing plan to destroy data	Set of destroyed data

Variation driver:

- regulations: as no personal data is collected, there is no specific regulation that the organization needs to be in compliance with.

- lifespan: the organization has decided to keep data for 10 years.

- category and sensitivity: data collected in this service is classified as public.

Relationship between phases: Phases are conducted in a sequential order (Plan, Collect, Storage, Use, Share, Delete).

Pre and Post requirements: these requirements are verified during each phase and activity to know if inputs and outputs have been met.

Role: programme manager, system users.

7 DISCUSSION

The frameworks analysed in section 5 show existing variations in their modelling. The analysis showed that none of the frameworks meet all requirements defined in this study (see Table 2). Another important point was a lack of connection between layers and entities with the exception of smartCityRA (Abu-Matar, 2016) and Government Enterprise Architecture for Big and Open Linked Data Analytics (Lnenicka et al., 2017) models. Furthermore, the majority of models do not take into account regulations, data category, sensitive data, and data lifespan, which are necessary components to define activities that will be conducted in a data lifecycle. Entity role is defined in all models, however, it does not show a relationship between them and other entities.

Due to the characteristics of a smart city, which is an integration of several components and having data as its main resource, it is relevant to show how entities of a model connect and data flows. The alignment of objectives with policy, regulations, business, and technical approaches are necessary, but these aspects are not reflected in the analysed frameworks (National Institute of Standards and Technology, 2018).

It is also important to emphasize the importance of classifying data so that it can be processed properly and this was another aspect absent from the analysed frameworks.

Overall, all these factors may result in a lack of alignment between citizen's needs and smart city implementation.

8 CONCLUSIONS

Smart cities have emerged as a solution to various challenges found in today's cities, it is a solution-focused on its citizens and mainly to improve their lives and there are many challenges to implement it. Due to its unique characteristics as multi-stakeholder, citizen-centric, data-centric, each smart city has its own implementation and particularity, leading to variations in its enterprise architecture models. The implementation of a smart city requires an alignment between services, policies, and security requirements, therefore it is necessary that frameworks reflect this requirement.

Data is considerably important in a smart city, however over the years, it has not been modelled adequately in order to provide essential information regarding some concerns, for instance, how data is being collected, processed, reused, and stored. Data is used to generate information and knowledge which are used in the decision making and to offer better services to citizens, therefore improvements are needed in its modelling.

The use case provided showed how the application of data lifecycle requirements can assist decision makers to have a holistic view of data processing to offer services to citizens. Sensitive data were not processed in the example, however, usage of data lifecycle requirements proved beneficial to have a better view of the process, especially when sensitive data are processed and also shared with third parties. Therefore, data lifecycle requirements can assist organizations to align services, regulations, and security requirements and moreover to assist in process improvements.

This paper analysed selected frameworks based on data lifecycle requirements, the investigation has shown limitations present in the modelling of smart cities using enterprise architectures. A natural progression of this work is to analyse how to integrate data lifecycle requirements identified in this work and their connections in a smart city framework and to conduct case studies to validate the findings.

ACKNOWLEDGEMENTS

This work was supported by the Science Foundation Ireland grant “13/RC/2094” and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero, the Science Foundation Ireland Research Centre for Software (www.lero.ie).

REFERENCES

Abu-Matar, M. (2016). Towards a software defined reference architecture for smart city ecosystems. *IEEE 2nd International Smart Cities Conference: Improving the Citizens Quality of Life, ISC2 2016 - Proceedings*. https://doi.org/10.1109/ISC2.2016.7580807

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology*, *22*(1), 3–21. https://doi.org/10.1080/10630732.2014.942092

Arass, M. El, Tikito, I., & Souissi, N. (2017). *Data lifecycles analysis : towards intelligent cycle*. https://doi.org/10.1109/ISACV.2017.8054938

Ball, A. (2012). *Review of Data Management Lifecycle Models*. Retrieved from University of Bath website: http ://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf

Bohli,J., Skarmeta. A., Moreno, V., García,D. and Langendörfer,P. (2015) SMARTIE project: Secure IoT data management for smart cities, *International Conference on Recent Advances in Internet of Things (RIoT)*, Singapore, 2015, pp. 1-6. doi: 10.1109/ RIOT.2015.7104906

Carlson, J. (2014). The Use of Life Cycle Models in Developing and Supporting Data Services. In *Research Data Management: Practical Strategies for Information Professionals* (pp. 63–86). Purdue University Press.

Cox, A. M., & Tam, W. W. T. (2018). A critical analysis of lifecycle models of the research process and research data management. *Aslib Journal of Information Management*, *70*(2), 142–157. https://doi.org/10.11 08/AJIM-11-2017-0251

Fernandez-Anez, V., Fernández-Güell, J. M., & Giffinger, R. (2018). Smart City implementation and discourses: An integrated conceptual model. The case of Vienna. *Cities*, *78*(December 2017), 4–16. https://doi.org/10.1 016/j.cities.2017.12.004

Guo, H., & Gao, S. (2020). Enterprise Architectures in E-Governments Studies: Why, What and How? In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-45002-1_1

Lankhorst, M. M. (2017). Enterprise architecture at work: Modelling, communication and analysis, fourth edition. In *Enterprise Engineering Series*.

Lim, C., Kim, K. J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities*, *82*(April), 86–99. https://doi. org/10.1016/j.cities.2018.04.011

Liu, X., Heller, A., & Nielsen, P. S. (2017). CITIESData: a smart city data management framework. *Knowledge and Information Systems*, *53*(3), 699–722. https:// doi.org/10.1007/s10115-017-1051-3

Lnenicka, M., & Komarkova, J. (2019). Developing a government enterprise architecture framework to support the requirements of big and open linked data with the use of cloud computing. *International Journal of Information Management*, *46*(December 2018), 124–141. https://doi.org/10.1016/j.ijinfomgt.2018.12.003

Lnenicka, M., Machova, R., Komarkova, J., & Pasler, M. (2017). Government enterprise architecture for big and open linked data analytics in a smart city ecosystem. *Smart Innovation, Systems and Technologies*, *75*(May), 475–485.https://doi.org/10.1007/978-3-319-59451-4_47

Mahbubur Rahman, M., Hassan, M. S., Md. Bahauddin, K., Khondoker Ratul, A., & Hossain Bhuiyan, M. A. (2019). Exploring the impact of rural–urban migration on urban land use and land cover: a case of Dhaka city, Bangladesh. *Palgrave Communications*, *7*(2), 222–239. https://doi.org/10.1057/s41599-019-0302-1

Manzi, L. (International O. for M. (2016). Migration from rural areas to cities: challenges and opportunities | Regional Office for Central America, North America and the Caribbean. Retrieved March 8, 2021, from https://rosanjose.iom.int/SITE/en/blog/migration-rural-areas-cities-challenges-and-opportunities

National Institute of Standards and Technology. (2018). Framework for improving critical infrastructure cybersecurity. *Proceedings of the Annual ISA Analysis Division Symposium*, *535*, 9–25.

Open Geospatial Consortium. (2015). Open Geospatial Consortium OGC Smart Cities Spatial Information Framework. *OGC Doc. No. 14-115*, 1–32. Retrieved from https://portal.opengeospatial.org/files/61188

Paskaleva, K. & Evans, J. & Martin, C. & Linjordet, T. & Yang, D & Karvonen, A. (2017). Data Governance in the Sustainable Smart City. Informatics. 4. 41. 10.3390/informatics4040041.

Pérez-Delhoyo, R., García-Mayor, C., Mora-Mora, H., Gilart-Iglesias, V., & Andújar-Montoya, M. D. (2016). Making smart and accessible cities: An urban model based on the design of intelligent environments. *SMARTGREENS 2016 - Proceedings of the 5th International Conf. on Smart Cities and Green ICT Systems*, (Smartgreens), 63–70. https://doi.org/10.5220/0005798100630070

Plale, B., & Kouper, I. (2017). The Centrality of Data: Data Lifecycle and Data Pipelines. In *Data Analytics for Intelligent Transportation Systems* (pp. 91–111). https://doi.org/10.1016/B978-0-12-809715-1.00004-3

Pouchard, L. (2015). Revisiting the Data Lifecycle with Big Data Curation. *International Journal of Digital Curation*, *10*(2), 176–192. https://doi.org/10.2218/ijdc.v10i2.342

Rabelo, A. C. S., Oliveira, I. L., & Lisboa-Filho, J. (2017). An architectural model for intelligent cities using collaborative spatial data infrastructures. *SMARTGREENS 2017 - Proceedings of the 6th International Conf. on Smart Cities and Green ICT Systems*, (Smartgreens), 242–249. https://doi.org/10.5220/0006306102420249

Siddiqa, A., Hashem, I. A. T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., & Nasaruddin, F. (2016). A survey of big data management: Taxonomy and state-of-the-art. *Journal of Network and Computer Applications*, *71*, 151–166. https://doi.org/10.1016/j.jnca.2016.04.008

Sinaeepourfard, A., Masip-Bruin, X., Garcia, J. and Marín-Tordera, E., (2015). A Survey on Data Lifecycle Models: Discussions toward the 6Vs Challenges. *Technical Report (UPC-DAC-RR-2015–18)*.

Sinaeepourfard, A., Garcia, J., Masip-Bruin, X., & Marín-Tordera, E. (2016). A comprehensive scenario agnostic Data LifeCycle model for an efficient data complexity management. *Proceedings of the 2016 IEEE 12th International Conf. on E-Science, e-Science 2016*, 276–281. https://doi.org/10.1109/eScience.2016.7870909

SmartCities. (2011). Creating Municipal ICT Architectures. *A Reference Guide from Smart Cities*. Retrieved from www.smartcities.info/ict-architecture

Urbaczewski, L., & Mrdalj, S. (2006). A Comparison of Enterprise Architecture Frameworks. *Issues in Information Systems*, *7*(2), 18–23.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly*, *26*(2).

APPENDIX A

Ali, I. M. et al. (2019) Measuring the Performance of Big Data Analytics Process, 97(14), pp. 3796–3808.

Alshammari, M & Simpson, A. (2018). Personal Data Management: An Abstract Personal Data Lifecycle Model. 10.1007/978-3-319-74030-0_55.

Attard J., Orlandi F., Auer S. (2016) Data Driven Governments: Creating Value Through Open Government Data. In: Hameurlain A. et al. (eds) Transactions on Large-Scale Data- and Knowledge-Centered Systems XXVII. Lecture Notes in Computer Science, vol 9860. Springer, Berlin, Heidelberg

Ball, A., (2012) Review of Data Management Lifecycle Models. University of Bath, Bath, UK.

Barnaghi, P. & Bermúdez-Edo, M. & Tönjes, R. (2015). Challenges for Quality of Data in Smart Cities. Journal of Data and Information Quality. 6. 1-4. 10.1145/2747881.

Carlson, J., (2014) The Use of Life Cycle Models in Developing and Supporting Data Services. In J. M. Ray (Ed.), *Research Data Management. Practical Strategies for Information Professionals.* West Lafayette: Purdue University Press., 63-86.

Coleman, S.L, (2013). Measuring Data Quality for Ongoing Improvement. A datta quality assessment framework. In MK Series on Business Intelligence, Morgan Kaufmann.

Cox, A. M., & Tam, W. W. T. , (2018) A critical analysis of lifecycle models of the research process and research data management. *Aslib Journal of Information Management*, *70*(2), 142–157. https://doi.org/10.1108/AJIM-11-2017-0251.

DAMA International, (2017). The DAMA Guide to the Data Management Body of Knowledge – DAMA-DMBOK. Technics Publications.

Elmekki, H., Chiadmi, D. and Lamharhar, H., (2019) Open Government Data: Towards a comparison of Data Lifecycle models. In Proceedings of the ArabWIC 6th Annual International Conference Research Track (ArabWIC 2019). ACM, New York, NY, USA, Article 15, 6 pages. DOI: https://doi.org/10.1145/3333165.3333180.

El Arass, M. Tikito, I. and Souissi, N. (2017). Data lifecycles analysis: Towards intelligent cycle, *Intelligent Systems and Computer Vision (ISCV)*, Fez, pp. 1-8.

El Arass, M. & Tikito, I. & Souissi, N., (2018) An Audit Framework for Data Lifecycles in a Big Data context. 10.1109/MoWNet.2018.8428883.

EMC Education Services, (2015). *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Wiley.

Erl, T., Khattak, W. and Buhler, P. (2016). Big Data Fundamentals: Concepts, Drivers & Techniques. 1st Edition. Prentice Hall.

Faundeen, J. & Hutchison, V. (2017). The Evolution, Approval and Implementation of the U.S. Geological Survey Science Data Lifecycle Model. Journal of eScience Librarianship. 6. e1117. 10.7191/jeslib. 2017.1117.

Gharaibeh, A. *et al.* (2017) 'Smart Cities: A Survey on Data Management, Security, and Enabling Technologies', *IEEE Communications Surveys and Tutorials*, 19(4), pp. 2456–2501. doi: 10.1109/COMST.2017.2736886.

Higgins, S. (2008). The DCC curation lifecycle model. International Journal of Digital Curation. 3. 453. 10.1145/1378889.1378998.

Inter-university Consortium for Political and Social Research (ICPSR). (2012). Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle (5th ed.). Ann Arbor, MI.

Levitin, A.V. and Redman, T.C., (1998). Data as a resource: properties, implications, and prescriptions. *MIT Sloan Management Review*, *40*(1), p.89.

Lim, C., Kim, K. J. and Maglio, P. P. (2018) 'Smart cities with big data: Reference models, challenges, and considerations', *Cities*. Elsevier, 82(April), pp. 86–99. doi: 10.1016/j.cities.2018.04.011.

Möller, K. (2013). Lifecycle models of data-centric systems and domains: The abstract data lifecycle model. Semantic Web. 4. 67-88. 10.3233/SW-2012-0060.

Pouchard, L. (2015). Revisiting the Data Lifecycle with Big Data Curation. International Journal of Digital Curation. 10. 10.2218/ijdc.v10i2.342.

Plale, B., & Kouper, I., (2017) The Centrality of Data: Data Lifecycle and Data Pipelines. In M. A. Chowdhury, A. Apon, & K. Dey (Eds.), Data analytics for intelligent transportation systems (pp. 91–111). Cambridge, MA: Elsevier Inc. Retrieved from https://doi.org /10.1016/B978-0-12-809715-1.00004-3.

Sutherland, M. & Cook, M. (2017). Data-Driven Smart Cities: A Closer Look at Organizational, Technical and Data Complexities. 471-476. 10.1145/308 5228.3085239.

Urbinati, A. & Bogers, M. & Chiesa, V. & Frattini, F. (2018). Creating and Capturing Value from Big Data: A Multiple-Case Study Analysis of Provider Companies. Technovation. forthcoming. 10.1016/j.te chnovation.2018.07.004.

van Veenstra, A. F. & van den Broek, T. (2015). A Community-driven Open Data Lifecycle Model Based on Literature and Practice. 183-198. 10.1007/978-3-319-08081-9_11.