Data Driven Governments: Creating Value through Open Government Data

Judie Attard, Fabrizio Orlandi, Sören Auer

University of Bonn
attard@iai.uni-bonn.de, orlandi@iai.uni-bonn.de,
auer@cs.uni-bonn.de

Abstract. Governments are one of the largest producers and collectors of data in many different domains and one major aim of open government data initiatives is the release of social and commercial value. Hence, we here explore existing processes of value creation on government data. We identify the dimensions that impact, or are impacted by value creation, and distinguish between the different value creating roles and participating stakeholders. We propose the use of Linked Data as an approach to enhance the value creation process, and provide a Value Creation Assessment Framework to analyse the resulting impact. We also implement the assessment framework to evaluate two government data portals.

Keywords: government data, value creation, smart city, data value network, assessment framework

1 Introduction

Especially in recent years, open government initiatives have gone way beyond the simple publishing of data. In fact, the end aims of open data movements such as the Public Sector Information (PSI) Directive[1], U.S. President's Obama open data initiative[2], the Open Government Partnership[3], and the G8 Open Data Charter[4] focus on achieving *transparency*, *participatory governance*, and *releasing social and commercial value*. In order to have a well-functioning, democratic society, citizens and other stakeholders need to be able to monitor government initiatives and their legitimacy. Transparency means that stakeholders not only can access the data, but they also should be enabled to use, re-use and distribute it. The success to achieve transparency results in a considerable increase in citizen social control. Furthermore, through the publishing of government data, citizens are given the opportunity to actively participate in governance processes, such as decision-taking and policy-making, rather than sporadically voting in an election every number of years. Hence, through open government data initiatives

[1] http://ec.europa.eu/digital-agenda/en/european-legislation-reuse-public-sector-information
[2] http://www.whitehouse.gov/open/documents/open-government-directive
[3] http://www.opengovpartnership.org/
[4] https://www.gov.uk/government/publications/open-data-charter

such as portals, stakeholders can also be more informed and be able to make better decisions [32]. This opportunity can have a major impact within so many dimensions, including, but definitely not limited to; urban management, marketing, service improvement, and citizens' quality of life.

All data, whether addresses of schools, geospatial data, environmental data, weather data, transport and planning data, or budget data, has social and commercial value, and can be used for a number of purposes that could be different than the ones originally envisaged. Governments are one of the largest producers and collectors of data in many different domains [15]. Considering its volume (huge amount of data produced), velocity (frequent gathering of data, especially sensor data), variety (different domains), and veracity (uncertainty of data), government data can be considered to be Big Data. By publishing such data the government encourages stakeholders to innovate upon it, and create new services. The main challenge in releasing social and commercial value is that open data has no value in itself, yet it becomes valuable when it is used [16], and there are many factors within an open government initiative that influence its success.

In this paper we attempt to answer the following research question: *What are existing processes of value creation on open government data?* With this research question we aim to address the niche in existing literature with regards to the creation of value on open government data. In this paper we hence identify the various processes in a government data value chain, as well as dimensions that, in some way or another, have an impact on value creation upon government data. We also distinguish between the different value creating roles of participating stakeholders within the government data value chain, and identify the resulting impacts of value creation and of exploiting data as a product. While we focus on government data, it is important to note that most of what we discuss is also valid for generic open data initiatives.

2 Methodology

In order to analyse existing approaches undertaken for creating value based on open government data, we review existing literature on open government data initiatives. We implement a systematic approach, where we define a number of search terms and perform a search on a number of digital libraries. Thereafter, we select which literature to include in our study by applying inclusion and exclusion criteria.

The search terms we defined are a combination of the following keywords: *government, data, portal, open, publishing, consuming*, and *public*. The latter were selected with the aim of obtaining results relevant to the research question defined in Section 1, or more specifically; any initiative that exploits government data on order to create value. We here stick to the definition of *Government Data* to entail any data that is government-related. It may or may not be produced or published by a governmental entity, and it may or may not be made openly available (it can have varying degrees of openness).

In order to cover the largest spectrum of relevant publications possible, we identified and used the most extensively used electronic libraries, namely: ACM Digital Library, Science Direct, Springer Link, IEEE Xplore Digital Library, and ISI Web of Knowl-

edge. To achieve relevant results that are sufficiently comprehensive and precise, we apply these search terms on both the title and the abstract search fields.

After the systematic search was completed, we led a manual study selection based on exclusion and inclusion criteria. Basically, we only considered literature to be relevant if it regarded the actual exploitation of government data. This resulted in 74 publications that form our set of primary studies[5].

With the research question in mind, we analysed the 74 publications with the aim of identifying current practices of creating value using government data. We hence provide our observations, comments, guidelines, and conclusions in the rest of this paper. Of course, apart from the above-mentioned literature, we also conduct further lookup and exploratory searches [22] to identify related literature on which to base the contributions within this paper.

3 Background Literature

Data is increasingly becoming a commodity in our information society. It is steadily becoming the basis for many products and services, such as open data, Linked Data, or Big Data applications. Using open data, specifically open government data, has the potential of not only resulting in economic benefits, but also has good social and governmental impacts. Releasing government data will impact transparency and accountability factors, while the release of specific datasets can encourage stakeholders to create innovative services and boost economic growth. The release of information will also aid stakeholders in making informed decisions based on relevant data.

In order to reflect such a data-centric society, the concept of *value chains* [29] was coined to identify how value is created in order to achieve a product. The value chain model describes value-adding activities that connect an industry's supply side, such as raw materials and production processes, to its demand side, such as sales and marketing. The value chain model has been used to analyse and assess the linked activities carried out within traditional industries in order to identify where, within these activities, value is created. This was done with the aim to identify what activities are the source of competitive advantage within these industries.

As successful as the value chain concept was to achieve this aim, during these last years products and services are becoming increasingly digital, and exist in a more non-tangible dimension [27]. In addition, the traditional value chain model does not consider when information is used as a source of value in itself [30]. Thus, the original concept of value chain is becoming an inappropriate method with which to identify value sources in today's industries that produce non-tangible products [27]. Newer definitions of the concept, such as in [9,19,20,24,27], cater for the digital dimensions; taking into account factors and activities which set this dimension apart from the more physical one.

Lee and Yang [20] define a value chain for knowledge, including the knowledge infrastructure, the process of knowledge management, and the interaction between the required components. Knowledge, a step further than information, is data organised in meaningful patterns. The process of reading, understanding, interpreting, and applying information to a specific purpose, transforms information into knowledge. This

[5] All primary studies can be accessed here: `http://mnd.ly/1LFgFQJ`

means that for an entity that is unable to understand knowledge, the knowledge is in fact still only information. This is the *data literacy* problem, where any effort invested in knowledge generation is lost if the target consumer is unable to actually understand the provided knowledge [37]. Similar to Porter, Lee and Yang classify the activities within the knowledge value chain in five categories, namely knowledge acquisition, knowledge innovation, knowledge protection, knowledge integration, and knowledge dissemination.

In [9], Crié and Micheaux provide us with a more generic value chain than Lee and Yang, including raw data in their definition. Within their paper, the authors aim to highlight any issues within the value chain, to provide an overview of the current progress, and also to encourage entities to view the benefits of participating within the data value chain. They focus on four aspects of the *Data Value Chain*, namely:

- *Obtaining the right data* – Capturing the right data is the first step to forming an information chain that aims to provide the best customer service and result in profits;
- *Data quality management* – Ensuring the data is of good quality increases the potential towards maximising returns from the data for both the entity and its customers;
- *Deriving information and knowledge from raw data* – The act of extracting information from data, and interpreting knowledge from information;
- *Using information and knowledge to satisfy customers and generate profits* – The use of good data increases the chance of making better decisions.

Peppard and Rylander [27] also discuss a value chain that is more suited where the product in question is digitised, and thus non-tangible. The authors introduce the concept of *Network Value*, where value is created by a combination of actors within the network. In contrast to the earlier definition of a value chain, network value does not necessarily follow a linear model, and accounts for the various interconnected actors that work together to *co-produce* value. While these actors or entities should be able to function independently, they operate together in a framework of common principles. This means that an action by a single entity can influence other entities within the network, or otherwise require further actions from them in order to achieve the final product. Morgan et al. [26] provide a similar discussion on the co-production of value through open-source software.

In line with more recent popular themes, Miller and Mork [24] and Latif et al. [19] focus on big data and Linked Data respectively. Similar to Crié and Micheaux [9], Miller and Mork discuss the data value chain concerning all required actions in aggregating heterogeneous data in an organised manner and creating value (information/knowledge) that can influence decision-making. The authors divide their data value chain in three main categories, namely data discovery, integration, and exploitation. In contrast, Latif et al. propose the *Linked Data Value Chain*. Motivated by the still limited commercial adoption of the Semantic Web, the authors aim to drive the Semantic Web and the use of Linked Data closer to commercial entities. The authors discuss the entities participating in the Linked Data value chain, their assigned Linked Data roles, as well as the types of data processed within the chain. An interesting aspect that distinguishes the proposed Linked Data value chain from the ones previously mentioned is that actors within the

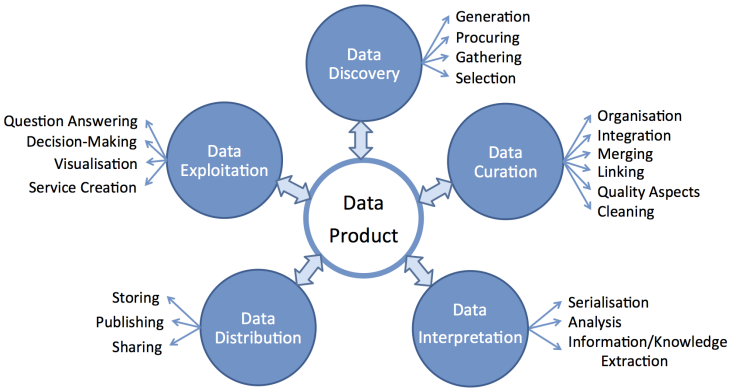

Fig. 1: The Data Value Network

chain are not necessarily bound to one specific role. Rather the assignment of roles to entities is more flexible where, in extreme cases, an entity can even occupy all roles at once.

Whichever model it follows, the data value chain is at the centre of a knowledge economy [1], where data products provide digital developments to more traditional sectors, such as transport, health, manufacturing, and retail. Being one of the largest producers and collectors of data in so many domains, governments play a vital role in data value chains. Essentially, the strategy beneath a data value chain is to extract the maximum value from data by building on the intelligent use of data sources [1]. The authors of [28] add that value-adding is one of the most important properties for information, where the objective of adding value to information is to develop information products and provide information service with social and economic value.

Based on the discussed literature, we provide our definition of a *Data Value Network*, as shown in Figure 1. This network differs from the classic definition of a data value chain in that the *activities* within the network do not follow a sequential structure; rather, activities can be executed in tandem, and other activities can be skipped or repeated. Furthermore, each activity can be further broken down into more specialised *value creation techniques* (the thin arrows in the Figure 1). While not exhaustive, the listed techniques are the most common and generic processes that can be executed on a data product. The Data Value Network also caters for multiple actors, where one or more actors can participate to co-produce value within an activity. We hence define the Data Value Network to be:

> *A set of independent activities having the aim of adding value to data in order to exploit it as a product*

where different **actors** can participate by executing one or more **activities**, and each activity can consist of a number of **value creation techniques**.

Data activities in a Data Value Network all have the purpose of adding value to data, which may or may not result in a new data product. We can consider ‘adding value’ to

be equivalent to 'making the data more usable, or making it more fit for use in a specific use case'. So, for example, while data in PDF format is easily human-readable, it's conversion to RDF would make it more usable where the use case requires data to be machine readable. The opposite can also stand true. We here provide a brief description of the activities that add value to data. The value creation techniques will be described in detail in Section 4.

- **Data Discovery**: Data discovery is the process of obtaining data. Sources of data can be as varied as sensor data, the Internet, private companies, governmental entities, and social media, amongst others.
- **Data Curation**: This is a very generic activity that can encompass a large number of different value creation techniques, all of which modify the data in some way or another. This activity can recur numerous times, until the required data product is obtained.
- **Data Interpretation**: This activity involves presenting the data in a different manner, in order for it to be more understandable.
- **Data Distribution**: Data distribution is the activity involving making the data available as a product. This means that other entities can search for and discover this data.
- **Data Exploitation**: This activity can be considered as the final goal of the Data Value Network, though it does not signify the Data Value Network is finished. It involves consuming the data as a product.

Within an urban environment such as a city, the Data Value Network can have major impacts on the citizens, especially where a data product is used in a decision-making process. This aspect is considered to be part of a smart city. The *decision-making process* is a very broad term used to encompass the practice of familiarising oneself with the relevant information before taking a particular decision. This concept was discussed as early as the 1970s, where Montgomery [25] describes the use of information systems to aid in the planning and decision-making processes within a marketing environment.

While there are various definitions of a smart city throughout literature, we consider a city to be smart *where investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance* [7]. There is increasing effort worldwide to transform cities into smart cities, particularly through the release of government data to the public, as well as through the exploitation of this data. Examples include Rio de Janeiro in Brazil[6], Dublin in Ireland[7], and London in the United Kingdom[8]. Whilst earlier attempts at "smartifying" a city mostly concerned the automation of some routine functions, more recent attempts are focusing on improving the management and sustainability of a city through monitoring and analysing the relevant data, with the intention of improving the quality of life if its citizens [3]. Smart cities can hence impact various dimensions in a citizen's life, for example:

[6] http://www.centrodeoperacoes.rio.gov.br/
[7] http://www.dublinked.ie/
[8] http://citydashboard.org/london/

– Transportation: The analysis of traffic data can aid citizens to check the best time to use certain roads, public transport can be better managed through better prediction of arrival times, whilst the government can attempt to lessen traffic by providing alternative transportation options. For example, a live view of the car boarding areas for the ferry between the islands of Gozo and Malta is streamed[9] in order to enable citizens to check if there is currently a long queue and plan their travels accordingly. Moreover, traffic supervisors can be dispatched to control and manage the boarding process.
– Energy Consumption: The use of smart meters and other sensors can help in reducing energy consumption through monitoring use in real-time. For example, an initiative throughout the European Union is currently ongoing with the aim of controlling energy consumption and providing for a more sustainable environment[10].
– Weather Emergencies: Weather information can be used to predict if a weather-related emergency is incumbent, such as flooding, landslides, earthquakes, etc. This prediction can be used to issue warnings or evacuation orders in time. The city of Rio de Janeiro is a good example of this use case, as an operations centre[11] was established with the aim to prevent weather-related disasters (amongst other aims).
– Health: Patient data can be used to generally monitor a patient during an ongoing treatment or to issue reminders when check ups or vaccinations are due. The Immunize India initiative[12] is an example of the latter.

4 Value Creation Techniques

Table 1 shows the various Value Creation Techniques within the Data Value Network. While not comprehensive, we included the most popular and frequently-used techniques from various stakeholders participating in the Data Value Network. The aim of all these techniques is to create or improve upon a data product, resulting in data that is (more) ideal to be used in the required application and increasing its value and re-use potential.

Data is produced in the day-to-day administration of a governing entity. The simple **generation** of this data is the first step towards its (re) use as a data product. As opposed to data generation, data **procurement** involves obtaining data generated by a different entity through performing some sort of negotiation. Data **gathering**, on the other hand, refers to the aggregation of data from different entities or locations. Finally, data **selection** requires the stakeholder in question to choose a subset of available data and extract it, potentially for the use in a different use-case then what the data was originally generated for. In order for the best value potential, all generated, procured, gathered, or selected data, need to be complete. This means a record has all the information required for an accurate representation of the described data.

[9] http://www.visitgozo.com/en/content/live-ferry-queue-streaming-beta-1538/
[10] http://my-smart-energy.eu/
[11] http://centrodeoperacoes.rio/
[12] http://www.immunizeindia.org/

Government Data Live Cycle Processes	Value Creation Techniques
Data Discovery	Generation Procuring Gathering Selection
Data Curation	Organisation Integration Merging Linking Quality Aspects
Data Interpretation	Serialisation Analysis Information/Knowledge Extraction
Data Distribution	Storing Publishing Sharing
Data Exploitation	Question Answering Decision-Making Visualisation Service Creation

Table 1: Value Creation Techniques categorised according to the Data Value Network

The value creation techniques falling under the Data Curation activity have the purpose of making the data more usable. Data **organisation** requires the structuring of data in such a way that the data is more understandable, or that the data follows some pattern; for example government budget data can be organised by year. Data **integration** has the purpose of enriching an existing dataset with new data, possibly with the intention to use the data in an unprecedented use. For example, the integration of weather data to accident information can be done by an insurance company to check the legitimacy of a claim. Another example is adding user feedback to product data in order to identify product faults. Data **merging** is somewhat similar, where different datasets are merged in order to obtain further information. For example, the merging of population data with geographical data can be used to obtain population density. On the other hand, the **linking** of different datasets is done in order to provide context, for example linking geographic data to textual descriptions about the locations in question. Finally, data **quality** involves the assessment and (if necessary) improvement and cleaning or repairing of data, such as removing duplicate data, ensuring the data is consistent, complete, timely, and trustworthy, and adding provenance data. This technique gives the data a higher level of quality and encourages its re-use. Similarly, metadata also enhances a datasets re-use potential. By enriching a dataset's metadata, a dataset is made more easily discoverable by potential users [31].

The Data Interpretation activity involves some sort of reasoning where the data in question is made more understandable. In the simplest way, data **serialisation** involves

the conversion of data into semantically richer or lower formats, such as PDF to RDF, or CSV to RDB. This conversion enables stakeholders with different backgrounds to still be able to exploit the data in question to its highest potential. Moreover, the use of non-proprietary, machine-readable formats will increase the value creation potential of the data in question. The implementation of **analysis** techniques, such as data mining, pattern identification, and trend analysis, enables stakeholders to identify any existing patterns, which can eventually aid actors in the Data Value Network in actions such as decision-making. **Information/knowledge extraction** has a similar purpose, where raw data is interpreted manually (non-machine), and along with the available context information and the knowledge from the stakeholders in question can be used to arrive to particular conclusions.

Techniques such as storing, publishing, and sharing, all have the purpose of adding the potential of the data to be distributed to different entities and re-used. The **storing** of data enables actors to re-use the data in question without requiring a local copy. By **publishing** the data in an open manner, and making it **shareable**, it is also made available to many more external stakeholders. This publishing process creates value simply by making data available for re-use. The data distribution activity is a vital node within the Data Value Network, as data that is not made available publicly is very limited in its re-use potential. Therefore, data that is provided in a timely manner (data is provided in a reasonable amount of time after creation/generation), without discrimination on its consumers (not requiring any registration), and made accessible for all, has the best value creation potential. Moreover, the addition of metadata enables the data to be more discoverable, thus enhancing this potential. Popular methods of publishing data include SPARQL[13] endpoints and Application Program Interfaces (APIs). Licensing is also vital here, as it has the purpose of declaring if and how data can be used. In the case of government data it is preferable that licences are of an open nature.

The Data Exploitation activity encompasses any value creation technique that involves consuming the data to solve a particular problem. **Visualisation** can be considered as an example of *passive* exploitation, where an actor consumes the data as information or knowledge. Visualisations involve a visual representation of data that, similar to data interlinking and data analysis, can provide us with a new insight. Visualisations can also be used to provide 'stories', since they are more easily interpreted than raw data. An example of a more *active* consumption of the data can be the use of data to influence **decision-making**, for example, a government might consider citizens' feedback before taking a decision. **Question answering** and **service creation** are other examples of active consumption of data. In the former data is collected and analysed in order to solve a specific question, whilst service creation is the provision of a service through the use of existing data, for example a mobile public transport timetable application.

4.1 Stakeholders: Beneficiaries, Contributors, and Their Roles

Government data, or public sector information, is a resource holding great potential for a large number of stakeholders. Governmental agencies, citizens, non-profit organ-

[13] `http://www.w3.org/TR/rdf-sparql-query/`

isations, and businesses, are but a few of the potential stakeholders who, through the exploitation of open government data, can reap substantial benefits. Since the efforts of the latter stakeholders remain largely uncoordinated, their motivations, levels of expertise, and priorities differ. In this section we proceed to identify and explore the various stakeholders who, either through value creation or other means of consumption, use open government data.

The most obvious role of **governments** in open government data initiatives is the role of a data provider. Yet, public entities are also the direct beneficiaries of their own published data. Through transparency as a motivation, the publishing of data can increase accountability, and moreover inhibits corruption. In turn this increases citizens' trust in their government. The analysis of government data, such as budget data, has the potential of increasing efficiency and influencing decision-making. Innovations based upon such data can also be used to provide more personalised public services, thus increasing the quality of the interactions between governments and their citizens.

Through publishing government data, **citizens** are given the possibility of participating in governance processes. Apart from being able to make more informed decisions, citizens are sometimes given the opportunity to take part in participatory governance. For example, in a participatory budget effort citizens are given a say as to how, or for what, budget should be prioritised. Citizens can also participate in open government initiatives by being data *prosumers*. By this we mean citizens who both produce and consume data. For example, the Fix My Street[14] application provides a platform where anyone can submit an existing problem in a street, in order to indicate the problem areas to the government. In this crowdsourced co-production of value, we have geographical data consumption, and street issues data production. Open government data certainly has the potential of increasing citizens' quality of life.

Non-profit organisations, such as non-governmental organisations (NGOs) or Civil Society initiatives, can have a huge difference in their goals. Examples of such organisations include the Sunlight Foundation[15] and the Open Knowledge Foundation[16], present in various countries. Organisations such as the latter usually share the goals of demonstrating the benefits of opening governmental data both to the general public and to the governments themselves. They also play a vital role as intermediaries who can identify key datasets that have the potential of being very valuable if published as open data.

Private companies, small to medium enterprises (SMEs), entrepreneurs, and other **businesses**, have the potential of not only making an economic profit through using government data, but can also create more jobs, and (depending on the nature of the service) also provide innovative services that increase the beneficiaries' quality of life and indirectly impact job creation in this field. While the sole access to data does not provide competitive advantage, private entities can innovate upon the available data to provide value-added services.

Whatever the stakeholder's nature (citizen, governmental entity, NGO, etc.), we identify six roles in which they can participate to create value, and in Table 2 we show how each role participates within the Data Value Network.

[14] `https://www.fixmystreet.com/`
[15] `http://sunlightfoundation.com/`
[16] `https://okfn.org/`

	Data Discovery	Data Curation	Data Interpretation	Data Distribution	Data Exploitation
Data Producer	✓				
Data Enhancer		✓	✓		
Data Publisher	✓			✓	
Service Creator		✓	✓		✓
Facilitator		✓	✓	✓	
Data Consumer			✓		✓

Table 2: The Activities in which each Actor participates within the Data Value Network

- **Data Producer:** A data producer is the entity that creates, obtains, or generates the data. The role of a data producer can be considered as one of the most important roles within the Data Value Network, as any activity or action in the network depends on the available data. If the data producer does not obtain relevant data for the use case at hand, then the Data Value Network will not reach its target to obtain the intended value out of this data.
- **Data Enhancer:** This role involves creating value through the actual manipulation of the data in a way that it is more usable for the target aim. A data enhancer can influence the outcome of the Data Value Network by adapting the data so that its highest value potential can be exploited.
- **Data Publisher:** This role involves the discovery and distribution of the data product. This distribution process enables other stakeholders to discover potentially useful data products.
- **Service Creator:** A service creator entity has the task of using open government data to provide a service. This can take the shape of a website, a mobile application, information access points, etc.
- **Facilitator:** This role involves entities that, in some way or another, aid the other stakeholders in using, re-using, or exploiting, open government data. This can be done through the provision of software, services, or other technologies. For example, the creator of a government data portal is facilitating the use and re-use of government data from other stakeholders by organising heterogeneous government data in a single location.
- **Data Consumer:** The data consumer role can be considered the final role in the Data Value Network, however, this is not always the case. For example, when a consumer gives feedback, the feedback can in turn be used as a data product by the product manufacturer. In the case of crowdsourcing, the data consumer also has the role of a curator, blurring the lines between both roles. Actors in the role of a data consumer can exploit the data product in many ways, as defined in the *Data Exploitation* activity.

4.2 Barriers, Enablers, and Impacts of Value Creation

Within the Data Value Network, value creation is both dependent on a number of dimensions, and also results in impact on other dimensions. Based on efforts in the primary

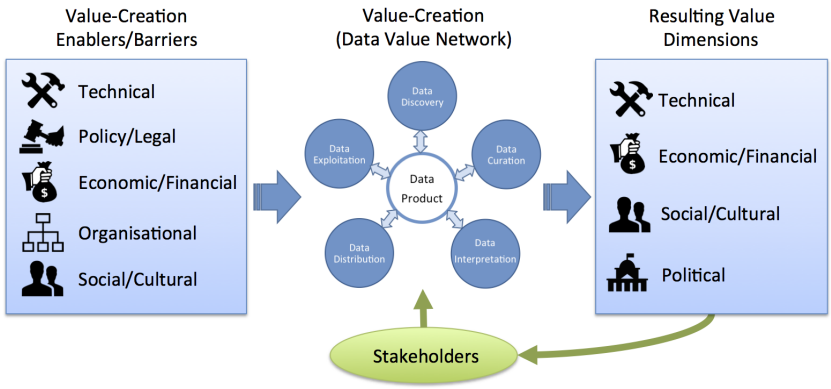

Fig. 2: Dimensions impacting, and impacted by, Value Creation

studies (See Section 2), and other literature such as [8, 16, 39, 41], we identify the dimensions with the strongest impact. Figure 2 maps their relationship, where a number of dimensions act as *enablers* or *barriers* towards value creation. In turn, the value creation process impacts a number of other dimensions. The stakeholders, while they give input for value creation, are also impacted through the results of their efforts.

4.2.1 Value Creation Enablers/Barriers: The latter dimensions have a great impact on value creation in that they control to what extent value is created.

The **Technical Dimension** mostly regards aspects concerning the data itself. The format of the data is an essential aspect. Two of the eight Open Government Data Principles[17], in fact, regard the format in which data is made available to the public. They state that such data should be available in a *machine-processable* format which is *non-proprietary*. Such data would enable easier and un-restricted use of the data for value creation. Furthermore, if a format such as Resource Description Framework (RDF) is used, data ambiguity is reduced due to the format's expressivity, making the data more *understandable*. Additionally, the use of common schema aids to reduce interoperability issues caused by the large heterogeneity of the existing data. In order to encourage its use, data must also be easily *discoverable*. This is possible through the use of good quality metadata. The implementation of agreed-upon standards would aid reduce some, if not most, of the issues within this dimension.

The **Policy/Legal Dimension** regards issues with existing laws or policies that, through their ambiguity or due to being out-of-date, prevent data from being used to create value. On the other hand, well thought out policies encourage and enforce the creation of value, for example the publishing of data as Linked Data. Fortunately, there are growing efforts towards amending such laws and policies, but there is still a long way to go. Copyright and licensing of data can inhibit its unrestricted use. The incompatibility of licences, due to the data being created by various entities, further aggravates the issue. Privacy and data protection is another important aspect. Data providers need

[17] `http://opengovdata.org/`

to strike a balance between making data freely available, whilst respecting the right to privacy.

The **Economic/Financial Dimension** is about aspects related to monetary issues and mainly concern the data provider and the data publisher roles. Being a relatively new concept, there might not be any budget allocation specifically for open government data efforts. In order to foster value creation, governmental entities cannot solely rely on existing data created in their day-to-day functionalities. Commitment is required, and hence also finances, for identifying and opening datasets with a high value creation potential.

The **Organisational Dimension** is concerned with the strategic aspects of the involved stakeholders. This dimension is especially relevant for governmental institutions. Considering there probably isn't an institution specifically in charge of open government data initiatives, data can get lost in the various hierarchical levels of a government. Adequate workflows need to be put in place for all the processes within a government data life cycle.

Finally, the **Social/Cultural Dimension** regards the feeling of the public towards open government data. While efforts are well under way to increasing awareness about the potential of open government data, not all stakeholders are ready to jump on the bandwagon. Workers within governmental entities might not understand the value of the data they are gathering/creating. This results in lack of motivation towards providing this data to the public. Stakeholders can also have misconceptions about the opening of public data. While open data can be considered as unfair competition for private entities (who invested to create their own data), public entities might consider the commercial appropriation of public open data unfair. The public also needs to be further informed on the advantages of public participation in creating value.

4.2.2 Impacts of Value Creation: As already discussed in the previous sections, value creation has a number of different dimensions of impact, which in turn affect the stakeholders. The term *public value* is used to define "what adds value to the public sphere" [4], where the public sphere is used to broadly indicate all of the following dimensions:

Technical Value is simply generated through the implementation of standards and the creation of services. As more value is created upon government data, the available data will be of better quality, and value creating services will increase.

Economic Value is defined as the worth of a good or service as determined by the market [17]. Value creation upon data enables the data itself to be considered as a product. Therefore, opening government data encourages its re-use in value creation, in turn stimulating competitiveness in the participating stakeholders and also encourages economic growth. For example, Mastodon C (a big data company) used open data to identify unnecessary spending in prescription medicine[18]. This will result in potentially huge savings from the National Health Service in the UK.

Social/Cultural Value is created first and foremost through the engagement of the public in open government data initiatives. The opening of data allows stakeholders to

[18] `http://theodi.org/news/prescription-savings-worth-millions-identified-odi-incubated-company`

scrutinise the data and provide feedback on it. If the governmental entities exploit this feedback, it can result in improvement of citizen services. This sort of participation also increases citizen social control. Social value is also generated through creating innovative services based on open government data. For example, the Walkonomics Application[19] uses open data to enable users to identify potential dangers in a street, such as fear of crime or road safety.

Political Value is created through the stimulation of democratic dialogue. Through participatory governance, citizens can gain a better insight as to how the governing process works. Stakeholders can possibly also participate in improving the policy-making process. Besides, the efforts of governmental entities to be more transparent and accountable increases citizens' trust in their government.

Through value creation, stakeholders are hence affected through all of the above dimensions. In line with the most relevant motivations behind open government data initiatives[20], namely transparency, releasing social and commercial value, and participatory governance, we identify four main levels of impact that are affected by the above dimensions and can be tangibly felt by the involved stakeholders.

1. Access to Information - Once data is re-used, the most directly tangible impact is access to information. The innovation and creation of services upon government data provides all stakeholders with more and more data and information that they can create value upon. In turn, the increase in availability of data products no only creates more jobs, but also affects the stakeholders' quality of life. This level of impact is directly affected through the Technical and Economic dimensions.
2. Transparency - By enabling stakeholders to create value upon government data, there can be a considerable increase in transparency. This is directly impacted by the Social/Cultural and Political Dimensions. Citizens are not only able to scrutinise data, but also create value upon it by providing relevant feedback. This sharing of responsibilities will allow them to interact with the government more actively, providing them with an opportunity to further exercise their duty and right of participation.
3. Accountability - Similarly to transparency, the creation of value on government data allows stakeholders to assess the legitimacy and effectiveness of the government's conduct. This helps citizens to establish a trusting relationship with the government. Affected by the Social/Cultural and Political dimensions, accountability enables citizens to be aware of how they are being governed, and have the relevant justifications.
4. Democratic Governance - Value creation on open government data not only promotes transparency and accountability, but also democracy. By participating in an open government initiative, stakeholders can provide feedback. The latter not only informs the governmental entity of the public opinion, but can also be used to improve service delivery. Affected by the Economic and Political dimensions, democratic governance essentially provides citizens with more social control.

[19] `http://www.walkonomics.com/`
[20] `http://opengovernmentdata.org/`

5 Linked Data

In recent open government data initiatives, Linked Data practices are being followed by an increasing number of data publishers/providers such as `data.gov.uk` and `data.gov`. Yet, the use of Linked Data in open government initiatives is still quite low [35]. This might be due to a number of reasons, as the use of Linked Data is a process involving a high number of steps, design decisions and technologies [40]. We here investigate the advantages and benefits of using Linked Data practices in an open government data initiative.

The term *Linked Data* is used to refer to a set of best practices for publishing and connecting structured data on the Web [5]. Therefore, Linked Data is published on the Web in a machine-readable format, where its meaning is explicitly defined. It is also linked to and from external datasets. This has the potential of creating the *Web of Data* (also known as Semantic Web); a huge distributed dataset that aims to replace decentralized and isolated data sources [13]. The benefits of applying Linked Data principles to government data as covered in literature include [10, 18]:

- Simpler data access through a unified data model;
- Rich representation of data enabling the documentation of data semantics;
- Re-use of existing vocabularies;
- Use of URIs allow fine-grained referencing of any information;
- Related information is linked, allowing its unified access.

While significant efforts in literature cover advantages of using Linked Data (for example [11,14,35,36]), there is no evident effort targeted towards the benefits of using Linked Data specifically in open government data value creation. We here therefore proceed to focus on the value creation techniques described in Section 4 and the benefits provided through the use of Linked Data. While still having similar barriers, enablers, and impacts, as described in Section 4.2, the use of Linked data can result in different levels of impact, since the use of Linked Data techniques directly reduces some barriers of the technical level.

5.1 Linked Data as a Basis for Value Creation

Linked Data and Semantic Web technologies have the potential of solving many challenges in open government data, as well as possibly lowering the cost and complexity of developing government data-based applications.

Starting from the most common starting point of creating value, in general, data **generation** is the least impacted from the use of Linked Data since essentially the data is still being created. Data **procurement** is similarly not impacted to a high level. Yet, the data **gathering** process can be enhanced through the use of Linked Data. Consider the example of providing feedback based on a linked open dataset consisting of budget data. The use of Linked Data enables feedback providers to have further context on the available data through the links. This would aid them in making a more informed decision. Furthermore, the high level of granularity of Linked Data has the potential of providing a deeper insight on the resource at hand. Also, since the data publisher is

not necessarily the data provider, Linked Data will enable the access to primary data through the use of provenance information located within the metadata. In the case of data **selection**, the use of Linked Data is particularly useful in querying for subsets of an existing dataset. Query languages such as SPARQL enable actors to generate complex queries and get very specific subsets of data.

The value creation techniques within the Data Curation activity are some of the highest impacted techniques within the Data Value Network through the use of Linked Data. Linked Data is based on models (schema) or ontologies that are best suited to represent the data at hand. In this way, the **organisation** of data is very easily achieved through the manipulation of the model at hand. If an entity is working with Linked Data, we can safely assume the data is represented in a semantically rich, machine-processable format. Hence, links with or between other datasets are more easily identified through the implemented models, and thus, the data **linking** process is simplified. Thereafter, data **integration** and **merging** follow easily through joining the existing models. Through the use of the standards required to obtain Linked Data, the *fitness for use* of data, and hence its **quality**, is immediately increased. For example, data ambiguity is decreased through the use of a semantically rich format, and data consistency can be ensured through the implemented data model. Moreover, in some instances, the quality assessment of data (and the ensuing data repairing/cleaning) can be more easily executed. For example, having a model for a linked dataset enables a stakeholder to assess the schema completeness for the dataset. Linked Data also enables (semi) automated cleaning and repairing of datasets through the use of reasoners. In this way, the violation of logical constraints is easily identified through the dataset's underlying model. Through the use of metadata, a consumer can also check the provenance of the data, and ensure that it is a reliable source. Timeliness and versioning information can be obtained in the same manner.

Having Linked Data means that the available data already conforms to some standards with regards to formatting, however this does not necessary make it easier to **serialise** to other formats. Yet, the use of agreed-upon standards positively affects the accessibility, discoverability, and re-usability potential of the data in question. Since Linked Data standards demand the use of a semantic representation such as RDF, Linked Data is automatically more accessible than other standards such as CSV or PDF. Data **analysis**, is also enhanced through the use of Linked Data. As explained above, Linked Data enables easier integration and merging of datasets, which in turn affect the implementation of analysis techniques. Moreover, through the existence of links it is easier to get further context and information on the data at hand, enhancing pattern identification. Similarly, the use of Linked Data in **information/knowledge extraction** also provides further insight and context to actors through links between the datasets, and within datasets themselves. This increased information directly affects the data interpretation process, as the data consumer can interpret the data in a more informed manner, and generate knowledge from the existing information.

The aim of the value creation techniques within the Data Distribution activity is to make the data more accessible as a data product. As mentioned above, the use of Linked Data standards automatically makes the data more accessible and discoverable. Hence, **stored** or **published** Linked Data has the potential to be easily accessed and

manipulated through a variety of manners, such as RESTful APIs and public endpoints (queryable through SPARQL). This means that while Linked Data alternatives might require a consumer to download a data dump, the use of Linked Data enables the same consumer to access the specific subset of data he/she needs, and manipulate it easily. Additionally, each data resource is dereferenceable, i.e. the resource URI can be resolved into a web document on the Web of Data. The **sharing** of data is also impacted through the use of Linked Data technologies, as the links in between different datasets make them more easily discovered through the crawling of web resources, which potentially could lead to the addition of the dataset to the more known LOD cloud[21].

Data Exploitation is possibly the activity that has the highest impact from the use of Linked Data. Similarly to the knowledge/information extraction process, **question answering** and **decision-making** are enhanced through the existence of links and the provision of further context. Hence a more informed stakeholder is more capable of making the best decision, or obtaining the best answer for the problem at hand. The creation of **visualisations** is also affected through the existence of links between multiple datasets. Visualising a dataset against a related dataset has the potential of providing the consumer with a new and different understanding of the data. Finally, **service creation** on top of Linked Data has the advantage of easier data consumption (through the use of standards), and more interoperability.

The above benefits of using data for value creation are only a few, yet they collectively encourage and enhance the exploitation of open (government) data. Of course, this does not mean the implementation of a Linked Data approach does not have its challenges. Various efforts in literature, such as [36], provide discussions on the topic.

5.2 Use Case of Linked Open Government Data

`publicspending.net` is a data portal created with the scope of demonstrating the power of economic Linked Open Data in analysing the situation with regards to market, competition conditions, and public policy, on a global scale. The creators of this portal consume and create value upon public spending data of seven governments around the world. Results of the analysis led on the data are then published on the portal as tables, graphs, and statistics. The stakeholders here participate through all six value-creating roles described in Section 4.1 and execute value creation processes accordingly. Firstly, the public spending data is produced by the various governments (Data Producers). The data is then subject to pre-processing and data-preparation. Through the role of a Data Enhancer, the stakeholders here homogenise and link the data through the Public Spending Ontology and other widely used vocabularies such as Dublin Core and FOAF. The resulting data in RDF is then published (Data Publisher) on the portal and is available both as bulk datasets and through a SPARQL endpoint. The Data Facilitator Role and the Service Creator Role are then fulfilled through the application built on top of the data. These stakeholders use the internal data, along with other cross-referenced and external data, to provide a portal acting as an information point. Finally, the Data Consumer can view and exploit the provided data in a myriad of ways, including exploring and scrutinising spending data that giving them a good insight as to what is being spent,

[21] `http://lod-cloud.net/`

where, and by whom. Such an open government data initiative enhances accountability and prevents corruption since it aids citizens to be more informed about how their country is being led, and if it is being led in a suitable manner. This can also help them decide who to vote for in an upcoming election.

6 Risks of Open Government Data

Whilst there are certainly numerous benefits and advantages of opening government data and creating value upon it, there still are a number of challenges that deter such initiatives from being successful and reaching their full potential, such as this discussed in Section 4.2. Moreover, if an open government data initiative is not implemented properly, the opening of data might also pose risks to some of the involved stakeholders. Within itself, this deters stakeholders from participating within an open government data initiative. We here proceed to outline some of the major risks of opening government data and creating value on it.

Conflicting regulations: Open government data initiatives have only become popular in recent years. Whilst there is certainly an increasing effort towards establishing policies, many open government data initiatives still belong to existing legal frameworks concerning freedom of information, re-use of public sector information, and the exchange of data between public entities. The risk here lies in the uncertainty of how such initiatives can interact. This issue concerns both data consumers, who are unsure how the available data can be used, and the data producers, who end up being sceptical of fully opening up their institutions' data, even if it is covered by a clear legal framework [33].

Privacy and Data Protection: Data protection and the right to privacy have some essential conflicts with the aims behind an open government data initiative and its motivations of transparency and accountability [23, 33, 41, 42]. Published data can certainly be anonymised, yet the merging or linking of different datasets can still possibly result in the discovery of data of a personal nature. For example, if garbage collecting routes are published, along with the personnel timetable, a data consumer would be able to identify the location of a particular employee. This issue requires more research in order to come up with guidelines that can provide a solution to this conflict, however a plausible approach would be to employ access control mechanisms which regulate data access. However, this restricts the openness level of such data.

Copyright and Licensing: The issue here lies with the incompatibility of used licences and copyright inconsistencies. Efforts in open government data initiatives strive towards publishing data in an open format, allowing the free and unrestricted use, reuse, and distribution of data. Since there are no agreed-upon standards, this results in a myriad of licenses that although all are of an open nature, they can be incompatible between them as they might contain restrictions that prevent data with different licences from being merged. Unclear dataset ownership resulting from data sharing, for example between different levels of public entities, results in copyright inconsistencies that hinders data from being published, as the rightful owner of the data is unclear [8, 42].

Competition: There are two perspectives to this risk: (i) open data can be considered as unfair competition for private entities, and (ii) public entities might consider

the commercial appropriation of public open data unfair [33]. In the first perspective consider business entities who invested in creating their own data stores. If the same data they created is made public through government open data initiatives, these companies will obviously deem it to be unfair competition as there is the possibility of new competitors who did not need to invest anything but could get the freely available open data. Thus, management mechanisms need to be applied in order to ensure that private companies do not suffer financial consequences due to opening up their data. On the other hand, public entities might be reluctant to publish their data openly due to not wanting data belonging to the public (and paid by taxes) to be used for commercial gain. A possible approach for the latter issue is to provide the data for a nominal fee. Yet, this limits the openness of the data in question.

Liability: Mainly, this risk is limited to data providers. The latter, in the context of this paper governmental entities, fear being held liable for damage caused by the use of the provided data due to it being stale, incorrect, or wrongly interpreted [12, 33]. To cater for this fear, many public entities either do not publish their data or otherwise impose restrictions on its use, resulting in data which is not truly open. In the worst case, due to fears of data being used against the publishing entity, such data might not even be collected/generated any longer [42]. A possible solution for these issues is to enable social interaction with regards to the data in question. A community of stakeholders within the data platform where the data is published can aid data consumers to better interpret and exploit the published data.

Considering the above risks or negative impacts, it is vital to find a trade-off for open government initiatives. One must keep in mind the numerous benefits associated with open data, but also cater and prepare for any risks, challenges and issues.

7 Value Creation Assessment Framework

In order to assess the success of open government data initiatives, there exist a large number of assessment frameworks that aim to evaluate the effectiveness of an initiative in achieving its goals and objectives. Yet, rather than assessing the resulting impacts of such an initiative, real-life assessments, as documented in literature (See Section 2), mostly involve checking whether open government data initiatives are obeying existing policies and regulations [34]. Since the latter are not necessarily up to date with current technologies and approaches, this assessment is not really representative of the success of an initiative.

Consider the example of a government publishing the data in PDF. While the entity would be obeying existing laws requiring opening up such data, the use of PDF makes it pretty inconvenient for re-use and re-distribution. In this case, one could argue that the open government initiative is not really a success. For this reason, a number of assessment frameworks analyse open government data initiatives based on different criteria [6, 21]. The latter include nature of the data, citizen participation, and data openness. In [2] we give a more in depth overview of existing assessment frameworks in literature. While there is still the problem that there is no agreed-upon assessment framework to evaluate open government initiatives, there is also limited literature (such as [38]) that focuses on the *impact of value creation*. Considering many resulting bene-

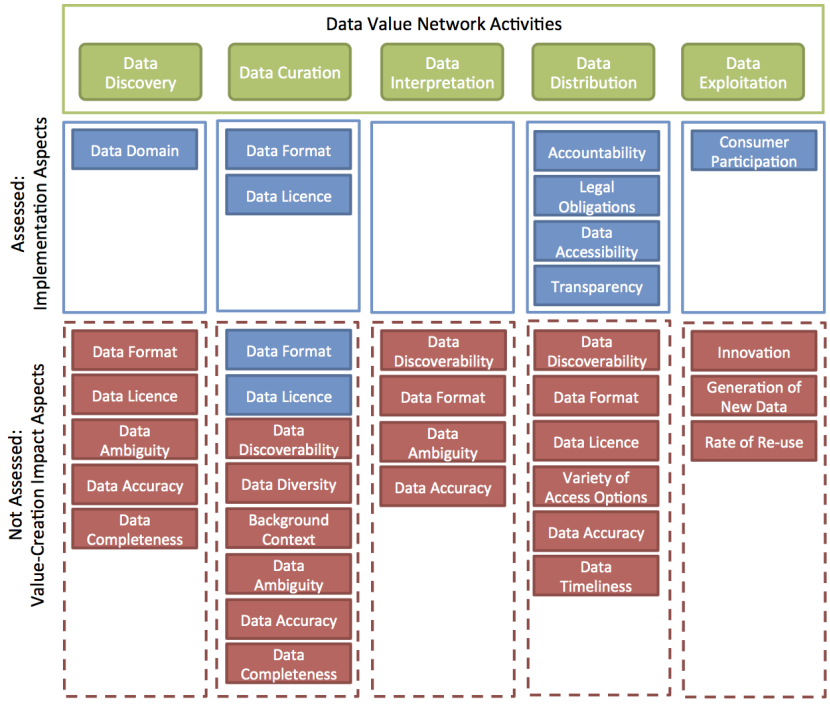

Fig. 3: Aspects assessed in existing frameworks (blue), aspects for Value Creation Assessment Framework (Red)

fits of open government data depend on the creation of value (through the execution of one or more value creation techniques), we deem it essential to assess open government data initiatives on their potential for enabling value creation.

In Figure 3 we provide an overview of commonly evaluated aspects (in blue) of an open government data initiative extracted from our primary studies. These mostly concern implementation aspects, such as the format of the data, and how the initiative respects the requirements set from existing laws and policies. The bottom part of the figure portrays the missing aspects (in red), i.e. those that are not considered when evaluating the success of an open government data initiative. We propose the latter aspects (together with a couple of aspects that are already being assessed) as part of a *Value Creation Assessment Framework*. The aim of this framework is to provide a guideline as to what aspects of an open government data initiative should be assessed to determine the potential of an open government data initiative to enable value creation, and thus exploit open government data to its highest potential. Here we briefly describe the aim of each aspect.

- *Data Format:* Formats such as CSV and RDF are much more usable then PDF. This is because they allow easier re-use of the represented data.
- *Data Licence:* Other than allowing for reasonable privacy, security, and privilege restrictions, data has the highest value creation potential if it is not subject to any lim-

itations on its use due to copyright, patent, trademark or other regulations. Hence, data with an open licence has the best value creation potential.
- *Data Ambiguity:* Data ambiguity is reduced when a representationally rich format (e.g. RDF) is used.
- *Data Accuracy:* The extent to which data accurately represents the respective information.
- *Data Completeness:* Data is complete when all required information is available, for the representation of the data in question.
- *Data Discoverability:* This aspect depends on the metadata annotating the data in question, and enables stakeholders to more easily find data that is relevant to their needs. Data Discoverability is also affected by the search functions provided by a government portal or catalogue.
- *Data Diversity:* In the Linking value-creation process, the use of diverse datasets has the potential of releasing new insights or unforeseen results.
- *Background Context:* The linking of datasets provides further context to the data in question, enabling stakeholders to have a deeper understanding.
- *Use of Standards:* Using agreed-upon standards throughout the life-cycle of government data encourages data re-use and integration.
- *Variety of Access Options:* Providing various access options to the available data, such as APIs and SPARQL endpoints, encourages stakeholders to create value upon the data as they are able to access the data in their preferred manner.
- *Data Timeliness:* Certain data might only be valuable if it is made openly available shortly after its creation.
- *Innovation:* Creating new products (data or otherwise) based on open government data is a direct impact of value-creation. Innovations include services and applications.
- *Generation of New Data:* The value-creation techniques in the Data Exploitation Process can result in the generation of new data, such as visualisations, that provide new interpretations or insight on the existing government data.
- *Rate of Re-use:* The participation of stakeholders in consuming the data is essential for value-creation. There is no use in having data made openly available if it is not exploited. The rate of re-use of open government data is directly indicative of the value-creation potential in the assessed initiative.

Since one of the major aims of open government initiatives is the release of social and commercial value, we deem that the proposed aspects are vital to determine the success of an initiative. Hence, these *value creation impact aspects* are used to assess the potential value that can be created through the use of the data product created as a result of each step within the Data Value Network.

7.1 Value Creation Assessment Framework in Action

In this section we implement the proposed assessment framework on two open government data initiatives, namely `www.govdata.de` and `www.gov.mt`, in order to portray its relevance and applicability in the context of value creation on open government data. Keeping in mind that this implementation is acting as a proof of concept, we

Value-Creation Impact Aspects	Assessment Metrics	Results govdata.de	Results gov.mt
Data Format	5 star scheme for LOD: 1-5 marks according to format	2.71 out of 5	2.39 out of 5
Data Licence	0 marks if no licence specified, 1 mark if licence has some restrictions, 2 marks if open and enabling re-use	1.85 out of 2	0 out of 2
Data Ambiguity	1 mark if using semantically rich formats (e.g. RDF)	0 out of 1	0 out of 1
Data Accuracy	Requires use of a gold standard[a]	-	-
Data Completeness	Requires use of a gold standard[a]	-	-
Data Discoverability	1 mark if metadata is available, 1 mark if portal offers search functions on the data (2 marks max)	2 out of 2	0 out of 2
Data Diversity	1 mark if there is more than one dataset on a specific domain	1 out of 1	1 out of 1
Background Context	1 mark if datasets are linked to other external datsets	0 out of 1	0 out of 1
Variety of Access Options	1 mark if more than one access option is available	1 out of 1	0 out of 1
Data Timliness	1 mark if data has a timestamp, 1 mark if recently updated data is available (2 marks max)	2 out of 2	0 out of 2
Innovation	1 mark if portal provides innovations based on published data, 2 marks if different innovations are provided (e.g. services, applications) (3 marks max)	3 out of 3	0 out of 3
Generation of New Data	1 mark if portal enables users to generate new data (e.g. visualisations)	0 out of 1	0 out of 1
Rate of Re-Use	1 mark if portal provides links and information on re-use of the published data	0 out of 1	0 out of 1
	Total	13.56 out of 20	3.39 out of 20

Table 3: Value-Creation Assessment Framework Metrics and Results

[a] This aspect cannot be assessed on a high level as it requires the use of an algorithm that analyses each dataset in a portal and compares it to a gold standard.

restrain our metrics to assess the portal on a high level, as we consider a through and more accurate implementation to require significant more research. We therefore base the provided metrics on ground research. In Table 3 we provide a description of the metrics used, and the results of the portals[22]. We assign marks according to the assessed aspect, and where relevant we average the marks out based on the number of available datasets. For example, to assess the data format of eight datasets, if four datasets are in RDF and linked to other datasets (4 x 5 marks) and four datasets are in CSV (4 x 2 marks), then the result for the data format aspect is 3.5 marks.

Having a value-creation potential of 13.56 marks out of 20, `www.govdata.de` can do with some improvements, especially with regards to the use of RDF and the linking to other documents. The portal could also benefit from enabling users to both create new innovations or data through the portal itself, and also from providing some sort of documentation to both portray any innovations based on the data in question. In summary, `www.govdata.de` is on the right track towards the opening of governmental data, however it definitely requires more effort towards encouraging stakeholders to create value upon the published data.

On the other hand, `www.gov.mt` does not really excel in publishing government data. Apart from providing very few datasets, some require logging in with a government-issued e-id to download, and others are not even available (404 error given). Moreover, no search functions are provided to aid a user search within the provided datasets, such as a faceted browser. Whilst there is a statement encouraging stakeholders to innovate upon the data, no actual data licence is provided, leading room towards uncertainty.

8 Concluding Remarks

The main challenge in public value is that open data has no value in itself, yet it becomes valuable when it is used. In our information society, value creation processes have the potential of extracting the maximum value from data by building on its intelligent use. All stakeholders of value creation can participate through different roles, yet they have one common goal; that of creating a data product. Different dimensions impact the creation of such a product, namely technical, policy/legal, economic/financial, organisational, and cultural. Some of these dimensions are in turn also impacted by value creation. The use of Linked Data in creating value enhances the process, and also aids us to gradually proceed through various degrees of data products: starting with data, to information, and ultimately to knowledge. In order to truly assess the value creation process of an open government initiative, we propose an assessment framework that focuses on the potential impact achievable from a data product generated through a value creating process, and implement it on a high level on two government data portals. As future work we intend to further explore more accurate metrics that can be used to assess the suggested aspects within the framework. Step by step the vision of having open government data exploited to its full potential can be acquired.

[22] As per 29th of December 2015

References

1. Elements of a data value chain strategy | Digital Agenda for Europe | European Commission, https://ec.europa.eu/digital-agenda/en/news/elements-data-value-chain-strategy
2. Attard, J., Orlandi, F., Scerri, S., Auer, S.: A systematic review of open government data initiatives. Government Information Quarterly 32(4), 399 – 418 (2015), http://www.sciencedirect.com/science/article/pii/S0740624X1500091X
3. Batty, M., Axhausen, K.W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., Portugali, Y.: Smart cities of the future. The European Physical Journal Special Topics 214(1), 481–518 (2012)
4. Benington, J.: From private choice to public value. Public value: Theory and practice pp. 31–49 (2011)
5. Bizer, C., Heath, T., Berners-Lee, T.: Linked Data - The Story So Far. Int. J. Semantic Web Inf. Syst. 5(3), 1–22 (2009)
6. Bogdanović-Dinić, S., Veljković, N., Stoimenov, L.: How Open Are Public Government Data? An Assessment of Seven Open Data Portals. In: Measuring E-government Efficiency, Public Administration and Information Technology, vol. 5, pp. 25–44. Springer New York (2014)
7. Caragliu, A., Del Bo, C., Nijkamp, P.: Smart Cities in Europe. Journal of Urban Technology 18(2), 65–82 (2011)
8. Conradie, P., Choenni, S.: Exploring Process Barriers to Release Public Sector Information in Local Government. In: Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance. pp. 5–13. ICEGOV '12, ACM, New York, NY, USA (2012)
9. Crié, D., Micheaux, A.: From customer data to value: What is lacking in the information chain? Journal of Database Marketing & Customer Strategy Management 13(4), 282–299 (jul 2006)
10. Cyganiak, R., Maali, F., Peristeras, V.: Self-service linked government data with dcat and gridworks. In: Proceedings of the 6th International Conference on Semantic Systems - I-SEMANTICS '10. p. 1. ACM Press, New York, New York, USA (sep 2010)
11. DiFranzo, D., Graves, A., Erickson, J.S., Ding, L., Michaelis, J., Lebo, T., Patton, E., Williams, G.T., Li, X., Zheng, J.G., Others: The web is my back-end: Creating mashups with linked open government data. Linking Government Data pp. 205–219 (2011)
12. Eckartz, S., Hofman, W., Van Veenstra, A.: A decision model for data sharing. In: Janssen, M., Scholl, H., Wimmer, M., Bannister, F. (eds.) Electronic Government, Lecture Notes in Computer Science, vol. 8653, pp. 253–264. Springer Berlin Heidelberg (2014)
13. Heath, T.: How Will We Interact with the Web of Data? IEEE Internet Computing 12(5), 88–91 (sep 2008)
14. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web, Morgan {&} Claypool Publishers (2011)
15. Janssen, K.: The influence of the PSI directive on open government data: An overview of recent developments. Government Information Quarterly 28(4), 446–456 (oct 2011)
16. Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management 29(4), 258–268 (2012)
17. Jetzek, T., Avital, M., Bjørn-Andersen, N.: Generating Value from Open Government Data. In: Baskerville, R., Chau, M. (eds.) Proceedings of the International Conference on Information Systems, {ICIS} 2013, Milano, Italy, December 15-18, 2013. Association for Information Systems (2013)

18. Kalampokis, E., Tambouris, E., Tarabanis, K.: A Classification Scheme for Open Government Data: Towards Linking Decentralised Data. Int. J. Web Eng. Technol. 6(3), 266–285 (jun 2011)
19. Latif, A., Us Saeed, A., Hoefler, P., Stocker, A., Wagner, C.: The Linked Data Value Chain: A Lightweight Model for Business Engineers. In: Proceedings of International Conference on Semantic Systems. pp. 568–576 (2009)
20. Lee, C.C., Yang, J.: Knowledge value chain. Journal of Management Development 19(9), 783–794 (2000)
21. Lourenco, R.P.: Open Government Portals Assessment: A Transparency for Accountability Perspective. In: Wimmer, M., Janssen, M., Scholl, H.J. (eds.) EGOV. Lecture Notes in Computer Science, vol. 8074, pp. 62–74. Springer (2013)
22. Marchionini, G.: Exploratory search. Communications of the ACM 49(4), 41 (apr 2006)
23. Meijer, R., Conradie, P., Choenni, S.: Reconciling Contradictions of Open Data Regarding Transparency, Privacy, Security and Trust. Journal of theoretical and applied electronic commerce research 9, 32 – 44 (09 2014), http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-18762014000300004&nrm=iso
24. Miller, H.G., Mork, P.: From Data to Decisions: A Value Chain for Big Data. IT Professional 15(1), 57–59 (Jan 2013)
25. Montgomery, D.B., UrbanL., G.: Marketing Decision-Information Systems: An Emerging View. Journal of Marketing Research 7(2) (1970)
26. Morgan, L., Feller, J., Finnegan, P.: Exploring value networks: theorising the creation and capture of value with open source software. EJIS 22(5), 569–588 (2013)
27. Peppard, J., Rylander, A.: From Value Chain to Value Network:. European Management Journal 24(2-3), 128–141 (Apr 2006)
28. Ping Wang, Haiying Hua: A model of government information value-added exploitation based on cloud computing. In: 2011 International Conference on Business Management and Electronic Information. vol. 2, pp. 518–522. IEEE (may 2011)
29. Porter, M.E.: Competitive Advantage:Creating and sustaining superior performance, vol. 15 (1985)
30. Rayport, J.F., Sviokla, J.J.: Exploiting the Virtual Value Chain. Harvard Business Review 73, 75 (1995)
31. Reiche, K.J., Höfig, E.: Implementation of Metadata Quality Metrics and Application on Public Government Data. In: COMPSAC Workshops. pp. 236–241 (2013)
32. Rojas, L., Bermudez, G., Lovelle, J.: Open Data and Big Data: A Perspective from Colombia. In: Knowledge Management in Organizations, Lecture Notes in Business Information Processing, vol. 185, pp. 35–41. Springer International Publishing (2014)
33. Dulong de Rosnay, M., Janssen, K.: Legal and Institutional Challenges for Opening Data across Public Sectors: Towards Common Policy Solutions. Journal of theoretical and applied electronic commerce research 9, 1 – 14 (09 2014)
34. Sandoval-Almazan, R., Gil-Garcia, J.: Towards an Evaluation Model for Open Government: A Preliminary Proposal. In: Janssen, M., Scholl, H., Wimmer, M., Bannister, F. (eds.) Electronic Government, Lecture Notes in Computer Science, vol. 8653, pp. 47–58. Springer Berlin Heidelberg (2014)
35. Shadbolt, N., O'Hara, K.: Linked Data in Government. {IEEE} Internet Computing 17(4), 72–77 (2013)
36. Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., Hall, W., Schraefel, M.: Linked open government data: lessons from Data.gov.uk (may 2012)
37. Shah, S., Horne, A., Capella, J.: Good Data Won't Guarantee Good Decisions - Harvard Business Review. Harvard Business Review [Internet] (April) (2012)

38. Susha, I., Zuiderwijk, a., Janssen, M., Gronlund, a.: Benchmarks for Evaluating the Progress of Open Data Adoption: Usage, Limitations, and Lessons Learned. Social Science Computer Review pp. 1–18 (2014)
39. Ubaldi, B.: Open Government Data (may 2013)
40. Villazón-Terrazas, B., Vilches, L., Corcho, O., Gómez-Pérez, A.: Methodological Guidelines for Publishing Government Linked Data. In: Wood, D. (ed.) Linking Government Data, chap. 2. Springer (2011)
41. Zuiderwijk, A., Janssen, M.: Barriers and Development Directions for the Publication and Usage of Open Data: A Socio-Technical View. In: Open Government, Public Administration and Information Technology, vol. 4, pp. 115–135. Springer New York (2014)
42. Zuiderwijk, A., Janssen, M.: The negative effects of open government data - investigating the dark side of open data. In: Proceedings of the 15th Annual International Conference on Digital Government Research. pp. 147–152. dg.o '14, ACM, New York, NY, USA (2014)