International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:14, No:4, 2020

Government (Big) Data Ecosystem: Definition, Classification of Actors, and Their Roles

Syed Iftikhar Hussain Shah, Vasilis Peristeras, Ioannis Magnisalis

***Abstract*—**Organizations, including governments, generate (big) data that are high in volume, velocity, veracity, and come from a variety of sources. Public Administrations are using (big) data, implementing base registries, and enforcing data sharing within the entire government to deliver (big) data related integrated services, provision of insights to users, and for good governance. Government (Big) data ecosystem actors represent distinct entities that provide data, consume data, manipulate data to offer paid services, and extend data services like data storage, hosting services to other actors. In this research work, we perform a systematic literature review. The key objectives of this paper are to propose a robust definition of government (big) data ecosystem and a classification of government (big) data ecosystem actors and their roles. We showcase a graphical view of actors, roles, and their relationship in the government (big) data ecosystem. We also discuss our research findings. We did not find too much published research articles about the government (big) data ecosystem, including its definition and classification of actors and their roles. Therefore, we lent ideas for the government (big) data ecosystem from numerous areas that include scientific research data, humanitarian data, open government data, industry data, in the literature.

***Keywords*—**Big data, big data ecosystem, classification of big data actors, big data actors roles, definition of government (big) data ecosystem, data-driven government, eGovernment, gaps in data ecosystems, government (big) data, public administration, systematic literature review.

I. Introduction

IN today's data-driven world, organizations are making efforts to create a data-driven culture in public organizations. To achieve a data-driven culture, organizations are adhering to basic (big) data principles. Such principles are about data generation, data storage, access to data, free flow of data, sharing of data, data publishing, data management, data analysis, data re(use), data protection, data privacy, and data preservation [1]-[3]. Globally, organizations are adopting state-of-the-art technological (big) data solutions to realize their value of (big) data, to promote data-driven decision making, and discover new business prospects [4]-[6].

Data are facts and figures about an object, and organizations usually process data, including raw data, as per their needs [4]. Organizations create, gather, and store data in different forms like textual, numeric, images, audio, and videos [7], [8]. Data are key strategic asset for the private, public sectors, and civil society. Organizations, particularly public sector organizations, are developing innovative abilities to transform data into information and knowledge for the data-driven government [4], [9].

Pospiech and Felden define big data as it can mean big volume, big variety, and big velocity. [10]. They further added that it is difficult to process big data without using cost-effective and unique innovative technological tools and analytical techniques [10]-[12].

Big Data is a paradigm shift in the perception of approaches to understand and study the world. Organizations are using big data to analyze fine-grained data to create numerous opportunities [13], [14]. In public sector organizations, such opportunities include efficient public service delivery, enablement of data-driven decision making for policymakers, enhancement of country digital economy, creation of new jobs for the youth, promote civic participation to define and improve public policies [12]-[15], and boost data value creation for the businesses [15]-[17].

We noticed different sources of (big) data in the literature. Examples of key (big) data sources include the smart mobile handsets, online social networks, Internet of Things (IoT), cloud computing solutions, and smart cities [18]-[20]. Public administrations process such sources of big data related to various public sectors like health, education, agriculture to promote data-driven administration [21]-[23].

We found different definitions of (big) data ecosystems in the literature. In these (big) data ecosystem definitions, we observed a common viewpoint amongst the research community that (big) data ecosystem is a network of different elements. We also noted that different elements of (big) data ecosystems like data, people, organizations, organizational procedures, and technology [24]. [25] define a big data ecosystem as a network of people and technologies to collect, handle, and use the (big) data and the interactions with each other [25].

Governments are aiming to create public value by accomplishing the needs and wishes of the public. Governments are implementing (big) data ecosystem in the public organizations to achieve such aims. Public administrations create, refine, store, analyze, access, manage, share, publish, re(use), protect, preserve data through (big) data ecosystem. Such data may be related to government employees, courts, taxes, agriculture crops, crimes. Moreover, (big) data ecosystem is the fundamental driver and enabler for

Syed Iftikhar Hussain Shah, Researcher and Ph.D. Candidate in eGovernment/Information Systems Management is with the School of Science & Technology, International Hellenic University, 14th km Thessaloniki - Moudania, 57001 Thermi, Greece (phone: +44-7932606032; e-mail: i.shah@ihu.edu.gr).

Dr. Vasilis Peristeras, Assist: Professor, and Dr. Ioannis Magnisalis, Associate of Assist Professor Vasilis Peristeras, are with the School of Science & Technology, International Hellenic University, 14th km Thessaloniki - Moudania, 57001 Thermi, Greece (e-mail: v.peristeras@ihu.edu.gr, i.magnisalis@ihu.edu.gr).

the data-driven government [24], [26].

The main goals of this research are to propose a robust definition of government (big) data ecosystem and a holistic classification of government (big) data ecosystem actors and their roles.

In the literature, we found about 25 research studies that had attempted to define the data ecosystem. However, these research studies had narrow perspectives and focused on a specific concept with limited details [27]-[30]. Moreover, [31], [32] explained generic ecosystems' definitions to fill the definition space in their research studies. In the literature, we did not find a well-established definition of government (big) data ecosystem.

To address the literature gap mentioned above, we proposed a definition of government (big) data ecosystem. Our proposed definition of government (big) data ecosystem consists of the following three main concepts that we revealed in the literature definitions of (big) data ecosystem. The first concept is "socio-technical network"; the second concept is "data functions", and the last central concept is "data value creation". We describe our proposed definition of government (big) data ecosystem and its associated information in the forthcoming sections.

In the literature, (big) data ecosystem research studies adopted a heterogeneous theoretical foundation to define (big) data ecosystems. The examples of most common such theories include socio-technical theory and value chain theory. Such mixed theories are adopted because (big) data field is in its infancy. Moreover, various research and industry communities have been exploring the (big) data field separately [33]-[35]. We also followed the same approach in terms of theoretical foundation to define government (big) data ecosystem.

Government (Big) data ecosystem actors represent distinct entities that provide data, consume data, manipulate data to offer paid services, and extend data services (e.g. data storage and hosting services) to other actors. For example, actors who provide or publish data to other actors in the government (big) data ecosystem is called data publishers or data providers. In the literature, actors are put together in different groups [32], [36]-[38]. We called actors group as a class.

Each actor's class has a different set of sectors and communities. For example, data publishers' class relevant sectors and communities include public sector/agencies [39]-[41], NGOs [42] municipal, state or federal government, and local government [38].

Each class has a different role and motivation. For example, the roles of the data publishers' class are to define data-driven plans/policies, create and gather data, store data, and publish data freely. The motive of the data publisher class is to work for better governance and to improve the quality of life of the citizens [37], [40], [41], [43], [44].

In the literature, we also did not find harmonization in (big) data actors and their roles. This problem took place as authors give different titles for actors and assigned contradictory roles [32], [42]-[44].

We also proposed a classification of government (big) data ecosystem actors and their roles. Detailed information about our proposed classification of actors, roles, is described in the forthcoming sections.

We found research articles on various areas based (big) data ecosystem. We noticed that these research articles based on areas like scientific research, semantic web & web content management, open government, and business. We did not find many research articles specifically on government (big) data ecosystem, including its definition and classification of the government (big) data ecosystem actors and their roles. Therefore, we borrowed ideas for the government (big) data ecosystem from the existing literature on the areas above based (big) data ecosystem.

This paper is a continuation of our last research paper, under review, and based on our systematic literature review regarding government (big) data ecosystem [37]. It is pertinent to mention that due to the common source of our past and current research articles, there is a possibility that readers of both articles may find some recurrence in the contents, especially in the initial sections.

The remainder of the research paper is organized as follows. In Section II, we present the research method and literature search process and its outcomes. In Section III, we describe our results. In Section IV, we include a detailed analysis of literature definitions of (big) data ecosystem and describe our proposed definition of government (big) data ecosystem. We present a graphical view of actors, roles, and their relationship in the government (big) data ecosystem. We also discuss our research findings, and lastly, in Section V, we describe our conclusion.

II. Research Method

We carried out a systematic literature review about the government (big) data ecosystem to find and organize research articles for the analysis of the literature review. We followed the research approach as per the proposed guidelines of Kitchenham [45]. The details about our research goals, research question, and the literature search process are as below:

A. Goal and Research Questions

The goal of our research is to emphasize the government (big) data ecosystem definition, and classification of actors, and their roles. For this, we framed research questions: RQ1: what is the definition of government (big) data ecosystem? RQ2: What is the classification of government (big) data ecosystem actors and their roles? The RQ1 aims to describe literature definitions of (big) data ecosystem, whereas RQ2 aims to explain the classification of government (big) data ecosystem actors and their roles. In the analysis and discussion section, we include a detailed analysis of literature definitions of (big) data ecosystem and describe our proposed definition of government (big) data ecosystem. We present a graphical view of actors, roles, and their relationship in the government (big) data ecosystem. We also discuss our research findings.

We attain our research goals and find out the answers to the mentioned-above questions by analysis of literature. We

explain the mentioned-above aspects in the forthcoming sections.

B. Literature Search Process

In this section, we describe information about the selection of electronic research libraries, relevant keywords, search criteria, and literature search process results.

We explored four electronic libraries: ACM Digital Library, IEEE Xplore Digital Library, Science Direct, and Springer Link. We did not specify articles publication year range to search research articles from the aforesaid electronic libraries. We only included English language articles in our literature review. We assigned high priority to choose a content type of journal. However, we did not find many research articles about (big) data ecosystem from the journals. We included relevant research articles from other sources to get. Such other sources include proceedings of international conferences, advanced countries' relevant policies, strategies, Acts, and case studies. We started the search process to find out relevant research articles in February 2019 and continue this process until December 2019. We also included secondary references in our subject literature review. We selected the secondary references from our primary paper references.

We performed the search process using online systems of above mentioned electronic libraries and manually explored official websites of some advanced countries, including the EU and its member states as well. We performed the search process in the following four phases to find out relevant research articles about require topics of government (big) data ecosystem. We performed searches on the following keyword strings in the research paper title, keywords, and abstract:

- *Phase-I:* We utilize mentioned-above electronic libraries to search the keywords "DATA ACTORS", "DATA ACTORS ROLES", "CLASSIFICATION OF DATA ACTORS", DATA PROVIDER", "DATA USERS", "DATA BUSINESS ENTITY", "DATA SUPPORT SERVICE PROVIDER", "Data Ecosystem", and "DATA-DRIVEN GOVERNMENT" along with choices "exact phase" and "matches all." We found a limited number of relevant research articles from the mentioned-above phase1 of our literature search process.
- *Phase-II:* In the phase-II of our literature search process, we performed a search on the keywords that are mentioned-above phase-I along with new choices like "matches any".

We investigated the results of phase-I and phase-II and compared them with our essential topics of government (big) data ecosystems to be examined and described in the subject research work. We observed that still, we need more relevant research articles.

- *Phase-III:* In the phase-III of our literature search process, we performed a search on the additional keywords like "data policy", "digital agenda", "data-driven government".
- *Phase-IV:* In the phase-IV of our literature search process, we manually explored official websites of some advanced countries, including the EU and its member states as well.

The final number of papers that we gathered after removing duplicates is 1021. Then, papers having titles irrelevant to our research were also identified manually and excluded in our next review phase, so it reduced the papers to 512. The procedure mentioned above was repeated by scanning the papers' abstracts, reaching 357 papers. We read the whole text of these papers, culminating in 294 of them for the future course of action. We presented our literature search process results in Fig. 1.

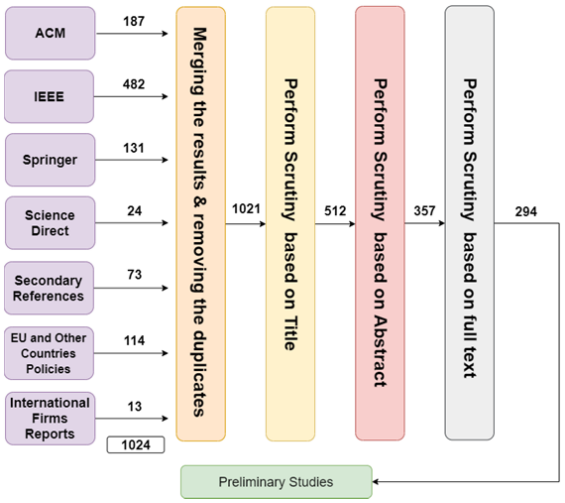

Fig. 1 Procedure for identifying preliminary studies

III. Results

We thoroughly examined 294 research articles that we already found papers through the above mentioned literature research process. We extracted relevant information from these papers. We organized information within the following categories: definition of government (big) data ecosystem, types of government (big) data, data lifecycle for the next generation data-driven government, government (big) data actors and their roles.

In this research article, we detailed our results, based on above mentioned research questions, about the following 2 elements, as shown in Table I.

TABLE I
Government (Big) Data Ecosystem 2 Key Elements and Their Respective Total Numbers of Papers

Elements	Government (big) data ecosystem Definition	Actors and their Roles
Total papers	36	18

In Table I, we mentioned that 36 research articles about the definition of government (big) data ecosystem and 18 research articles about government (big) data ecosystem actors and their roles that we included in our preliminary studies. The details of government (big) data ecosystem and government (big) data ecosystem actors and their roles are as below:

A. Results for RQ1 – Definition of Government (Big) Data Ecosystem

We have found about 25 research studies, which had attempted to define the data ecosystems. Here we present some definitions of data ecosystems found for the data ecosystem concept in different contexts. References [40], [46], [47] define the data ecosystem as a network of humanitarian actors, governments, and private sector organizations, and affected communities in which they interact with each other to produce, collect and analyze digital data about vulnerable populations — the actors in the humanitarian data ecosystem exchange data for disaster management through close coordination. Reference [47] added that sources of crisis-affected communities data include mobile phone records, social media posts, satellite imagery, sensor data, financial transactions. The said definition pertains to the humanitarian domain. Parsons et al. define the data ecosystem, specific to the scientific research field, as the people and technologies are collecting, handling, and using the (big) data and the interactions between them [25]. References [32] and [48] define data ecosystems as socio-technical complex networks in which actors (organizations and individuals) interact and collaborate to exchange and use data as the primary source to foster innovation, create value, and support new businesses. Reference [45] defines (big) data ecosystem as a heterogeneous network of software, hardware, and networking resources, human capital (such as skills), industry applications and methodological techniques, social actors, and the new ideas and concepts those actors coin [49].

We noted that some data ecosystems clearly stated about open data and its entire ecosystem in the literature [31], [50]-[52], [39]. Reference [37] defines an open data ecosystem. For example, [32] defines open data ecosystem as a complex of various actors, having light interaction with each other, and usually perform open data functions like create, find, store, open, access, share, protect, preserve, and feedback [37].

During the literature review, we also noted that some data ecosystems definitions be relevant to certain domains like the humanitarian [47], [40], [46], and personal [53] data ecosystems. Reference [35] was of the view that definitions of such domains-based data ecosystems are associated with an environment where an ecosystem emerges [35].

To address the mentioned above literature gap, we proposed a robust definition of government (big) data ecosystem. First, we explain the following identified three main concepts in the literature regarding the definition of government (big) data ecosystem:

1. Identification of the Concepts in Literature about Government Data Ecosystems Definitions

During the investigation of existing literature definitions, we identified three main concepts. The first concept is "socio-technical network," which is about the collaboration of socio-technical elements such as people, processes, technology, organizations, data, and infrastructure. The second concept is "data functions," which consist of different phases (data collection, data integration, analysis, data storage, sharing, use, data security and protection, and data archive) to transform data/information into knowledge. The third and last central concept is "data value creation," which is about the extraction of value from the data (big). These values (public service delivery, data-driven administration, transparency, data economy, and new businesses) are the outcome of the data functions performed on data in a collaborative environment by the stakeholders.

In the literature, mentioned above three concepts did not exist together in any single data ecosystem definition. Moreover, we noticed that some literature definitions consist of first and third concepts. Some (big) data ecosystem definitions consist of second and third concepts, and other definitions consist of first and second concepts.

2. Grouping Three Main Concepts

For a robust definition of the data ecosystem, we grouped three identified main three concepts. The pictorial presentation of our proposed definition of government (big) data ecosystem that consists of n these concepts extracted from the literature definitions is shown in Fig. 2.

3. Explanation about How Mentioned Concepts Help Someone to Understand Better the Government (Big) Data Ecosystem

Our government (big) data ecosystem definition consists of following three concepts. Our self-explanatory definition gives a holistic understanding of the data ecosystem and its core components to the readers. Our proposed definition first concept gives understanding to the readers that (big) data ecosystem consists of socio-technical elements, including organizations require collaborative efforts to process the data (big) as per their needs [54]. The second concept emphasizes on a data lifecycle through which organizations process raw data and transform information into knowledge. Our definition of the last concept conveys a clear message to the stakeholders that besides keeping big data, they should also focus on the extraction of value from the data (big) for their benefits.

We describe a detailed analysis of literature definitions of (big) data ecosystem and present our robust definition of government (big) data ecosystem in the forthcoming section *"Analysis and Discussion"*.

B. Results for RQ2: Classification of Government (Big) Data Actors, and Their Roles

Actors represent distinct entities that provide data, consume data, manipulate data to offer paid services, and extend data services (e.g., data storage and hosting services) to other actors in the government (big) data ecosystem. Actors are put together in distinct groups, and we call this group as a class. Each class has a different set of sectors and communities. Moreover, such a class has a different role and motivation as well.

As we mentioned above in the literature, we did not find harmonization in types of actors and their roles. This problem took place as authors give different titles for actors and assigned contradictory roles [32], [42]-[44].

To describe literature on government (big) data ecosystem

actors and their roles in a better presentable way, we establish a classification of government (big) data ecosystem actors and their roles, as shown in Fig. 3.

The explanation of each class of government (big) data ecosystem actors and their roles are as below that include definition of each actor class, sub-types, relevant sectors and communities, actor class roles and motivation, and some real examples as well:

1. Data Publishers (Providers)

Definition

The data publishers (providers) typically represent distinct entities that amply provide data to other relevant actors of the data ecosystem [32]. The provided data can represent raw data, redefined data/information, or analyzed information. The data publishers give data to the data users for free or with some licenses that restrict the use of data for commercial purposes. For example, a Creative Commons (CC) license is a public copyright license that enables the free distribution of an otherwise copyrighted work [42], [32].

Types of Data Publishers

There are two types of data publishers. The first type represents the entities that provide data for free and without any condition or with some licenses. Such a data license restricts the use of data. The second type of data publishers that do business from selling access to the data [55], [56].

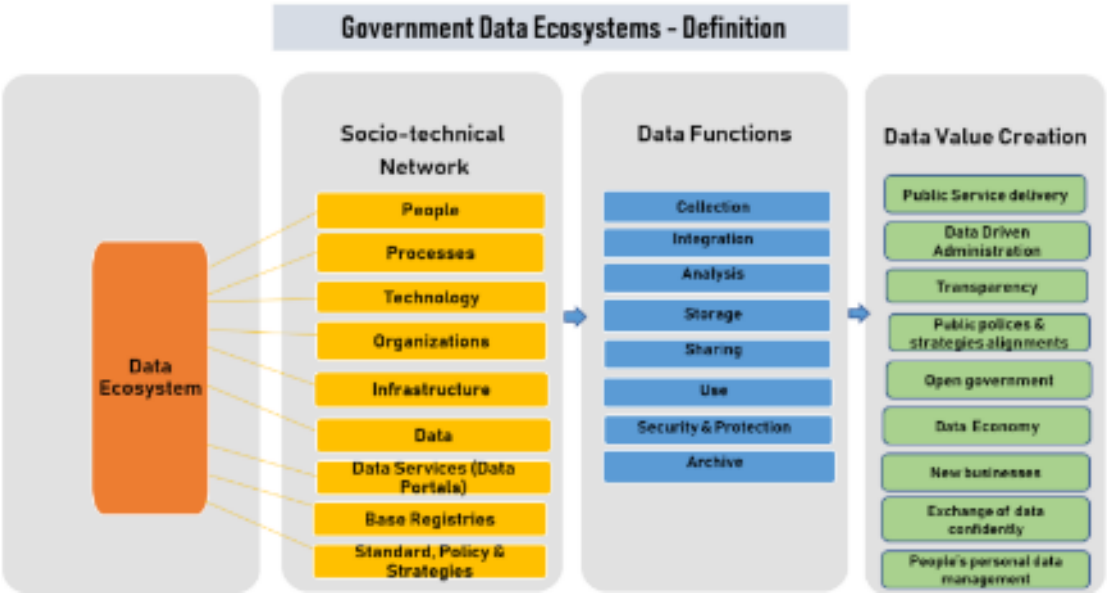

Fig. 2 Three main concepts - Definition of Government (Big) Data Ecosystem

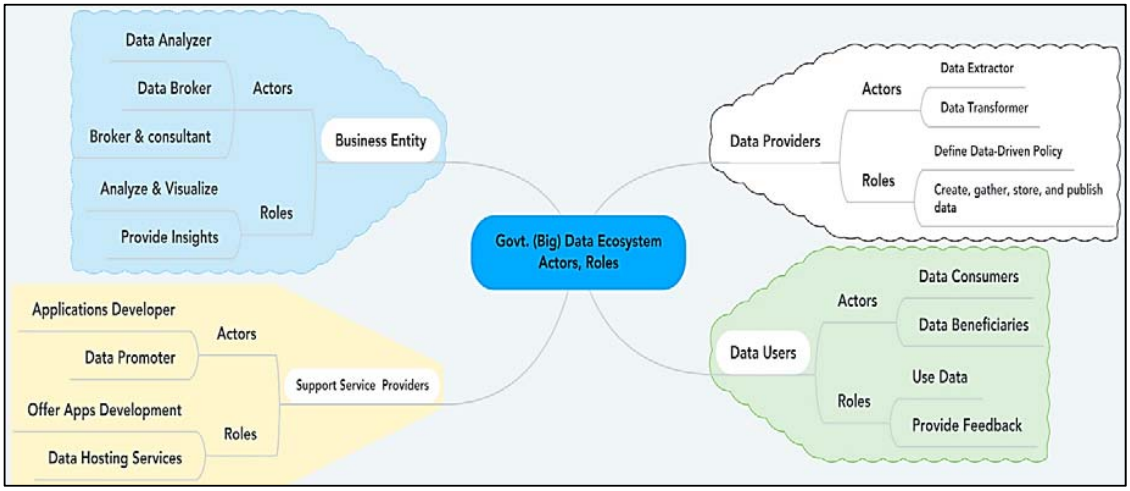

Fig. 3 Classification of Government (Big) Data Actors and their Roles

The first type of entity usually represents public administrations or other public entities. They provide data to

improve the national economy, enabling enterprises and citizens to exploit the data [55]. The second type of entity usually sells data. They engage users to pay for data. Such entities provide only a subset of their data as open data, provide access to specific data by users as per signed data-sharing agreement [56].

Relevant Sectors and Communities

Data publishers' relevant sectors and communities include the public sector/agencies [43], [40], [41], NGOs [42], municipal, state or federal government, and local government [38]. The provided data can represent raw data, redefined data/ information, or analyzed information [32], [38].

Roles and Motivation

The role of the data publishers is to define data-driven plans/policies, create and gather data, store data, and publish data freely. Data publishers also participate in stakeholder joint activity "feedback and discussion". The motive of the data publisher is to work for better governance and to maximize the quality of life of the citizens [37], [40], [41], [43], [44]. Zuiderwijk et al. were of the view that data providers and data users depend critically on each other. In the literature, we also noted that a commercial value of raw data could only achieve when both actors collaborate [50]. The data publishers also stimulate the participation of citizens in governmental processes of decision making and policymaking [57], [32]. The responsible government data agencies may not only share data with data users freely but also get back in return value-added input from them [41]. Data users and business entities perform supportive roles in publishing the data in the data ecosystems.

Real Examples of Data Publishers (Providers)

In this sub-section, we share real examples of data publishers. The first example is the Helsinki Regional Transport Authority (HRTA), a data publisher, has developed a free API that provides access to information to data users. The data users, including service providers, use this API to help customers with trip planning. Through HRTA API, data users can access different kinds of information. These kinds of information include public-transportation routes and timetables, service disruptions, and live data for vehicle location and tracking. Moreover, mobile Apps developers used the HRTA API to create about 30 mobile trip-planning applications [38]. Another example is the Russian City "St. Petersburg" open government portal [58]. The portal contained 195 datasets from 46 data providers until July 2019. The data providers include the administration of St. Petersburg, Administration of the Governor of St. Petersburg, the Committee on Energy and Engineering. [54]. The EU open data portal contained 13801 datasets of different areas from 12 data providers until July 2019 [59]. Data providers include the European Parliament, the Council of the European Union, the European Commission. The EU and St. Petersburg portals also introduced mobile apps to discover the potential of both open data portals.

2. Data Users

Definition

The data user is an entity that consumes data that are provided by the data publishers in a data ecosystem. Data users may consume information manually or with the help of data-based applications and services [36]. Dawes et al. stated that when the government itself is an active data user, it seems more capable of discovering the needs of external users and more likely to offer appropriate data and information to people [54]. Haak et al. were of the view that data publishers and users should mutually trust and be transparent to each other while providing data and employing it in the data ecosystems [40].

Types of Data Users

A data user can be a consumer, a citizen, or an enterprise user. A consumer is a user that bought a commercial application from an application store. A citizen can be the user that uses the provided application as a citizen. An enterprise user is a user who uses the applications in business. For example, the application can enable a user to produce information from the environment and then consume the information provided by the public administration [36]. Rao et al. were of the view that data consumers include citizens in general, who will use services developed by re-users. Reference [42] added that data consumers also include a civil society that may use and publish research outcomes and request more information and services. Zuiderwijk et al. stated that data users include data infomediaries or intermediaries. The infomediaries process the raw data and add value to the data through cleaning, analyzing and integrating the data. Other users may prefer to make use of the processed data or services and tools that are derived from raw data by infomediaries [32].

Relevant Sectors and Communities

Data users' relevant sectors and communities include local community/individuals, public sector [43], NGO, civil society [42], private sector, academics [60].

Roles and Motivation

The roles of the data users are to use data, provide feedback, do R&D to investigate new algorithms, technologies, publish research [54]. Smith et al. stated that data users perform the roles of finding data, analysis of data, data processing, and participate in the stakeholder joint activity 'feedback and discussion' [37]. Immonen et al. stated that data users' roles include an application developer to utilize the data as part of the service and an interpreter to interpret the data [36]. Zuiderwijk et al. explained that data users could conduct different activities on the data and provide feedback to the data publishers and other stakeholders. The examples of such activities include searching, finding, integrating, filtering, analyzing, visualizing data [50]. Infomediaries are providing data management as a service. Other organizations and individuals use this service [41]. The motivations of the data users are to promote community

welfare, business growth, and to encourage civil society participation to enhance the quality of the data [54], [36].

3. Data Business Entity

Definition

The data business entities integrate their data with the publisher's data to offer data services, and they gain income from the usage of their services by external organizations. They offer essential data service to both data publishers and users [37], [61], [36].

Types of Data Business Entity

There are different types of data business entities. Such types of business entities include data application developers [36], data aggregator [36], [44], data harmonizer [36], [42], data analyzer [38], data enablers [61], data brokers [37], [36], data facilitators, data re-users [42], data marketplace companies, data consultants [37]. The application developers innovates applications around the available data and integrates their organizational data. Data publishers and data support service providers deliver data and services for applications [36]. Data aggregators combine and modify data. They collect data from different sources [36], [43]. Data harmonizers perform standardization and homogenization of data. Data harmonizers also gather different data sources [36], [44]. The data analyzer gathers and analyzes data [38]. Data enablers provide solutions and services to data publishers to combine different types of data. Such types of data include users' locations, medical records, a product design by crowdsourcing, and third-party data. Data brokers have their sub-types, which include data promoters, distributors, and matchmakers. Data broker sells personal data about individuals to other third parties without a data owner's permission [62]. This way of selling personal data creates doubts among the community, and they do not trust the current ecosystems.

The data promoter finds out data and advertises it to the actors. Data distributor provides the communication and distribution channels of data. Data matchmakers match the data demand with the best available data source [38]. Data facilitators aid with the exchange of data between the data publishers and data users in the data ecosystem [61]. Online data marketplace companies use websites or mobile applications. Such technological solutions connect buyers and sellers in an open, cooperative environment [37]. Data re-users include entrepreneurs, companies, IT organizations, and universities that use data to develop applications or services aimed at citizens or to do research [42]. Data consultants provide consultancy services to the other actors in the data ecosystem

Relevant Sectors and Communities

Data business entities' relevant sector and communities include the private sector only [43].

Roles and Motivation

Data business entities analyze gathered data from various sources and provide channels for communication and distribution of data. They visualize data, advertising data to the other actors, and provide insights into the data. Data business entities match the data users' demand with the best available data source. Data business entities provide data consultancy services to the companies. The consultancy services are helpful for the companies to determine customers' needs and utilization areas of the present data [38].

In some cases, they performed an intermediator role between data providers and data users and arranged links between these parties [43], [42], [38]. Data business entities perform a crucial role in the exchange of data between data publishers and data users [61]. The motivations of the data business entities include commercialization, business growth, and earn money [36].

Real Examples of Data Business Entities

Lindman et al. presented an example of a data analyzer company. The company uses open government data and private data acquired from Finnish firms, to produce credit ratings and other financial information for sale. They described another example of a data analyzer company that analyzes business' financial data and draws an easy-to visualize, a tree-shaped image of their balance sheets [38]. Magalhaes et al. highlighted an example of data enablers companies 'Captricity' and 'Xcential' that assist governments in transforming static documents into actionable data [61]. Quandl, a data enabler tool that helps users' to explores complex finance, economics, society, health, energy, and demography data. PolicyMap is a data enabler tool that supports entities to map public data [61]. eBay, an e-commerce platform provider, Uber, and Cream, ride-hailing service providers, are examples of online data marketplace companies [37].

4. Data Support Service Providers

Definition

The data support service provider provides support services to the other actors in the data ecosystems. Such services include data hosting, data storage, and the design & development of mobile applications and websites [36], [38].

Types of Data Support Service Providers

There are different types of data support service providers. These types of service providers include cloud computing service providers, mobile applications, website design and development service providers, and user-experience providers [38], [36]. The cloud computing service providers deliver the physical facilities for the data ecosystem and receive income from the facilities 'rent'. The mobile application and website development service providers offer software application development expertise and visual design to their clients in the data ecosystem [36], [61]. The user-experience providers gather and combine data sources and offer user interfaces to manipulate data through a web browser or mobile application [38].

Relevant Sectors and Communities

Data support service providers' have various relevant

sectors and communities. Such sectors and communities include the private sector, cloud computing services entities, mobile application & website development entities, and other public sector supporting agencies [32], [36], [38].

Roles and Motivation

Data support service providers are providing cloud-based data hosting services, mobile applications, and website development expertise to their clients in the data ecosystem Data support service providers' motivations are to generate revenue by provisioning data hosting services, selling mobile applications, and website development [37], [50].

Real Examples of Data Support Service Providers

Amazon Web Services (AWS) offers cloud-based platforms on a pay-per-use basis to the clients. These clients may include individuals, companies, and governments. For example, a company is providing data support services to their clients by creating a website showing job advertisements [38]. A company website also improves through additional information from other sources. The examples of such resources include online maps, social media, financial data, and news feeds. In the literature, we also observed that a company is getting revenue by selling website subscriptions or advertisements shown within the sites [37].

In the literature, we did not find a holistic classification of government (big) data ecosystem. Our aforesaid proposed classification of government (big) data ecosystem actors and their roles address the mentioned above literature gap and provides a comprehensive set of actors and gives clarity in actors' roles.

IV. Analysis and Discussion

In this section, we describe a detailed analysis of literature definitions of (big) data ecosystem, describe our proposed definition of government (big) data ecosystem. We present a graphical view of actors, roles, and their relationship in the government (big) data ecosystem. We also discuss our research findings.

A. Analysis of Literature Definitions of (Big) Data, and our Proposed Definition of Government (Big) Data Ecosystem:

In the literature, we found about 25 research studies, which had attempted to define the data ecosystem. However, these research studies had a narrow perspective and focused on a specific concept with limited details [27]-[30].

To address the mentioned above literature gaps, we performed an analysis of the three concepts, shown in Fig. 2, regarding the definition of government (big) data ecosystem and proposed a definition of Government (big) data ecosystem.

In the literature, (big) data ecosystem research studies adopted a heterogeneous theoretical foundation to define (big) data ecosystems. The example of most common such theories includes socio-technical theory and value chain theory. Such mixed theories are adopted because (big) data field is in its infancy. Moreover, various research and industry communities have been exploring the (big) data field separately [33]-[35]. We also followed the same approach in terms of theoretical foundation to define government (big) data ecosystem.

1. Overview of Three Concepts (C) Regarding Definition of Government (Big) Data Ecosystem, and Our Proposed Definition of Government (Big) Data Ecosystem

An overview of three Concepts (C), as shown in Fig. 2, and our proposed definition of government (big) data ecosystem is as below:

C1 - Socio-Technical Network

In the literature, we observed that the "socio-technical network" concept mostly exists in all literature definitions. However, different authors assigned different labels to this concept in their respective definitions. The examples of such labels include "interconnected human & technological resources", "complex network of individuals and organizations", "complex interconnected, multilayered ecosystem", and "heterogeneous network of software, hardware, people, and processes". We reviewed all these labels and assigned a unique label to this concept as a "socio-technical network". This section also possesses different characteristics of ecosystems such as: i) emphasize on the behavior of the people supported by the technology [51], ii) interdependent elements (e.g., networks of interactions, risk assessments in organizations) and interdependent components [63] and iii) composed of interconnected, interrelated and interdependent digital species situated in a digital environment.

We found these concept elements ('people', 'processes', 'technology', 'organizations', 'data', and 'infrastructure') from literature definitions. Moreover, we also opted for some elements from literature other than the part of the definition. The examples of such elements are 'data services (data portals)', and 'base registries'. We will explain these elements in our next research work segment "component of the data ecosystem." In this concept, some elements represent the social part (e.g. people, and organizations.) while other elements (e.g. technology and data portals.) represent the technological part.

Our proposed definition starts from the first concept which we identified from the literature definitions i.e. "A socio-technical network [37], [43], [64], [48], [49], [65] of people, processes, technology, infrastructure, data services, base registries standards & policies, processes, organizations …...". This concept focuses on the social and technical resources to work jointly to transform data into knowledge. However, our definition also contains aforesaid second and third identified main concepts as well.

C2-Data Functions

This concept is found in some literature definitions and not in all definitions as compared to the first concept. Some authors [54], [66], [32], [52], [49], [65] consider the first and third concepts only in their respective data ecosystems definitions. This concept consists of different phases such as data collection, data integration, analysis, data storage,

sharing, use, data security and protection, and data archive. These phases are important to transform data/information into knowledge. This concept is also helpful in identifying dataflows and work processes for stakeholders in the data ecosystem [14]. This section also possesses different characteristics of ecosystems such as i) data cycles with feedback loops, sharing of data back to publishers and sharing between so-called infomediaries, and ii) infomediaries to publish and share what public organizations produce [67].

We found "data functions" concept elements like 'collection', 'integration', 'analysis', 'storage', 'sharing', 'use', 'security & protection', and 'archive' from literature definitions. We also found such elements from other 50 (big) research studies about data lifecycle models that are published during the last 25 years period. We considered these 8 data functions as mandatory, and the other eleven phases (e.g. 'planning', 'filtering', and 'destruction') are optional.

Our definition starts from the first concept and followed by the second concept which we already identified from the literature i.e. "A socio-technical network [37], [43], [64], [48], [49], [65] of people, processes, technology, infrastructure, data services, base registries standards & policies, processes, organizations and resources jointly work to perform data functions (collection, integration, analysis, storage, sharing, use, data security & protection and Archive) [68]-[75] in order ….". This concept focuses on the data functions. These data functions usually perform to transform data into knowledge. However, our definition also contains identified the third main concept as well.

Data Value Creation

This concept is found in a few literature definitions as compared to the first and second concepts. Some authors [25], [29], [30], [37], [49], [52], [64], [67], [76]-[78], did not consider this concept in their respective data ecosystems definitions. It is about the extraction of value from the data (big). Harrison et al. [51] mentioned that someone gauge data value by knowing about the context of the use of data. The examples of data value creation from data (big) are data-driven administration, public service delivery, and open government. Data by itself has no value. However, the extraordinary value-adding potential of data lies in the ability to extract meaningful and actionable information from it. Hence, the data value creation concept focuses on value creation junction with our identified first and second concepts.

We found a list of data value creation from literature definitions and other relevant twenty plus research studies on this concept as well. The possible values are public service delivery, data driven administration [79], new businesses [53], [50], [32], data economy [54], innovation [35], [54], [51], transparency [80], [54], [63], [38], [18], public polices & strategies alignments [81], [79], open government [38], exchange of data confidently [49], people's personal data management [82], government performance [63], democratic governance, and political participation, trust in government, data reuse and integration of public and private data [54].

We further explore (big) data ecosystem literature definitions that somehow consist of one or two of the concepts mentioned above. We presented a summary of the literature (big) data ecosystem definitions and occurrence of the mentioned above three main concepts in Table II.

TABLE II
SUMMARY OF RESEARCH (BIG) DATA ECOSYSTEM DEFINITIONS AND THE EXISTENCE OF THREE MAIN CONCEPTS

S#	Literature Definitions References	C1	C2	C3
1	[27]	×	-	×
2	[35], [32], [48]	×	-	×
3	[77]	-	×	×
4	[50]	-	×	×
5	[51]	×	-	×
6	[76]	×	×	-
7	[54]	-	-	×
8	[52]	×	-	-
9	[83]	×	×	-
10	[25]	×	×	-
11	[64]	×	×	-
12	[67]	-	×	-
13	[29]	×	×	-
14	[49]	×	-	-
15	[30]	×	-	-
16	[78]	×	-	-
17	[47], [40], [46]	×	×	-
18	[84]	×	-	-
19	[37]	×	×	-
20	[85]	×	-	×
21	[28]	-	×	×

C1 - Socio-technical network, C2 - Data Functions, C3- Data Value Creation

In Table II, we describe literature definitions references due to paper space limitation. We describe the following two literature definitions as an example for the readers of this research paper. Moreover, readers of this research article may read above mentioned referenced published research studies about definitions of (big) data ecosystem. Reference [27] defined (big) data ecosystem as a system of interconnected human and technological resources working together to extract value from data and to use it for decision-making. The said definition of (big) data ecosystem consists of C1 and C3 concepts whereas [77] defined (big) data ecosystem as a kind of a data-based system where stakeholders of different sizes perform different data functions and consume open government data in connection with online tools, services, and societies. The data functions include find, manage, archive, publish, reuse, integrate, and mashup. The second definition of (big) data ecosystem consists of C2 and C3 concepts.

Our definition started with the first concept and followed by the second and third concepts that we already identified from the literature. From the above, we define the government data ecosystem as under: "A socio-technical network [37], [43], [64], [48], [49], [65] of people, processes, technology, infrastructure, data services, base registries standards & policies, processes, organizations, and resources jointly work to perform data functions (collection, integration, analysis, storage, sharing, use, data security & protection, and Archive

[68]-[75] in order to extract value [27], [43], [32], [51], [36] from data to ensure better services delivery, support the businesses, promote data-driven administration & open government, boost data economy and policy creation to benefit citizens, businesses, and government bodies itself".

Our robust definition covers the main aspects of the data (big) ecosystems identified in the literature definitions. We did not find any single data ecosystem definition in the literature, which is based on all three main concepts and so self-explanatory for the readers.

B. Classification of Actors, Roles, And Relationships

In Fig. 4, we present a graphical view of our proposed classification of actors, roles, and their relationships in the government data ecosystem.

In the use case diagram, we described eight use cases only to represent data actors' interactions in the government data ecosystem. The eight use cases consist of publish data, consume data, clean data, insights of data, data hosting, data access, feedback, and re-use. Fig. 4 also demonstrates the relationship between the actors and the different use cases in which the actor is involved. For example, in the use case 'publish data', our first class of actors *'data publisher'* is overall responsible for managing the data publishing whereas other two classes of actors *'business entity'* , *'support service providers'* perform a nurturing role. In Fig. 4, we equally represent sub-actions of the use case 'publish data.

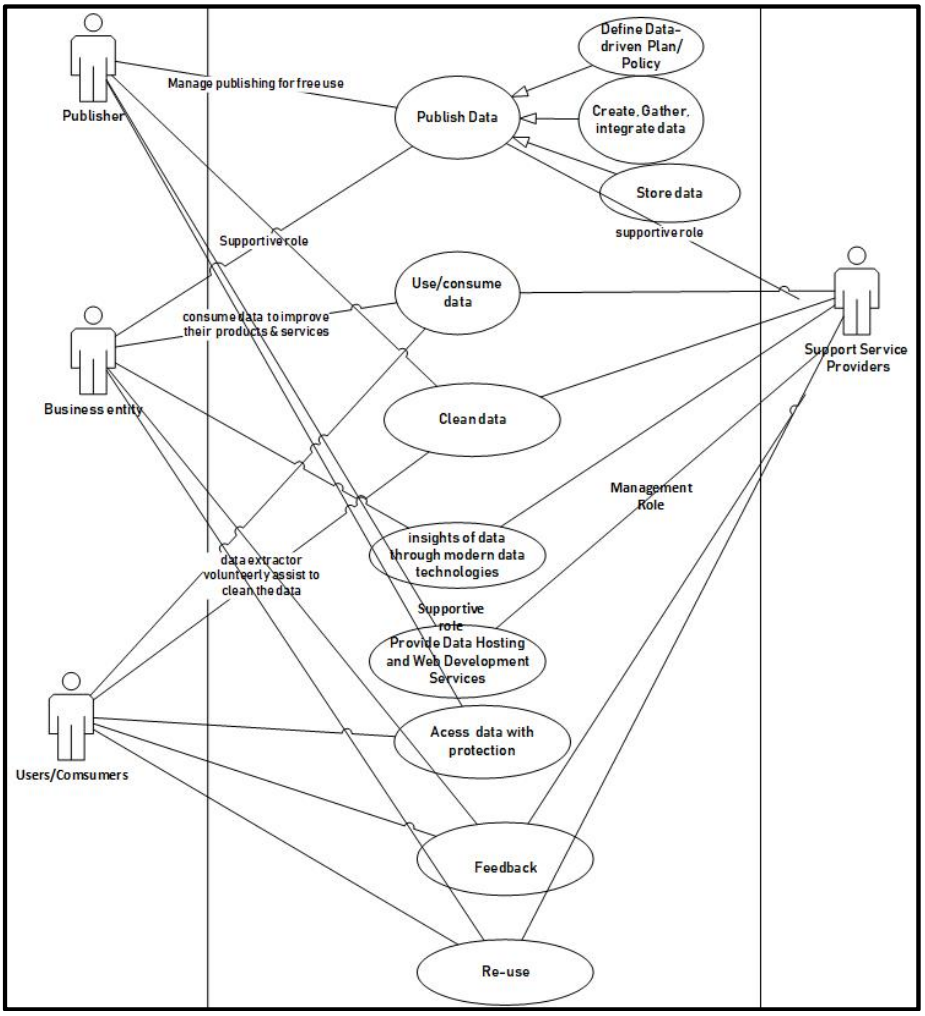

Fig. 4 Classification of actors, roles and their relationships

The sub-actions include defining the data-driven plan, creating, gathering, integrating data, and storage of data. Data publishers can perform requisite actions on raw data by themselves or may obtain assistance from other actors. In Fig. 4, another use case 're-use' has one class of actors "data users", and the other three classes of our actors perform a set of actions on the published data. They suggest improvements and corrections in the published data to the publishers.

V. Conclusion and Future Work

Government (big) data ecosystems can help the public administration to take evidence-based decisions making,

ensure data interoperability and data privacy, prioritize the problems, encourage civic participation in government in developing holistic policy development processes, and contribute to a better government.

State-of-the-art research in government (big) data ecosystem is still far behind in maturity, has an enormous scope of research; continuous and active research participation is required at our end to have new findings.

During this research work, we have found out about 25 research studies, which had attempted to define the data ecosystems. The authors use the term data ecosystem without providing details. We did not find a well-established definition of government (big) data ecosystem. To address this literature gap, we proposed a definition of government (big) data ecosystem. Our proposed definition consists of above mentioned three main concepts that we found from the literature (big) data ecosystem definitions. The three concepts include a socio-technical network. The first concept is "socio-technical network. The second concept is "data functions", and the last main concept is "data value creation". We define government (big) data ecosystem as "A socio-technical network [37], [43], [64], [48], [49], [65] of people, processes, technology, infrastructure, data services, base registries standards & policies, processes, organizations and resources jointly work to perform data functions (collection, integration, analysis, storage, sharing, use, data security & protection and Archive) [68]-[75], in order to extract value [27], [43], [32], [51], [36] from data to ensure better services delivery, support the businesses, promote data-driven administration & open government, boost data economy and policy creation to benefit citizens, businesses, and government bodies itself".

In the literature, we also did not find harmonization in (big) data actors and their roles. To address this literature gap, we propose a holistic four-dimensional classification of government (big) data ecosystem actors and their roles. The classification mentioned above provides a comprehensive set of actors and gives clarity in actors' roles as well.

As future work, we intend to explore further and analyze the literature to address the aforesaid identified gaps in the subject ecosystems. So, we will offer the following aspects: i) a new data life cycle for the data-driven governments, ii) development of a CSF framework for the government data ecosystem by incorporating six leading data ecosystem CSF dimensions: Organizational, Economic, Social, Technical, Legal, and Semantic (OESTLS), iii) a set of components of the government data ecosystem based on adequate criteria, and iv) creation of a theoretical framework of data-centric (conceptual) architecture for the data-driven government.

ACKNOWLEDGMENT

The paper is funded by the European Union-funded project Co-Inform, Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020) H2020-SC6CO-CREATION-2016-2017 (CO-CREATION FOR GROWTH AND INCLUSION), Grant Agreement 770302 and by The European Union-funded Project: Digital Europe for All (DE4A), Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020), H2020-SC6-GOVERNANCE-2018-2019-2020 (GOVERNANCE FOR THE FUTURE), Grant Agreement: 870635.

REFERENCES

[1] A. Fabijan, P. Dmitriev, H. H. Olsson, and J. Bosch, "The Evolution of Continuous Experimentation in Software Product Development: From Data to a Data-Driven Organization at Scale," *Proc. - 2017 IEEE/ACM 39th Int. Conf. Softw. Eng. ICSE 2017*, pp. 770–780, 2017.

[2] R. Fahy, J. Van Hoboken, and N. Van Eijk, "Data Privacy, Transparency and the Data-Driven Transformation of Games to Services," *2018 IEEE Games, Entertain. Media Conf. GEM 2018*, pp. 136–146, 2018.

[3] U. Aftab and G. F. Siddiqui, "Big Data Augmentation with Data Warehouse: A Survey," *Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018*, pp. 2785–2794, 2019.

[4] A. R. Pathak, M. Pandey, and S. Rautaray, "Construing the big data based on taxonomy, analytics and approaches," *Iran J. Comput. Sci.*, vol. 1, no. 4, pp. 237–259, 2018.

[5] O. R. D.J. Patil, "Building Data Science Teams." Oreilly Radar, 2011.

[6] F. Halper, "Driving Digital Transformation Using AI, and Machine Learning." TDWI, p. 31, 2019.

[7] A. Immonen and J. Kalaoja, "Requirements of an Energy Data Ecosystem," *IEEE Access*, vol. 7, pp. 111692–111708, 2019.

[8] D. R.-M. and P. T. Michael Schroeck, Rebecca Shockley, Janet Smart, "Analytics : The real-world use of big data." Saïd Business School at the University of Oxford, pp. 1–22, 2012.

[9] G. Netherlands, "Open Govt. Action Plan 2018-2020." Government of Neitherland, p. 8, 2018.

[10] M. Pospiech and C. Felden, "A descriptive big data model using grounded theory," *Proc. - 16th IEEE Int. Conf. Comput. Sci. Eng. CSE 2013*, pp. 878–885, 2013.

[11] [11] Gartner, "What is a big data." Gartner Publications, pp. 1–3, 2019.

[12] [12] and J. M. Nobubele, Angel Shozi, "Big data privacy and security: A systematic analysis of current and future challenges." University of South Africa, South Africa, pp. 1–9.

[13] OECD, "Data-driven Innovation for Growth and Well-being." OECD, pp. 1–86, 2014.

[14] S. Allard, "DataONE: Facilitating eScience through Collaboration," *J. eScience Librariansh.*, vol. 1, no. 1, pp. 4–17, 2012.

[15] S. Mazumdar, D. Seybold, K. Kritikos, and Y. Verginadis, *A survey on data storage and placement methodologies for Cloud-Big Data ecosystem*, vol. 6, no. 1. Springer International Publishing, 2019.

[16] S. J. Divakar Mysore, Shrikant Khupat, "Introduction to Big Data Architecture," pp. 1–14, 2017.

[17] KnowledgeHut, "Types of Big data." KnowledgeHut, pp. 1–4, 2019.

[18] L. Ding, V. Peristeras, and M. Hausenblas, "Linked Open Government Data [Guest editors' introduction]," *IEEE Intell. Syst.*, vol. 27, no. 3, pp. 11–15, 2012.

[19] D. Lee, "Building an open data ecosystem," pp. 351–360, 2015.

[20] D. Misra, A. Mishra, S. Babbar, and V. Gupta, "Open Government Data Policy and Indian Ecosystems," pp. 218–227, 2017.

[21] European Commission, "Data Strategy for Digital Transformation." European Commission, pp. 1–22.

[22] Statista organization, "Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions)." Statista Survey Organization.

[23] B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in iot-based information system for emergency medical services," *IEEE Trans. Ind. Informatics*, vol. 10, no. 2, pp. 1578–1586, 2014.

[24] M. M. Rantanen, S. Hyrynsalmi, and S. M. Hyrynsalmi, "Towards Ethical Data Ecosystems: A Literature Study," *Proc. - 2019 IEEE Int. Conf. Eng. Technol. Innov. ICE/ITMC 2019*, 2019.

[25] M. A. Parsons *et al.*, "A conceptual framework for managing very diverse data for complex, interdisciplinary science," *J. Inf. Sci.*, vol. 37, no. 6, pp. 555–569, 2011.

[26] IBM, "Data-driven government: Challenges and a path forward." p. 8, 2015.

[27] E. Commission, *Data Action Plan Implementing Data Strategy*. 2018.

[28] J. Attard, F. Orlandi, and S. Auer, "Data Value Networks: Enabling a New Data Ecosystem," *Proc. - 2016 IEEE/WIC/ACM Int. Conf. Web Intell. WI 2016*, pp. 453–456, 2017.

[29] J. J. Zubcoff *et al.*, "The university as an open data ecosystem," *Int. J.*

Des. Nat. Ecodynamics, vol. 11, no. 3, pp. 250–257, 2016.
[30] R. L. Grossman, "data lakes, commons and clouds : A review for analyzing and sharing genomic data." Arxiv, pp. 1–5, 2018.
[31] S. S. Dawes, L. Vidiasova, and O. Parkhimovich, "Planning and designing open government data programs: An ecosystem approach," *Gov. Inf. Q.*, vol. 33, no. 1, pp. 15–27, 2016.
[32] A. Zuiderwijk, M. Janssen, and C. Davis, "Innovation with open data: Essential elements of open data ecosystems," *Inf. Polity*, vol. 19, no. 1–2, pp. 17–33, 2014.
[33] C. W. Clegg, "Sociotechnical principles for system design," *Appl. Ergon.*, vol. 31, no. 5, pp. 463–477, 2000.
[34] C. Chyi Lee and J. Yang, "Knowledge value chain," *J. Manag. Dev.*, vol. 19, no. 9, pp. 783–794, Nov. 2000.
[35] M. I. S. Oliveira, G. de F. B. Lima, and B. F. Lóscio, "Investigations into Data Ecosystems: a systematic mapping study," *Knowl. Inf. Syst.*, pp. 1–42, 2019.
[36] A. Immonen, M. Palviainen, and E. Ovaska, "Requirements of an open data based business ecosystem," *IEEE Access*, vol. 2, pp. 88–103, 2014.
[37] G. Smith, H. A. Ofe, and J. Sandberg, "Digital service innovation from open data: exploring the value proposition of an open data marketplace," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2016-March, pp. 1277–1286, 2016.
[38] J. Lindman, T. Kinnari, and M. Rossi, "Business Roles in the Emerging Open-Data Ecosystem," *IEEE Softw.*, vol. 33, no. 5, pp. 54–59, 2016.
[39] R. M. Corrado Moiso, "Towards a User-Centric Personal Data Ecosystem," in *2012 16th International Conference on Intelligence in Next Generation Networks, ICIN 2012*.
[40] E. Haak, J. Ubacht, M. Van Den Homberg, S. Cunningham, and B. Van Den Walle, "A framework for strengthening data ecosystems to serve humanitarian purposes," *ACM Int. Conf. Proceeding Ser.*, 2018.
[41] S. Chattapadhyay, "Access and use of government data by research and advocacy organisations in India: A survey of (potential) open data ecosystem," *ACM Int. Conf. Proceeding Ser.*, vol. 2014-Janua, pp. 361–364, 2014.
[42] V. Köster and G. Suárez, "Open data for development: Experience of Uruguay," *ACM Int. Conf. Proceeding Ser.*, vol. 01-03-Marc, pp. 207–210, 2016.
[43] C. Moiso and R. Minerva, "Towards a user-centric personal data ecosystem the role of the bank of individuals' data," *2012 16th Int. Conf. Intell. Next Gener. Networks, ICIN 2012*, no. February 2018, pp. 202–209, 2012.
[44] M. Heimstädt, 'The Institutionalization of Digital Openness,' in *Proceedings of The International Symposium on Open Collaboration - OpenSym '14*, 2014, pp. 1–2.
[45] Kitchenham, "Guidelines for performing Systematic Literature Reviews in Software Engineering." Keele University, UK and University of Durham, UK, pp. 1–57, 2007.
[46] M. Van Den Homberg, J. Visser, and M. Van Der Veen, "Unpacking data preparedness from a humanitarian decision making perspective: Toward an assessment framework at subnational level," *Proc. Int. ISCRAM Conf.*, vol. 2017-May, no. May, pp. 2–13, 2017.
[47] N. Raymond, Z. Al Achkar, S. Verhulst, J. Berens, L. Barajas, and M. Easton, "Building data responsibility into humanitarian action," *OCHA Policy Stud. Ser.*, no. May, p. 18, 2016.
[48] M. I. S. Oliveira, L. E. R. A. Oliveira, M. G. R. Batista, and B. F. Lóscio, "Towards a meta-model for data ecosystems," pp. 1–10, 2018.
[49] B. Chae, "Big Data and IT-Enabled Services: Ecosystem and Coevolution," *IT Prof.*, vol. 17, no. 2, pp. 20–25, 2015.
[50] A. Zuiderwijk, M. Janssen, G. Van De Kaa, and K. Poulis, "The wicked problem of commercial value creation in open data ecosystems: Policy guidelines for governments," *Inf. Polity*, vol. 21, no. 3, pp. 223–236, 2016.
[51] T. M. Harrison, T. A. Pardo, and M. Cook, "Creating Open Government Ecosystems: A Research and Development Agenda," *Futur. Internet*, vol. 4, no. 4, pp. 900–928, 2012.
[52] A. Poikola, P. Kola, and K. A. Hintikka, *Public data-an introduction to opening information resources*. 2011.
[53] C. Moiso and R. Minerva, "Towards a user-centric personal data ecosystem the role of the bank of individuals' data," *2012 16th Int. Conf. Intell. Next Gener. Networks, ICIN 2012*, pp. 202–209, 2012.
[54] S. S. Dawes, L. Vidiasova, and O. Parkhimovich, "Planning and designing open government data programs: An ecosystem approach," *Gov. Inf. Q.*, vol. 33, no. 1, pp. 15–27, 2016.
[55] K. Smith *et al.*, "'Big metadata': The need for principled metadata management in big data ecosystems," *Proc. 3rd Work. Data Anal. Cloud, DanaC 2014 - Conjunction with ACM SIGMOD/PODS Conf.*, pp. 21–24, 2014.
[56] E. Przeybilovicz and M. A. Cunha, "Open government data programs," in *Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age - dgo '18*, 2018, pp. 1–2.
[57] NIST Big Data Public Working Group: Definitions and Taxonomies Subgroup, "NIST Big Data Interoperability Framework: Volume 2, Big Data Taxonomies," vol. 2. NIST, p. 31, 2015.
[58] R. G. Department, "Russian City "St. Petersburg" open government portal." Russian govt. department, pp. 1–10, 2019.
[59] EU, "EU Open Data Portal." EU, pp. 1–12, 2019.
[60] J. María and C. Edward, *New Horizons for a Data-Driven Economy*. Library of Congress, 2016.
[61] G. Magalhaes, C. Roseira, and L. Manley, "Business models for open government data," *ACM Int. Conf. Proceeding Ser.*, vol. 2014-Janua, pp. 365–370, 2014.
[62] H. Oh, S. Park, G. M. Lee, H. Heo, and J. K. Choi, "Personal Data Trading Scheme for Data Brokers in IoT Data Marketplaces," *IEEE Access*, vol. 7, pp. 40120–40132, 2019.
[63] B. Ubaldi, "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives." OECD, pp. 1–61, 2013.
[64] S. Buteau, P. Rao, A. K. Mehta, and V. Kadirvell, "Developing a framework to assess socioeconomic value of open data in India," *Proc. 14th Int. Symp. Open Collab. OpenSym 2018*, pp. 1–6, 2018.
[65] B. (Kevin) Chae, "A General framework for studying the evolution of the digital innovation ecosystem: The case of big data," *Int. J. Inf. Manage.*, vol. 45, no. January 2018, pp. 83–94, 2019.
[66] D. H. Shin and M. J. Choi, "Ecological views of big data: Perspectives and issues," *Telemat. Informatics*, vol. 32, no. 2, pp. 311–320, 2015.
[67] W. Street, "Building the (Open) Data Ecosystem the Present: A One-Way Street the Future: An Ecosystem Annexe: Some Illustrations of Our Current One," pp. 1–14, 2011.
[68] S. Orenga-Roglá and R. Chalmeta, "Framework for implementing a big data ecosystem in organizations," *Commun. ACM*, vol. 62, no. 1, pp. 58–65, 2019.
[69] M. Kaufmann, "Towards a Reference Model for Big Data Management," 2016.
[70] A. Ali, J. Qadir, R. ur Rasool, A. Sathiaseelan, A. Zwitter, and J. Crowcroft, "Big data for development: applications and techniques," *Big Data Anal.*, vol. 1, no. 1, 2016.
[71] H. D. A. do. Santos, M. I. S. Oliveira, G. de F. A. B. Lima, K. M. da Silva, R. I. V. C. Rayelle, and B. F. Lóscio, "Investigations into data published and consumed on the Web: a systematic mapping study," *J. Brazilian Comput. Soc.*, vol. 24, no. 1, 2018.
[72] D. S. R. Group, "Overview of the DDI Version 3.0 Conceptual Model." Structural Reform Group, 2004.
[73] Research data Management Team, "Data Life Cycle & Data Management Planning." University of Essex, UK, 2013.
[74] K. Möller, "Lifecycle models of data-centric systems and domains The abstract data lifecycle model," *Semant. Web*, vol. 4, pp. 67–88, 2013.
[75] K. Crowston and J. Qin, "A capability maturity model for scientific data management: Evidence from the literature," *Proc. ASIST Annu. Meet.*, vol. 48, no. 1, pp. 1–22, 2011.
[76] Y. Demchenko, C. De Laat, and P. Membrey, "Defining architecture components of the Big Data Ecosystem," in *2014 International Conference on Collaboration Technologies and Systems (CTS)*, 2014, pp. 104–112.
[77] L. Ding *et al.*, "TWC LOGD: A portal for linked open government data ecosystems," *J. Web Semant.*, vol. 9, no. 3, pp. 325–333, 2011.
[78] A. Immonen, E. Ovaska, and T. Paaso, "Towards certified open data in digital service ecosystems," *Softw. Qual. J.*, vol. 26, no. 4, pp. 1257–1297, 2018.
[79] P. O. T.D., "Public service Data Strategy 2019-2023." Government of Ireland, pp. 1–40, 2018.
[80] D. Koznov, O. Andreeva, U. Nikula, A. Maglyas, D. Muromtsev, and I. Radchenko, "A Survey of Open Government Data in Russian Federation Open Data Movements in Different," vol. 3, no. Ic3k, pp. 173–180, 2016.
[81] European Commission, "European Commission Digital Strategy." European Commission, pp. 1–33, 2018.
[82] S. Fan, R. Y. K. Lau, and J. L. Zhao, "Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix," *Big Data Res.*, vol. 2, no. 1, pp. 28–32, 2015.
[83] H. B. James E. Moore, "The Death of Competition -Leadership and

Strategy in the Age of Business Ecosystems." New York: Harper Business, pp. 1–6, 1996.
[84] M. Heimstädt, F. Saunderson, and T. Heath, "Conceptualizing Open Data Ecosystems: A timeline analysis of Open Data development in the UK," *Proc. Int. Conf. E-Democracy Open Gov.*, vol. 1000, no. October 2013, pp. 1–11, 2014.
[85] M. I. and R. Levien, *The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability*. Harvard Business School Press, 2004.