

Introducing Data Mesh paradigm for Smart City platforms design

Marcin Krystek¹
Poznan Supercomputing
and Networking Center, Poznań, Poland
mkrystek@man.poznan.pl

Mikołaj Morzy
Poznań University of Technology,
Poznań, Poland
mikolaj.morzy@put.poznan.pl

Cezary Mazurek
Poznan Supercomputing
and Networking Center, Poznań, Poland
mazurek@man.poznan.pl

Juliusz Pukacki
Poznan Supercomputing
and Networking Center, Poznań, Poland
pukacki@man.poznan.pl

Abstract

The concept of a smart city imposes a unique set of requirements for the underlying ICT technologies for a successful implementation of services and applications for citizens. At the core of these requirements lies the complex data platform architecture which must be carefully designed. The selection of a particular data platform architecture incurs significant technical debt to be serviced in the future, as well as the integration challenges involving hundreds of stakeholders. Since services developed within the smart city ecosystem have significant impact on human well-being and quality of life, the process of designing data platform must be robust.

Data mesh paradigm is a new approach for building complex information systems. It is particularly suited as the blueprint for designing data platforms for smart cities. In this paper we present the overview of the data mesh concept. Building upon 25 years of experience of developing applications and providing data infrastructure for the city of Poznan, we identify key challenges when using the data mesh approach to build data platform tailored to a smart city. We provide guidelines for successful introduction of the data mesh at sociological, technological, and infrastructure level. We also point to the usability of the data mesh paradigm in the context of digital twins, a promising vision of future services for smart cities.

Keywords: smart city, data mesh, data platform, digital twin.

1. Introduction

While many definitions of smart city coexist together, all of them have single common denominator - using digital data for making city a better place and citizens' life easier. Since data is the fuel for optimisation processes and business scenarios, appropriate data management and governance, as well as its effective accusation and delivery, are key factors determining the success of the smart city. Over the recent years a number of different approaches to build smart city solutions were implemented, which enabled to gather experience and learn lessons about main challenges. Facing a rapid technology development leading to fundamental changes on how data can be generated, transferred and processed, both city government and its technological partners must consider new methods for creating next generation data platforms, able to address challenges of the future. Data mesh concept with its major shift of perspective from technology to data itself might be an interesting alternative for modern smart city platform design.

Smart city development is strictly coupled with data gathered inside city space and utilised for process optimisation, governance, or building business revenue. All smart applications, analytical models, and decision support systems are built on top of data. Therefore, data itself are the main point of interest for all stakeholders. The ability to effectively collect and distribute data is the key feature of the modern, digitalised smart city. Rapid technology development results in creating large volumes of digital data that

¹ The research was conducted as part of the Polish Ministry of Education and Science's Industrial Doctorate Program implemented from 2020 to 2024 (Contract No. DWD/4/24/2020)

can be consumed for the benefit for citizens. The evolution of the smart city data platforms must reflect technology changes and be prepared for supporting new user scenarios. Smart city data systems developed for over past 20 years suffer from technological debt, which makes them difficult to quickly respond to upcoming changes. However, choosing the right, adequate and most promising technology from the stack of technologies available on the market is also not a straightforward task.

In this article, a data mesh approach for building data platforms is introduced in the context of the smart city. Since data mesh concept rejects the main role of technology and puts data in the primary role, it is an excellent approach which can be followed to upgrade existing platforms in the evolutionary manner. It can be also used as a guidance for brand new designs, helping to avoid commonly made mistakes which usually leads data platforms to fail in the long-term perspective.

This article is organised as follows: Section 2 focuses on the definition of smart city in the context of data. Section 3 provides the description of the former approach of developing smart city solution based on the city of Poznan example. The experience and lessons learned during this development are shortly summarised in Section 4. Section 5 highlights the key aspects of the data mesh paradigm. Discussion about how to apply this concept to smart city environment can be found in Section 6. Section 7 stresses future application possibilities of well-designed smart city data platform. The paper concludes in Section 8 with a brief summary.

2. Smart city data perspective

Smart city concept can be understood differently depending on the organisation which provides its definition, broader context in which it is introduced, or terms used to describe it. At least 24 different definitions of the concept of a “smart city” can be found in the literature [1], which exemplifies the fuzziness of the concept. For the purpose of this article, the definition provided by the European Commission will be adopted - "A smart city is a place where traditional networks and services are made more efficient with the use of digital solutions for the benefit of its inhabitants and business" [2]. It is still a very broad definition, however, it underlines the need to use digital technologies for improving existing processes with the aim of creating a better, more hospitable environment for living. This kind of improvement

requires at first gathering a detail source data which will characterise investigated phenomena from different perspectives, and then conduct an analysis which will provide new knowledge, based on which appropriate actions can be taken. This process might be repeated until the optimisation goals are reached. It is clear that without source data no optimisation process can take place. Continuing along this line of reasoning, the ability to gather, deliver, process and analyse data is a crucial factor creating a smart city. From the data perspective a smart city is a city which takes advantage of various kinds of data describing and characterising ongoing processes and transforming them into knowledge useful for future city governance or creating new business opportunities. Supporting this process is a challenging task, requiring continuous improvements, overcoming emerging difficulties and following latest technology achievements.

3. Former approach to develop smart city

In 1997 City Council of Poznan and Poznan Supercomputing and Networking Center as a technological partner, established the cooperation aiming at introducing city assets into the digital world. As a result, the Municipal Multimedia Guide (MIM) was built. Since the beginning MIM was designed as a communication platform between the city, its inhabitants and visitors [3]. In 2019 the platform was serving nearly 6 million users per year, collecting data submitted by 2500 editors, providing REST APIs for more than 20 data collections [4]. Based on the APIs, the number of different services for residents, tourists and officials were created including: digital plan, platform for public consultations, civil budget services, air quality monitoring, reporting interventions, smart zoo application, just to mention a few. The proposed architecture of the platform successfully supported number of different smart applications which are utilising data published on the platform. Smart city concept is understood in this context as innovative services operating within and across key areas of city life, having dynamic impact on decision-making processes [5].

Platforms and technologies used for smart cities should be extended to support integrated and cross-functional analysis of data from many sources. This includes the development of predictive models that capture the behaviour of individual urban systems as well as the interdependencies with related systems, and further enable automated system controls and

interactions. Ultimately, they address the capabilities of decision-makers at all levels - from individual residents to government officials, and from small businesses to large-scale service providers - to use these integrated data sources and models to inform their own interactions and decisions [6]. The experience gained during deployment of existing smart applications gives some overview about complexity of introducing new scenarios in the current model. On the other hand, sophistication of future scenarios is continuously rising due to the involvement of many new data sets and the willingness to use artificial intelligence solutions for every day applications. Technological advancement and maturity reached by many solutions creates new circumstances in which there is a place for designing next generation data platform, that will help to cope with many upcoming challenges.

4. Challenges of building data platforms for smart cities

Over the past 25 years Poznan MIM platform went through the transformation process from simple informative platform towards data platform aggregating and delivering data for smart applications. During this evolution, a number of different challenges were encountered. Below some of these challenges are summarised.

4.1. Technology

A rapid development of digital technologies creates new opportunities for finding better, more effective ways to process data. On the other hand, the same technological progress allows to generate massive volumes of new data due to progressing digitalisation. There is a general tendency to build and deploy dedicated IT systems with the aim to organise and optimise processes that are designed for public transport coordination system, intelligent transport system, CCTV system, air condition monitoring system, just to name a few. The increasing number of IT systems, as well as growing density of potential measurement points, creates significant volume of data which needs to be stored and processed. From the process business perspective, the raw data can be effectively utilised by the dedicated systems in the real time. However, if sharing and new feature generation is considered, it has to be stored and prepared for distribution with proper governance rules. The desire of having direct insight into ongoing processes,

knowing their status and variables related to these processes at any moment of time, cause the IT real time systems to generate a significant velocity of data, which needs to be also handled properly. The data itself is not only numerical, as the measurements values, but it also comes as plain text, images, video streams, event streams, it can be well structured, semi-structured, unstructured, etc. This brings another challenge which is data variety. Modern data platforms for the smart city should address these challenges by providing means for effective processing and storing data. This is crucial to support different scenarios development in the future.

Creating a dedicated IT systems or deploying vendor specific solutions brings a risk of building data silos. In theory, data is available in the digital form, however it is not directly accessible from the outside of system. This is especially the problem of legacy systems, which have a very little or no support, are no longer developed, and were built without any requirements for sharing data. Even modern vendor specific solutions not always provide means for future integration and data sharing. A lot of effort must be dedicated to extract and integrate these kinds of data sources. It can be avoided, if data sharing and future integration abilities are specified as strict requirements at the system design phase.

Technological development abounds in new kinds of devices and processing concepts. The emergence of IoT devices brought a revolution in perception of what and how can be measured and controlled. This technology opened new horizons for utilising digital technologies in everyday life. However, to exploit data generated by IoT devices, gain true knowledge and holistic view of the city system, the infrastructure, tools and processing workflows need to be properly designed and deployed. From the smart city perspective, it is not about having IoT devices in the city space, it is about getting knowledge and use it for city governance. Another aspect of the technological development is related to new models of data processing. Since cloud computing is available for a longer time, the appearance of IoT devices accelerates emergence of new architectures like edge computing or fog computing. While designing a next generation data platform, existence of these architectures must be also considered.

4.2. Usability

Since technology development enables widespread digitalisation, there must be also a technical means for storing, processing and delivering

this data. Depending on scale, two most common approaches in this area are data warehouse and data lake. Unfortunately, both of them have significant disadvantages. Data warehouse requires a very strict data schema, which means that data transformation process must be implemented before inserting new data into the warehouse. This reduces overall throughput. It is also difficult to represent all variety of data with a single schema, as well as serve future scenarios for which schema might not be optimal. Data lakes can store raw data and, in this way, solve warehouse issues with supporting volume, velocity and variety of data. Processes of curation, transformation and schema application are applied in data consumption phase. It means that this effort will be repeated every time data is accessed. Lack of data quality control and unclear data ownership are the main disadvantages of data lake architecture [7].

In the face of downsides of the existing architectures, it is worth reconsidering user scenarios in the context of data utilisation. Defining these kinds of scenarios for the data platform in advance seems to be nearly impossible. User scenarios are strongly related to data sets available on the platform and these are not known in advance. By adding a single new data set, a number of different user scenarios can appear. The conclusion is that data platform should not be designed to support a particular scenario. It should rather be able to serve any future user requirement and be able to evolve towards user expectations.

Users, and more broadly, stakeholders, are parties whose interest must be represented. At the time of data platform design, it is possible to identify those stakeholders who will be directly involved in the processes around the platform: data owners, analytical teams, data consumers, infrastructure maintenance teams and possibly others. Having in mind the ongoing technology development on the one hand, and evolutionary aspect of the platform on the other, the future stakeholders might differ from the ones identified so far. Automation processes may exclude the need for manual infrastructure maintenance, human consumers might be replaced by AI. Therefore again, it is more important to follow or define standards and provide rules for utilising platform and accessing data rather than focus on supporting particular stakeholders. They are important in case of data governance, but the platform must be flexible enough to introduce new stakeholder at any moment of its existence.

4.3. Sustainability

Taking into account the above challenges arising from technological requirements and platform usability perspectives, and in order to build a long-lasting data platform, the platform must be designed in a flexible way. Otherwise, it will soon be outdated and all efforts invested in the platform development will be lost. Sustainability regarding smart city data platform means:

- ability to consume new types of data: there should be no restrictions or limitations regarding type of data published on the platform,
- ability to deliver new products: whatever result is obtained through applying data processing and analytical processes, it should be possible to materialise it on the platform,
- ability to serve new scenarios: the platform should not put any limitations regarding user scenarios implemented on the platform, this applies to both using delivered data, as well as implementing analytical processes,
- ability to support new stakeholders: the platform should provide technical solutions able to serve and being consumed by any future stakeholder with the ability to reflect interest of new stakeholders in the governance rules.

4.4. Governance

There might be different approaches to introduce data governance rules into data platform. The centralised approach collects power and responsibility for setting rules which are applicable across the whole platform in a single body. This might be a case for small, centralised platforms like warehouses, where there is a single data owner and single data set. For more distributed approaches, where there is a need to support multiple data sets, number of data owners and very different user scenarios, it is nearly impossible to deliver common rules for all of them because of excluding requirements and different expectations. This lack of governance rules is pointed out as one of the data lake drawbacks, which lead to poor quality of data. [7] So neither strict, centralised governance, nor the lack of governance at all are appropriate solutions. Governance is about defining rules, making decisions and being in charge to execute them. However, rules should not be enforced. They should reflect needs and interests of all platform stakeholders in order to be accepted and respected.

5. Basic tenets of Data Mesh

Designing a data platform to support smart city development is a challenging task, as shown in the previous section. It involves number of different aspects that need to be properly addressed in order to deliver a functional platform, prove its usability and make it successful in a long term. Fostering this long-term perspective seems to be crucial as urban data platform can be considered as a public investment, and approaches used so far for implementing similar cases in the corporate environments frequently fail for many different reasons [8]. Data mesh concept for building next generation data platforms propose an alternative approach. According to the data mesh paradigm, technology is no longer in the centre of concern. By focusing on the data, all other aspects become a means to deliver a valuable product, as the data itself should be considered [9]. This paradigm shift, as well as new concepts introduced by data mesh (like data domain, distributed data governance), are trying to address the most important weaknesses of the first and second generation data platforms (warehouses and data lakes respectively), the weaknesses that lead them to final failure, understood as an inability to deliver expected values [8].

Paradigm shift is not only about changing the focus from technology to data. It introduces a spectrum of changes on different levels starting from a sociological perspective involving human interactions, through technical aspects of how specific tools are used, how infrastructure is delivering computing power and storage, on the data governance ending.

Sociological level

Delivering data product in some sense is similar to delivering any other product on the market. Additional effort is required to prepare and advertise it to potential users. Data product advertisement should include description of the following data traits:

- discoverable: it should be described in the way, that allows web crawlers and indexing engines to easily find and reference to particular resource,
- understandable: the description must be clear and explicit in describing what the data are about,
- addressable: is should be possible to create and manage direct reference (address) to data set, proper governance of this address space is crucial in order to design and maintain reusable workflows, powered by the data,

- interoperable: whenever it is possible, any procedure regarding data - its acquisition, processing, describing should follow well defined standards,

- secure: authorisation and authentication mechanisms need to be present in order to control who has access to data, and which data is available for particular user,

- trustworthy: the precise procedures must be defined in order to classify data source and data itself, the verification process must ensure that system will not be populated with low quality data,

- valuable: estimated parameter, describing business potential of the data set, understood as a possible profit which can be gained due to exploiting particular data set; in the smart city environment, it can also represent data usability in the city management processes.

The approach of treating data as a product should change the perception of data delivering process. There is no more place for best effort approach, realised in the meantime just in order to fill the checklist. Data delivery should be a well-defined process, which produces a valuable product at the end. Understanding this major change is crucial for building well qualified teams of experts, who will be truly devoted to reaching expected goals. It is also important to mention that product thinking can significantly change data owner's perspective. By following data governance rules, a number of conscious decisions must be made in order to design and deliver data products. In this context, data sharing should be considered as an opportunity to deliver additional value to the community and/or business rather than a threat to the organisation security or revealing business secrets. Due to organisational complexity of the city management council, its division into different departments and distributed responsibility, as well as the existence of many city owned companies, it will be nearly impossible to force every data owner to follow exactly the same procedures to deliver data. However, by defining general rules of data governance and providing common understanding of the data sharing purpose, as well as responsibility for its quality, all teams are equipped with tools to follow data product delivery process in their own manner.

Technology level

Providing technologies which will support teams in creating and delivering new products to data

platform, as well as other platform users, might be an important concern. Fortunately, existing technologies already provide capabilities to set-up required environment. The key to success is taking advantage of cloud platforms and resource virtualisation technologies. Considering a user toolkit, this approach is able to deliver any requested software in the form of ready to use services. The existing Software as a Service solutions provided by the global technological companies like Google, Amazon, Microsoft already proved their usability and ability to realize successful business model. Providing custom service in the similar model, tailored to the exact user needs, will significantly reduce time and cost required for preparation and maintenance of the software stack. As software management tasks will be dedicated to DevOps teams, data management teams can focus only on their analytical responsibilities. This is crucial to avoid bottlenecks in the cooperation between different teams and allows to reuse the same resources for different scenarios. At the same time, providing software resources in the cloud enables flexible management. The amount of computing power and storage resources can be easily adjusted to meet exact requirements defined by particular scenario. It is foreseen that the BigData-like platforms delivering dedicated technologies for distributed storage (HDFS), processing (SPARK), task management, user control, together with object store technologies providing limitless scalability will become a foundation for the next generation data platforms.

Infrastructure level

The infrastructure aspect was already mentioned in the previous paragraph, as an environment hosting services. However, the infrastructure itself, as well as teams maintaining these resources, have also a special place in data mesh platform concept. It is very common for a proof-of-concept scenario carried out with direct and hands-on support from the infrastructure team to mature over time and become a production solution. This fact accelerates emerging many new scenarios which might be more complicated and time consuming. In very short time, the organization of such projects in which the infrastructure team has been somehow involved in the development process, will become riskier due to the higher priority given to infrastructure maintenance as opposed to development work. Having this in mind, the infrastructure for the data mesh must be provided from the early beginning in the self-service manner. This means that data management

teams must be equipped with a set of tools, which will allow them to individually execute e.g. CI/CD pipelines, processing templates for repeatable tasks or take advantage of ready to used general purpose services like e.g. common storage layer. Data management team must be able to perform on their own as much technical and infrastructure tasks as possible, firstly to avoid possible bottlenecks, secondly because as implemented scenarios become more and more complicated, any decision will require detailed domain knowledge that is available among data team.

6. Introducing Data Mesh into Smart City

In the previous section some very basic aspects of data mesh were shortly introduced. Authors of this concept provide a comprehensive overview of their motivation to pursuit new design for a data platform, considering present technologies that can be utilised and downsides of the existing architectures [9]. Data mesh concept was introduced based on their experience gained while putting into practice many user scenarios related to any form of data platforms. It is a concept built on very practical approach and has a strong corporate background. This might be an advantage since it is based on real user scenarios. However, perceiving city management as a form of corporate management seems to be unjustified and inappropriate. The differences become clear when organisational structure of the modern city inhabited by more than 1 million people is considered. Also, objectives of each city department are focussed on providing services to citizens rather than gaining commercial profit. Urban environment might be managed differently on different levels, there are different competences and level of understanding between officials, it involves subsidiary companies operating on behalf on city council but having their own status and rules. Considering the issues mentioned above, it would be difficult to apply data mesh concept directly within the city space. Data mesh needs to be rethought and correctly adjusted to fit the specificity of city environment. However, in order to make it possible, at least the following best practices should be considered:

- Build a tech team or find a technological partner who will help in development rather than interested in selling ready to use solution; building data platform is an ongoing process, not a single deployment.

- Build a deep understanding of concepts like data product and data domain, they should be widely accepted and introduced to everyday life on all city management levels.

- Understand value of the data products for the community, business and city management; openness for data commercial usage can bring new solutions or services for citizens; data driven decision are always better justified and more conscious.

- Consider delivering data products as an opportunity and foundation for becoming smart city rather than threat to anyone's interests.

- Provide clear data governance rules which will open potential of sharing data and advance interest of data product owners.

- Enable crowd-source aspect of data product creation, take advantage of human creativity and open data platform for experiments carried out by city residents, they experience daily inconvenience and can help solving them if possible.

- Consider to monetise data products, at the end data is a product and is commercially utilised, define data product usage rules which will bring revenue to the city budget.

7. Beyond data mesh - foundation for building digital twin

Digital twin is a digital model of a physical space, which is built based on information gathered from sensors, IoT devices and others, using advanced machine learning models and other artificial intelligence technologies. The purpose of this model is to gain insights into performance, functioning or efficiency of the physical space [5]. In the industry world, this approach is already a standard for different use cases like: decision making support, better understanding of the market and client needs, creating better products and services, service operation improvements or data sale. When combined with city governance processes digital twin can support better decision-making by simulating consequences of the undertaken actions, explaining hidden relationships and reducing the risk related to making wrong decisions. Quality and reliability of the digital twin model depend on data quality and model quality. Since AI model correctness depends also on input data quality, it is clear that high quality input data is key for creating valuable digital twin. In this context, data platform built according to data mesh principles emerges as an interesting data source alternative,

ensuring compliance with the requirements by design. By introducing data product approach, it supports data quality improvement, by introducing data domains it brings transparency and responsibility, data governance rules ensure reliable access to data, and exploitation of standards helps to ensure interoperability of its services and hence amount of various data feeding digital twin model. Data platform delivered along data mesh design rules is a solid foundation for developing future, complex digital twins which can represent complicated phenomena of the city space.

8. Conclusions

In this article the question of building next generation data platform for modern smart city according to data mesh principles has been considered. Discussion started with the definition of a smart city in the context of data. Then, a former approach based on city of Poznan example was introduced. The experience gathered from 25 years of development were shortly summarised in the form of identified challenges. A brief description of data mesh was provided to indicate a new perspective in data platforms development principles. Since data mesh is a general concept, its adaptation for smart city specificity has been also considered. It is expected, that well designed and effectively operating data platform built according to data mesh principles will enable opportunity to implement digital twin concept for smart city.

Data mesh is a new concept, introducing a very fresh overview on how modern data platforms should be built, organized and delivered. It takes advantage from the experience of creating number of different warehouses and data lakes, pointing out weaknesses of these approaches. Data mesh concept addresses these weaknesses and proposes an alternative approach on many different levels. Originating from commercial environment, it is an interesting alternative for delivering data platform for a complex organisation, such as a modern city. Despite the fact that many aspects are already addressed, applying data mesh concept do smart city introduces some additional steps that need to be undertaken. Potential of data product thinking, a way it changes perspective of data owners, how it influences data quality, enables effective utilisation of computing resources and introduces positive relationships between different stakeholders, significantly increasing chances for building long

lasting, effective and flexible data platform - an investment defining a modern smart city.

9. References

- [1] Albino, Vito & Berardi, Umberto & Dangelico, Rosa. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*. 22. 2015. 10.1080/10630732.2014.942092. y
- [2] https://ec.europa.eu/info/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en, access 06.2022
- [3] Binczewski, A., Mazurek, C., Meyer, N., Nabrzyski, J., Niwinski, S., Stroinski, M. (1997). Information Infrastructure of the Poznan Science Society. *Proceedings of EUNIS*
- [4] Mazurek, C., Stroinski, M. (2019). Common Data and Technological Partnership - The Foundation for the Development of Smart Cities - Poznań Case Study. 10.24251/HICSS.2019.889.
- [5] Mazurek C., Stroinski M. (2022). Technology pillars for digital transformation of cities based on open software architecture for end2end data streaming. 10.24251/HICSS.2022.930.
- [6] Nahrstedt K., Cassandras C., & Catlett C., (2017). City-Scale Intelligent Systems and Platforms. arXiv preprint arXiv:1705.01990
- [7] Schultze M., Wider A. (2021). *Data Mesh in Practice* O'Reilly Media, Inc. ISBN: 9781098108496
- [8] Dehghani, Z. (2019) How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh <https://martinfowler.com/articles/data-monolith-to-mesh.html> , access 06.2022
- [9] Dehghani, Z. (2022) *Data Mesh*. O'Reilly Media, Inc. ISBN: 9781492092391
- [10] Farsi, M., Daneshkhah, A., Hosseinian-Far, A., Jahankhani, H.. (2019). *Digital Twin Technologies and Smart Cities*. 10.1007/978-3-030-18732-3.
- [11] Dehghani, Z. (2020) *Data Mesh Principles and Logical Architecture* <https://martinfowler.com/articles/data-mesh-principles.html> , access 06.2022
- [12] Technology for smart cities: the pillars of urban planning of the future. *Tomorrow City*, <https://tomorrow.city/a/technology-for-smart-cities-the-pillars-of-urban-planning-of-the-future> Accessed on 20210730
- [13] Budden, R. (2021), What “smart city” means for 2021: How digital twins, AI and other innovations drive smart transformation <https://www.americancityandcounty.com/2021/06/16/what-smart-city-means-for-2021-how-digital-twins-ai-and-other-innovations-drive-smart-transformation/> access 07.2021
- [14] Elmaghraby AS, Losavio MM. (2014) Cyber security challenges in Smart Cities: Safety, security and privacy. *Journal of advanced research*. 5(4):491-7
- [15] Europe's way to the information society: an action plan COMMISSION OF THE EUROPEAN COMMUNITIES COM(94) 347, Brussels, 19.07.1994
- [16] Borkowski M., (2017) “Civil society needs and expectations: What does NGI do for the people?”, *The New Generation Internet Workshop*, Poznań, 8.06.2017r
- [17] Asch et al., (2018) Big data and extreme-scale computing: Pathways to Convergence Toward a shaping strategy for a future software and data ecosystem for scientific inquiry. *The International Journal of High Performance Computing Applications*, Vol. 32(4) 435–479