



D2.4 OPEN DATA STAKEHOLDER REQUIREMENT REPORT 1

PROJECT

Acronym: **OpenDataMonitor**

Title: Monitoring, Analysis and Visualisation of Open Data Catalogues, Hubs and Repositories

Coordinator: SYNYO GmbH

Reference: **611988**

Type: Collaborative project

Programme: FP7-ICT

Start: November 2013

Duration: 24 months

Website: <http://project.opendatamonitor.eu>

E-Mail: office@opendatamonitor.eu

Consortium: **SYNYO GmbH**, Research & Development Department, Austria, (SYNYO)

Open Data Institute, Research Department, UK, (ODI)

Athena Research and Innovation Center, IMIS, Greece, (ATHENA)

University of Southampton, Web and Internet Science Group, UK, (SOTON)

Potsdam eGovernment Competence Center, Research Department, Germany, (IFG.CC)

City of Munich, Department of Labor and Economic Development, Germany, (MUNICH)

Entidad Publica Empresarial Red.es, Shared Service Department, Spain, (RED.ES)

DELIVERABLE

Number:	D2.4
Title:	Open Data Stakeholder Requirement Report 1
Lead beneficiary:	IfG.CC
Work package:	WP2: Research Studies and Stakeholder Analysis
Dissemination level:	Public (PU)
Nature:	Report (R)
Due date:	September 30, 2014
Submission date:	September 26, 2014
Authors:	Sirko Hunnius , IfG.CC Bernhard Krieger , IfG.CC
Contributors:	Nacho Sanchez , RED.ES Sonia Castro , RED.ES Wolfgang Glock , MUNICH
Reviewers:	Bernhard Jäger , SYNYO Ejona Sauli , SYNYO Michael Heil , SYNYO

Acknowledgement: The OpenDataMonitor project is co-funded by the European Commission under the Seventh Framework Programme (FP7 2007-2013) under grant agreement number: 611988.

Disclaimer: The content of this publication is the sole responsibility of the authors, and in no way represents the view of the European Commission or its services.

TABLE OF CONTENTS

1	Introduction	6
2	Stakeholders in Open Data	8
2.1	Stakeholder Roles in the Open Data Process	8
2.2	Stakeholder barriers in open data.....	11
2.3	Preliminary ODM Stakeholder Classification.....	14
3	Methodology	15
4	Strategic Processes Regarding Open Data	16
4.1	Policy-Making around Open Data	16
4.2	Decisions about Platforms, Format Standards, Meta Data and Vocabulary.....	21
4.3	Decisions about Publishing Data Sets.....	27
5	Processes of Open Data Production	30
5.1	Data Generation, Storage and Management	30
5.2	Data Extraction, Preparation and Publication	31
5.3	Feedback Processing and Analytics of Open Data Usage.....	33
6	Processes of Open Data Use and Consumption	35
6.1	Data Detection	35
6.2	Data Retrieval and Use	36
7	Stakeholder Requirements	38
8	Consequences for ODM.....	43
9	References.....	45



LIST OF ABBREVIATIONS

AENOR	Spanish Association for Standardisation and Certification
API	Application Programming Interface
CIO	Chief Information Officer
EC	European Commission
EU	European Union
FOI	Freedom of Information
ICT	Information and Communications Technologies
MINETUR	Ministry of Industry, Energy and Tourism
MINHAP	Ministry of Finance and Public Administration
NEGS	National E-Government Strategy
NIF	National Interoperability Framework
OGD	Open Government Data
OGP	Open Government Partnership
OKF	Open Knowledge (Foundation)
ODI	Open Data Institute
ODM	OpenDataMonitor
PSI	Public Sector Information
SME	Small and Medium-sized Enterprises
STRAC	IT-Strategie und IT-Steuerung/IT-Controlling of the city of Munich
USP	Unique Selling Proposition
W3C	World Wide Web Consortium

LIST OF TABLES

Table 2.1 Barriers of Open Data; reshuffled, complemented and condensed list from Zuiderwijk, Janssen et al. (2012)	11
Table 3.1 Composition of Interview Participants	15
Table 4.1 Stakeholders and their exemplary interests in policy-making around open data	21
Table 4.2: Stakeholders and their exemplary interests in decisions about open data endeavours	26
Table 4.3: Stakeholders and their exemplary interests in decisions about publishing data sets	29
Table 5.1: Stakeholders and their exemplary interests in data generation, storage and management	31
Table 5.2: Stakeholders and their exemplary interests in data extraction, preparation and publication	33
Table 5.3: Stakeholders and their exemplary interests in feedback processing and usage analytics ..	34
Table 6.1: Stakeholders and their exemplary interests in data detection	36
Table 6.2: Stakeholders and their exemplary interests in data retrieval and use	37
Table 7.1: Stakeholder Requirements and ODM Potential	38

1 INTRODUCTION

The topic of open data is generating considerable interest among researchers, technology developers and practitioners in public administration as well as in the private sector. However, especially the public sector is pressured to make its data openly available. Here, the topic is not entirely new, as the re-use of public sector information (PSI) has been the subject of longer debates and EU directives (K. Janssen & Dumortier, 2003) which in some respect have prepared the ground for open data (K. Janssen, 2011). Also, debates about freedom of information and transparency show some connections to open data (Owen, Cooke, & Matthews, 2013) and illustrate its particular value laden nature in the public sector (Cerrillo-i-Martínez, 2012). Beyond the more philosophic argument that the data has already been paid for by the public, this draws attention to the impact of transparency on e.g. trust in government (Bannister & Connolly, 2011), social inclusion (Gurstein, 2014) and accountability (Yu & Robinson, 2012). Thus, the debate about open data is often reduced to open government data. The possibility of open data supplied by private companies receives only scant attention (Deloitte, 2012; Immonen, Palviainen, & Ovaska, 2014).

Furthermore, the conversation so far often circles around the potentials of open data (Geiger & von Lucke, 2012; McKinsey Global Institute, 2013; Shadbolt & O'Hara, 2013). However, up until now there is little evidence of any significant economic or societal impact (Huijboom & van den Broek, 2011). Regarding the sheer amount of available open data Anglo-American governments seem to provide far more data and render the latter in a more sophisticated way than governments in continental European countries (Davies, 2013; Open Knowledge Foundation, 2013). Even though publishing the data is not an end in itself, the different extent of available data is remarkable in its own right. This difference is puzzling as open data seems to be an international trend, fostered by an international community and pushed by international advocacy groups (e.g. Open Knowledge Foundation). Multinational initiatives such as the Open Government Partnership are taken up by countries as diverse as the United States of America, Chile, Austria, Russia, Kenya and Malaysia. Considering this heterogeneity of actors on the one hand and the differences in the implementation of open data activities on the other institutional factors appear to play an important role in how open data is perceived and adopted in the different administrative traditions and public sector organisations (Davies & Bawa, 2012). A more thorough understanding of the adoption process seems necessary to learn why open data catches on faster and differently in some instances and what the various impacts are.

Thus far, how open data is adopted by countries and organisations remains scarcely understood. This is because to a large extent, research has focussed on the operative routine processes around open data. Models around open data exclusively take the operational day-to-day processes into account (such as extracting, cleaning, publishing and maintaining data), while at the same time neglecting the

strategic processes (such as policy production, decision making and administrative enforcement). However, it appears evident that these latter processes play an influential role in the shaping of open data (Courmont, 2012; Heimstädt, Saunderson, & Heath, 2014; Hunnius & Krieger, 2014).

More comprehensive conceptualisations of the “system of people, practices, values, and technologies” (Nardi & O’Day, 1999, p. 49) around open data have gained some recognition recently as open data ecosystems (Harrison, Pardo, & Cook, 2012; Heimstädt et al., 2014). The (information) ecology metaphor points to the symbiotic relationships between actors and how they are embedded in a specific ecology. Looking at the current technical open data landscape with its diverse sources of data, scattered hubs, various formats for data and meta data not to speak of data structures and vocabularies it becomes evident that the landscape is still fractured (Mayer-Schönberger & Zappia, 2011). Actors have largely withstood initiatives which aimed for consolidation or at least systematisation. Therefore, this report aims to understand why this is the case and what could level the landscape.

This report attempts to paint a comprehensive picture of stakeholders in open data – what exactly their stake is, which role(s) they have in the ecosystem, what their interests are and what requirements need to be met so they can fulfil their function in the ecosystem. The report is structured as follows: After describing the methodology used for this research the literature about open data is reviewed to get an initial understanding of stakeholders in open data. Thereupon, we suggest based on the empirical data, a more nuanced and detailed understanding of processes around open data. Looking at the roles the stakeholders have in these processes, we delineate their requirements. Finally, we contextualise how their requirements can be met by the ODM project.

The ODM project will build a platform that automatically monitors open data in Europe across the various open data catalogues and hubs. The goal is to increase transparency not only of the amount of available open data, but also of the spread of certain artefacts (licenses, data formats, meta data etc.). This enables stakeholders to get a more detailed understanding of what is available and can be used as well as what others make available. To satisfy the various interests of the stakeholders in open data it is necessary to understand their role, their goals and their interests in order to thereupon deduce how ODM could serve them. Within the ODM project this impacts on the design of the platform, the functionalities it will offer and the available analytics.

2 STAKEHOLDERS IN OPEN DATA

The term stakeholder is fairly often used in the context of open data, however, often by name only, without explicitly drawing on the stakeholder approach. The stakeholder approach was developed in business sciences as a means to analyse how groups or individuals with vested interests in a firm are or should be considered by its management. The reasons why a company might be interested in such analysis are very diverse, from moral (normative) to efficiency-oriented (instrumental) reasons (Jones & Wicks, 1999). Despite its initial focus on private entities, the approach has been applied to public sector settings (Tennert & Schroeder, 1999), in particular to study technology adoption in this space (Flak & Rose, 2005; see Hans J Scholl, 2001; Tan, Pan, & Lim, 2005), mostly with positive results. Notwithstanding criticism of such an expansion (Donaldson & Preston, 1995), it can be argued that due to outward accountabilities (see e.g., Romzek, 2000), largely constitutive externalities (see e.g., Batley, 1994; Haque, 2001) and the network-type interdependencies of public sector management (see e.g., Kickert, Klijn, & Koppenjan, 1997; O’Toole, 1997), the stakeholder approach seems suitable and beneficial for public sector management (Hans J Scholl, 2001; Tennert & Schroeder, 1999). The stakeholder approach seems especially promising for the context of open data, since it provides an analytical frame that is not restricted to intra-organisational actors, but offers considerable conceptual breadth (Phillips, Freeman, & Wicks, 2003). On the downside, it does not provide a clear cut in-out-distinction who is a stakeholder and who is not (Mitchell, Agle, & Wood, 1997). To mitigate this imprecision, we will resort to the processes in and around open data identify relevant stakeholders.

2.1 Stakeholder Roles in the Open Data Process

Various models of processes in and around (linked) open data have been put forward under different headings. They have been termed the open data life cycle, the open data value chain or plain open data process (Zuiderwijk, Janssen, Choenni, Meijer, & Alibaks, 2012). The different terminologies illustrate different purposes – practical guidance (Hyland & Wood, 2011) or analytical separation – and foci. Whereas value chain models focus more on the creation of value during open data usage (Julien, 2012), the life cycle models aim to structure the handling of the data itself. Existing process models focus on activities within public administration, such as generating, editing and publishing the data without paying too much attention on the outside-use.

Most models contain similar elements and differ only regarding semantics, granularity or the extension of the process. Hyland et al. (2011) provide a six-step guidance model that contains the steps to (1) identify, (2) model, (3) name, (4) describe, (5) convert, (6) publish the data and the reverse activity to maintain it, similar to Villazon-Terrazas et al. (2011). Another model by Hausenblas et al. (2010) also includes the user perspective, adding the steps “discovery”, “integration” and “use

cases". With the ambition to build tools to support creating linked data, the LOD2 project developed a more granular 8-step lifecycle model (Auer et al., 2012). LOD2 broadly distinguishes citizens, public administration, politics and industry as the main stakeholder groups and additionally media and science. Thereupon user types are derived, namely: "producer and publisher", "user and producer", "user and consumer" (M. Martin, Kaltenböck, Nagy, & Auer, 2011). Synthesizing various models, van den Broek et al. (2011) derive a lifecycle model comprising the steps (1) identification, (2) preparation, (3) publication, (4) re-use and (5) evaluation.

All of these models have in common that they describe a consecutive, one-dimensional arrangement of activities that an unspecified set of actors repeatedly undertake in order to provide a formerly unexposed amount of data to an abstract general public. Furthermore, these models incorporate only one analytical level. They predominantly take the operational day-to-day processes into account (such as extracting, cleaning, publishing and maintaining data), while largely neglecting the strategic processes (such as policy production, decision making and administrative enforcement). Therefore the decision making processes of which data will be published, who extracts data, how are data edited, how data can be accessed, which licenses are available, how data privacy and liability issues are treated, who is involved in these decisions etc. remain under-appreciated. These issues point to another deficiency of most open data process models: These process models are mostly actor-blind. Van den Broek et al. (2011) assign five internal stakeholder roles to the various steps of the life cycle: "top management, information manager, legal advisor, community manager and data owner". Furthermore, they make some reference to the strategic issues, but intermingle them with the operative process.

However, there is some literature that takes a broader perspective at the processes around open data, at the policy-making-level as well as at the implementation process (see e.g., Blakemore & Craglia, 2006; Courmont, 2012; Heimstädt et al., 2014; Zuiderwijk & Janssen, 2014). With regard to policy-making content-related analyses illustrate considerably different emphases (Huijboom & van den Broek, 2011), however with little regard to stakeholders involved and the role they might play (see Huijboom & van den Broek, 2011; Zuiderwijk & Janssen, 2014). With special reference to the European level, Blakemore and Craglia (2006) point out the role the European commission, in particular its Directorate General responsible for the information market, plays in the shaping of the understanding and regulation of public sector information, as well as the national governments represented in the Council of Ministers. The latter largely act as advocates of PSI producers who in general favour a restrictive understanding of PSI and want to preserve their rights to charge for the dissemination of data (K. Janssen & Dumortier, 2003). Due to the limited authority the EC can exert in this area, the national government largely retained their autonomy to decide how to disseminate data. Merely in regard to geospatial data has a wider agreement been reached that also involves conventions about standards (quality, data and meta data harmonization) (Blakemore & Craglia,

2006). Here, the inclusive approach also involved domain experts and various online public consultations (Blakemore & Craglia, 2006). Thus, while the EC is lacking far-reaching legal authority, it shapes the discussion by influencing the agenda and reaching various stakeholders.

Critically looking at who these stakeholders are especially in the PSI re-use industry, Bates (2012) distinguishes between multi-national corporations and conglomerates from various industries, SMEs, micro enterprises, independent developers and voluntary civic hackers (see also Mayer-Schönberger & Zappia, 2011). This distinction is largely based on size, only the latter category taking into account the different motives. Nevertheless, Bates draws a distinction between benevolent and naïve transparency activists and profit-seeking and exploitive corporations (Bates, 2012). Similar distinctions are sometimes drawn between transparency and accountability advocacy on the one hand and commercial re-use on the other (K. Janssen, 2012; Yu & Robinson, 2012). In regards to open data, the two groups interests largely overlap, but also show significant differences in terms of contents, shape and rights of use of the data. Regarding the content of data “for innovation and economic growth this generally includes geographic data, postcodes, transport data, corporate data and other business information [whereas a]ccountability advocates will rather be interested in budget and spending data, legal information, and procedural items such as meeting minutes and reports” (K. Janssen, 2012). Regarding the shape of the data, the role of technology has become more prominent in open data compared to freedom of information, stressing issues like machine-readability, formats etc. (Yu & Robinson, 2012). Open data activists also tend to be more technology-savvy than traditional transparency advocates (K. Janssen, 2012). The most pronounced difference between the transparency and re-use is the rights-debate: Whereas transparency is about access rights in the context of freedom of information, re-use of PSI puts stronger emphasis on rights of use in terms of licensing (commercial vs. non-commercial, liabilities etc.). Considering these differences, we will separate these groups in our stakeholder classification in terms of advocacy as well as (intermediary) users, e.g. with civic activists coming from the transparency movement and independent developers putting stronger emphasis on re-use.

The political nature of decisions during the implementation of open data (portals) has been discussed by Courmont (2012) who focuses on the politics of legal, economic and technical decisions. The actors involved are not at the centre of the article and only cursorily mentioned, such as open data infrastructure providers (e.g. Socrata) and advocacy groups from civil society. Courmont states that these political choices are rarely discussed and often “imposed by public authorities without any debate” (Courmont, 2012), thereby treating them as a monolithic bloc. With some more detail, a distinction between the policy-level and the government agencies actually owning the data is sometimes made (see e.g. Huijboom & van den Broek, 2011; van den Broek et al., 2011). However, a more fine-grained understanding of the stakeholders involved in implementation seems necessary to comprehend the shaping of open data in this crucial phase.

2.2 Stakeholder barriers in open data

The term stakeholder is in the context of open data often used when identifying “barriers” (Barry & Bannister, 2013; S. Martin, Foulonneau, Turki, & Ihadjadene, 2013) or “impediments” (Zuiderwijk, Janssen, et al., 2012) to open data. The literature about barriers and impediments often makes only implicit reference to stakeholder roles or treats them very broadly, e.g. as users. Furthermore, the identified barriers are to some extent as structural, legal and technical and not directly related to specific stakeholder roles. However, the literature is very extensive and detailed in identifying impediments which pose a challenge specifically for users of open data (see esp. Zuiderwijk, Janssen, et al., 2012). This also gives some indication for requirements users have. Table 2.1 gives a summarised overview of the barriers found in the literature, drawing upon a modified classification of Zuiderwijk et al. (2012).

Table 2.1 Barriers of Open Data; reshuffled, complemented and condensed list from Zuiderwijk, Janssen et al. (2012)

Category	Barrier, Impediment etc.	Sources
Availability and Quality	Completeness: Incomplete, partially available data; lacking data; unpublished data; non-original, processed data; insufficient meta data	Blakemore & Craglia, 2006; Zuiderwijk, Janssen, et al., 2012; Zuiderwijk, Jeffery, et al., 2012
	Timeliness: outdated, non-updated data	M. Janssen et al., 2012; Lee & Kwak, 2012
	Relevance: largely irrelevant, uninteresting data	Blakemore & Craglia, 2006; M. Janssen et al., 2012; Zuiderwijk, Janssen, et al., 2012
	Quality: Bad, inaccurate or indeterminable data quality, ontologies and meta data	Conradie & Choenni, 2012; Huijboom & van den Broek, 2011; M. Janssen et al., 2012
Accessibility and Findability	Search Barriers: finding the proper dataset; no advanced search facilities; missing, incomplete or incorrect meta data	Conradie & Choenni, 2012; M. Janssen et al., 2012; Zuiderwijk, Janssen, et al., 2012
	Access Barriers: Registration requirements; requests necessary; exclusive, restrictive access	Blakemore & Craglia, 2006; M. Janssen et al., 2012; Meijer & Thaens, 2009; Napoli & Karaganis, 2010; Zuiderwijk, Janssen, et al., 2012
	Cost Barriers: fees due; inappropriate pricing	Huijboom & van den Broek, 2011; K. Janssen, 2011; S. Martin et al., 2013; Zuiderwijk, Janssen, et al., 2012

	Fragmentation: fragmented sources; duplicated data	Conradie & Choenni, 2012; Vickery & Wunsch-Vincent, 2006
Usability	Licensing: Restrictive licenses; incoherent licenses; incomprehensible licenses	K. Janssen, 2011; S. Martin et al., 2013; Zuiderwijk, Janssen, et al., 2012
	Machine-Readability: non-machine-readable formats; lack of good API	Zuiderwijk, Janssen, et al., 2012
	Reliability: no reliable long term preservation of data; unclear provenance of data and trustworthiness of source	S. Martin et al., 2013; O’Riain, Curry, & Harth, 2012; Zuiderwijk, Janssen, et al., 2012
	Manual Effort: various arbitrary data transformations necessary; data cleaning	Ding et al., 2011; M. Janssen, Charalabidis, & Zuiderwijk, 2012
	Linking: difficulties to link data and meta data	M. Janssen et al., 2012; King, Liakata, Lu, Oliver, & Soldatova, 2011; Zuiderwijk, Janssen, et al., 2012
	Interoperability: open data infrastructures not interoperable with other systems; fragmentation of software and applications	M. Janssen et al., 2012
	Compatibility of Vocabulary and Structure: different definitions of data, terminologies; too much vocabularies; general lack of standards	Conradie & Choenni, 2012; Huijboom & van den Broek, 2011; M. Janssen et al., 2012; S. Martin et al., 2013; Zhang, Dawes, & Sarkis, 2005
Understand-ability	Meaning: domain knowledge necessary to understand data; jargon in data and meta data; meaning and meaningful interpretation of data are unclear; statistical expertise necessary	M. Janssen et al., 2012; King et al., 2011; S. Martin et al., 2013; Zuiderwijk, Jeffery, & Janssen, 2012
	Validity: methods of data gathering unclear	Zuiderwijk & Janssen, 2012
	Support: no expert advice available; lack of service by data providers to use raw data; lack of dialogue between data producers and consumers	M. Janssen et al., 2012; S. Martin et al., 2013
	Visualisation: data and meta data are not visualised	Zuiderwijk, Janssen, et al., 2012



The barriers listed in table 2.1 are largely expressed from a user point of view. However, they also give some indication of the multitude of other stakeholders involved in the more general decisions, e.g. regarding licensing, costs and fragmentation/consolidation of sources. As Huijboom and van den Broek (Huijboom & van den Broek, 2011) state, “whereas the drivers lie predominantly outside government, the barriers are within government organisations.” Whether or not this is entirely true, various barriers mentioned in the literature indirectly refer to the vertical and horizontal differentiation between different governmental levels. Especially regarding policy making in the context of open data, these levels' roles are rarely appreciated. Accordingly, open data process models treat the public sector as a uniform entity. However, differences i.a. in regard to authority, amount of data, technical capabilities are notable among the local level, the regional level, the national level, the supranational level (e.g. the EU), the international level (e.g. the World Bank), and non-political levels (e.g. academic institutions). Predominantly the legal literature appreciates the role this administrative and political complexity plays (K. Janssen & Dumortier, 2003; K. Janssen, 2011).

One core element of open data is making data publicly available that is previously stored and kept internally by public agencies. Similarly, inter-agency information sharing involves making data available beyond the organisation's boundaries, if only within the public administration. Here, previous research highlights how different values and cultures within public administration impact on information sharing practices (Yang & Maxwell, 2011). For example the kind of data that is collected and stored, how data are defined and how it shall be interpreted and used is affected by professional values (Dawes, Cresswell, & Pardo, 2009). Making data available thus means giving up the prerogative of definition and interpretation, because the data can be interpreted and used very differently in ways, what public agencies might view as detrimental (Pardo, Cresswell, Thompson, & Zhang, 2006). The completeness, accuracy of information and the misusing of the data by others which might even incur liabilities is considered as a further barrier for data providers (Yang, 2012). The lack of resources, especially staff shortage, is also mentioned as to inhibit data sharing. (Landsbergen Jr. & Wolken Jr., 2001; Yang, 2012). On the other hand trust has been shown to positively impact knowledge sharing practices (Willem & Buelens, 2007). These factors can be presumed to have a negative effect on open data as well, especially since most of them should be more favourable within the public sector than beyond sector boundaries. However, how these factors impact on implementing open data and what the results are, is so far barely understood.

2.3 Preliminary ODM Stakeholder Classification

Looking at the current literature about open data, we derive a preliminary three-layer stakeholder distinction. At the core are the **stakeholders involved in the operational day-to-day activities around open data**. A middle layer consists of those **stakeholders who are involved at the policy-level of the open data case in point**. The outer layer consists of **stakeholders who are involved in open data at a more general, abstract level** that have an indirect effect on the open data project at hand.

Involved at the operational level of open data are the so-called data owners, often the administrative department or some other public entity that generates the data in the course of their day-to-day activities. Depending on the institutional arrangements, the data owners or support units, like e.g. the IT department, can be responsible for extracting, cleaning and curating the data. Subsequently involved are the platform providers who host the open data hub. These providers can be an internal ICT service provider or an external platform provider. Various types of users are involved at the operational level too. The user side comprises to a large extent intermediary users, who work with open data to provide a product (e.g. an application or a newspaper article) for end-users (Heimstädt et al., 2014). These end-users of open data can be individual citizens, private for profit or non-profit organisations or public sector organisations. There are, however, also end-users who make direct use of open data for their own purpose, e.g. business intelligence, market analysis (Deloitte, 2012). Intermediary users can e.g. be application developers, researchers or data journalists. A special case are (meta-)platform providers/aggregators who facilitate data use but do not necessarily use or process open data themselves. They contribute parts of the infrastructure around open data (e.g. Socrata).

The immediate policy-level contains stakeholders that exert some influence on whether and how open data is taken up in a specific jurisdictions or organisation. These are various policy-makers, parliamentarians involved in legislation about open data, the political executive and senior level bureaucrats/mandarins as well as the strategy unit in charge of open data and other administrative departments. Furthermore, external stakeholders who have a stake in the specific project, like e.g. civic and corporate advocacy groups engage in deliberations. The wider context of open data includes international advocacy groups and networks, like OKF or ODI. Also in this category are policy-levels that indirectly influence a specific open data project, for instance higher and lower-level governments, supra- and international bodies.

This is a generic model, but the specific stakeholders of an open data case in point presumably depend on institutional characteristics, e.g. status of the ICT industry, state structure, ICT and data governance, privacy legislation and administrative culture. How some of these factors interact and play out will be more specifically analysed based on the empirical data.

3 METHODOLOGY

In order to analyse this exploratory question, a qualitative research strategy appears adequate, because of its ‘thickness’ and conceptual openness to the phenomenon. Thereby, numerous potentially relevant aspects, like actor’s perceptions and constructions, as well as institutional aspects can be apprehended. The research goal is to gain a deeper understanding of the processes around open data within public administration. Thus, the empirical data collected and analysed for this article are interviews and policy documents.

For this research effort about 30 semi-structured interviews were carried out with public administrators involved with open data as well as open data users, advocates and consultants. In order to grasp open data as broader, multi-governmental and multinational phenomenon interviewees from two continental European countries (Spain and Germany) were selected. In order to reflect the whole spectrum of governmental levels, we interviewed participants working in municipalities, at the regional and at the national level. The interviews lasted around one hour each and included three topics: (1) perception of open data, (2) governance structure around open data and (3) actual processes of data publication and usage. The interviews were recorded, coded and analysed with the help of the qualitative data analysis software Nvivo. The transcripts were coded based on a coding scheme derived from the relevant literature on the topic which was integrated into the interview guideline. This coding scheme was refined during the coding process considering the empirical data.

Table 3.1 Composition of Interview Participants

Sector Affiliation		Country of Origin			Governmental Level			
Public	Private	Spain	Germany	third party c.	national	regional	local	n.a.
20	13	13	15	5	8	2	10	13
33		33			33			

4 STRATEGIC PROCESSES REGARDING OPEN DATA

4.1 Policy-Making around Open Data

Topics strongly related to open data are virulent at various political levels for many years by now, as outlined above. Especially at the European level, the EC has pressed to facilitate PSI re-use by attempting to pass general rules about licensing, terms of re-use and pricing. However, lacking a clear legal mandate and with infeasible consensus, the EU Member States have largely retained their influential role¹ (K. Janssen & Dumortier, 2003; K. Janssen, 2011). Nevertheless the EC consistently promotes PSI re-use, now also under the term open data, to strengthen the European information economy. Also, a PSI directive finally passed (Public Sector Information Directive 2003/98/EC) that required member states to implement PSI re-use regulation into national law, thereby putting the topic on the national political agendas. Lacking wide-ranging formal legal power, the European Commission does also have instruments and means to push initiatives, like the smart cities-concept or research projects, which act as a transmission belt for topics like open data between the European level and the local level. Also in this category of influential stakeholders are other **inter- and supra-national bodies and associations**, like the G8 who passed an Open Data Charta or the Open Government Partnership (Bates, 2012) or EuroCities. These are associations of countries or cities, though less formal and comprehensive than the European Union, who come together to agree on common terms around open data. Somehow surprisingly, countries agree in these deliberations that otherwise resisted the attempt of European harmonisation.

Involved in deliberations at the international level is the ICT industry that promotes the topic in general, also through associations such as the PSI Alliance (Bates, 2012). They have vested business interests and thus stress specific issues in the discussion, e.g. licensing terms, competition in the information market and costs. These **corporate advocacy groups** differ quite significantly from **civic advocacy groups**, even though their interests in some respect overlap. While the ICT industry comes from the PSI re-use perspective, civic advocacy groups are strongly influenced by the transparency tradition, putting government under scrutiny and unearthing troves of administrative data. Also, some **media outlets** join in the debate and act as advocates for open data (see Arthur & Cross, 2006).

Despite prominent and high-voiced advocacy at the international level, the legal framework around open data is largely national. Here, the legal framework allocates the mandate in regard to open data, or it does not, but leaves it an open question. The role of the political-administrative system, e.g. state structure (federal/unitary; centralised/decentralised) or administrative culture (public interest/legal state) has been highlighted in regard to other reform waves (Pollitt & Bouckaert, 2011).

¹ The directive indicates the minimal compromise, stating that it „establishes a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public sector bodies of the Member States“. (European Parliament & Council of the European Union, 2013)

Assumptions can be made that these factors as well as e.g. the role of the individual in relation to the state and the history of freedom of information and transparency legislation do play a role in open data too and the superior results in Anglo-Saxon countries suggest an influence. However, this has not been systematically researched so far.

In general, notwithstanding any general open data paradigm, the decision whether or not to publish a dataset is left to the administration that “owns” the data, which means generates and stores it in the course of their public task. Which organisation this is happens to be strongly influenced by the overall state structure as well as political decisions not related in any way to open data. Thus it can be stated, that who is the data owner is an arbitrary assignment in regard to open data. As a consequence, highly decentralised, diversified, federal states will, all else being equal, find it more difficult to act jointly, swiftly in regard to open data, because coherent action requires intensive negotiations and compromises, sometimes called the “joint-decision trap” (Scharpf, 1988, 2005). Implementing European regulations, such as the PSI directive, in countries like Germany and Spain thus involves several levels of government, possibly having to pass bi-cameral parliaments and negotiations of the administrative levels with various policy fields each. During this process, what is considered open data is shaped by the various stakeholders involved, as will be shown.

Spain

In Spain, the law 37/2007 of 16 November on the Reuse of Information in the Public Sector passed by **Parliament** transposed the PSI directive into Spanish national law, more than two years after the established deadline (Garcia & Soriano Maldonado, 2012).² The law goes beyond the pure transposition of the directive, e.g. requiring public administrations to provide electronically accessible catalogues of the available data. It is seen as a general framework for open data, which is, however, not specific and action-oriented enough, since it leaves the burden to request largely on the potential re-user. Therefore, the Ministry for Industry, Energy and Tourism (MINETUR) in coordination with the Ministry of Finance and Public Administration (MINHAP) – the **ministries pushing open data** – initiated and launched the Aporta project in 2008 (Garcia & Soriano Maldonado, 2012). The project focused initially on promoting the topic of open data, especially in the central government, organised community engagement and fostered studies estimating the economic impact of PSI and open data reuse (see Proyecto Aporta, 2011, 2012). It also published an initial PSI catalogue early in 2010. In order to overcome some of the barriers open data in Spain faced (see Garcia & Soriano Maldonado, 2012), a royal decree (Royal Decree 1495/2011, of 24 October) was issued. It contains as a main principle that all public sector information is reusable but exceptions that have to be justified, a standard open data license for the national government, conditions for re-use, requires every government organisation to establish some sort of open data officer, prescribes

² For a discussion of preceding legislation on public sector information re-use in Spain see Cerrillo-i-Martínez (Cerrillo-i-Martínez, 2011).

measures for the operative open data process (interoperability guidelines for DCAT and URLs) and contains a soft open by default-clause. Furthermore, it mandates all national government organisations to list their data in a central PSI catalogue. The royal decree only applies to the national level, however.

The cross-level coordination structure around open data in Spain follows to a great extent the established modes of coordination within its federal-like system (Colomer, 1998; Lijphart, 1999), not only for ICT or open data, but in regard to any topic of cross-level relevance. Thus, MINHAP engages in regular working groups for open data with the regional level and the Spanish federation of local entities which represents the municipalities, provinces etc. These are affected by the 2007 PSI-law, but not by the royal decree. The autonomous regions cannot be directed by the central government to open their data, how to publish it or where to publish it. Spain is a unitary parliamentary constitutional monarchy divided into 17 autonomous communities and two autonomous cities. The autonomous communities and cities have far-reaching and exclusive competences which renders the state government with less formal powers than in a centralised state. As a third and fourth governmental tier below the autonomous communities, there are 51 so called provinces and about 8.000 municipalities. Hence, there are multiple governmental layers, each of which exercises its own rights. The various (independent) governmental layers also cause a strong vertical fragmentation regarding the public administration and policy issues in Spain. This fragmentation can also be seen in the approaches to e-administration (electronic administration), which is dominated by divided structures as well (Muñoz-Cañavate & Hípola, 2011).

However, the Spanish public administration has increased efforts in recent years to raise coherence and interoperability in regard to e-government. Various **coordination bodies for e-government and ICT** have been established, such as the High Council for E-Government and the E-Government Sector Committee, coordinating ICT-use across departments and governmental levels respectively. Besides the institutional structure, a joint legal basis has been developed. Subsequently to the law on electronic access of citizens to public services (Law 11/2007), the Spanish national interoperability framework was drafted, which applies to all public administrations in Spain. An interoperability agreement about the re-use of information resources was added to the national interoperability framework in 2011 (Ministry of the Presidency, 2011), making detailed specifications for open data in Spain.

Germany

In Germany, the PSI-Directive was transposed into national law through the “Informationsweiterverwendungsgesetz” (Law on the Re-use of Public Sector Information) in 2006. Like most other European countries, Germany failed to implement the law within the required time frame until 2005. As a consequence, the European Commission opened treaty violation proceedings



against Germany (Wirtz, 2014), thus spurring the implementation of the directive. Due to the rules of concurrent legislation, which allow the national level to decide single-handedly on norms concerning economic activities in Germany, the German federal government drafted and passed the law, without approval necessary by the second chamber of Parliament (Püschel, 2006), the Bundesrat, which represents the state governments. Hence, the Federal Ministry for Economic Affairs and Energy drafted the law. The state chamber was consulted, however, but did not raise any objections to the regulations. The German law on the re-use of public sector information is mostly identical to the regulations set forth in the directive 2003/98/EG (Wirtz, 2014). The PSI law does not regulate access to public sector information, as this is regulated by each state and the federal level individually through FOI legislation.

Parliamentary initiatives around the topic remained scarce. Electoral successes of the Pirate Party at the state level in 2011 and 2012 spurred the public discourse about FOI, transparency as well as internet- and ICT-related topics. Also, other **political parties in Parliament** started to take up related issues and positioned themselves. However, after the Pirate Party did not win any seats in the 2013 elections to the Bundestag, the media and political attention faded.

The German federal government was rather slow to pursue an Open Data policy. First steps were taken with the government program “linked and transparent public administration” in August 2010. The program was issued by the conservative-liberal government, which had taken office in October 2009. Henceforth, it was the announced aim of the German government to publish any data that the public administration collects – as long as legally possible and appropriate. Even though issues such as transparency, participation, and collaboration played an important role, the government program also emphasised the economic potential of open data to strengthen Germany as an innovation hub, to increase the competitiveness of German enterprises, and to support the public administration in fulfilling its tasks. Hence, both transparency and economic considerations appear to have been equally important for the German government. Furthermore, the government program directed that only a shared (nationwide) understanding of open data and a joint initiative towards increased openness of government could generate added value for Germany. Thus, the program prescribed a coordinated approach across different governmental levels. Therefore, open data became a topic for the then recently developed **governance structures for cross-level collaboration in e-government and ICT** among the different governmental levels, which had been established for a coordinated use of information technology.

In 2009, the German parliament had passed several amendments to the *Grundgesetz* pertaining to the relation between the federal and the state level. Amongst others, the parliament incorporated Article 91c which stipulates that the federal and the state level should collaborate and cooperate closely in the field of IT. In order to implement the requirements, the federal level and the states signed a state treaty, which mandated (1) to set up a nationwide IT Planning Council and (2) to lie

down ground rules for collaboration in IT. The subsequently established IT Planning Council coordinates e-government projects, sets IT-standards regarding interoperability and security, etc. Even though the IT Planning Council comprises high-level representatives from the federal level and the states, such as the federal and state CIOs, three representatives of the municipal associations and the Federal Data Protection Officer, the actual coverage of its decisions remains low (Hunnius, Schuppan, & Stocksmeier, 2014).

In 2010, the IT Planning Council passed the National E-Government Strategy (NEGS) which also includes a clause that relevant information by the government and the public administration should be accessible for the public. Thereupon, late in 2011 the IT Planning Council installed a project “Promotion of Open Government”, which dealt directly with the opening of Government Data on all levels of the public administration. It assigned the Federal Ministry of the Interior together with one state to lead the project management. The project commissioned and published a feasibility study about open government data in Germany in the summer of 2012. A Fraunhofer Institute, an **applied research centre** that had conducted the study was subsequently charged with developing a prototype of an open data platform. It had previously done both for its native state government in Berlin. Also a number of other state governments and especially city governments had already running open data portals, when early in 2013, GovData – The Data Portal for Germany went online as a public beta version. In parallel, the 2013 E-Government Act includes some general clauses that prescribe machine-readable formats and recommend providing meta data, without further specifications or mandates.

Germany also has a voiceful **activist movement with regard to open government data**. Especially through the purposeful use of public relations, several non-profit organisations have become well-known among and beyond the Open Data community: First of all, there is the “Wikimedia Deutschland-Society for the Advancement of Free Knowledge”, which is engaged in open knowledge in general. Another important actor in the realm of the civil society is the Open Knowledge Foundation Germany. The Open Knowledge Foundation’s aim is to foster the idea of open knowledge in Germany through application-oriented projects.

In Germany, the civil society and its organised bodies (i.e. Wikimedia or the Open Knowledge Foundation) generally seem to choose a cooperative rather than confrontational course towards the government concerning their demands for an increased openness of government data. This may have several causes: To start with, the federal government offered and offers several possibilities for the civil society to partake in the process of developing open government data in Germany. For example, the federal government initiated an open consultation process during the development phase of the project “Promotion of Open Government”. Next to that, so-called community workshops were held. In these Workshops, open data activists had the opportunity to get involved in the process of developing a German open data policy. Furthermore, the government decided to

publish the source code of GovData on GitHub. Thereby, everybody who is interested can get involved and make suggestions for improvements of the open data portal.

Table 4.1 Stakeholders and their exemplary interests in policy-making around open data

Stakeholder	Exemplary Interests
Inter- and supra-national bodies and associations	Strengthen efficiency of governmental bodies, facilitating economic growth, esp. in the technology industries and knowledge economy
Corporate advocacy groups	Push governments to publish harmonised PSI, ideally for free, in order to support business intelligence or build a business case on the data itself
Civic advocacy groups	Raise transparency in government, unearth administrative and political data to involve citizens in governmental decisions
Media outlets	Use open data for data journalism, in particular to analyse economic, societal and political issues
Parliament	Support economic growth and innovation at a general level, while retaining decision rights
Ministries pushing open data	Support economic growth and innovation, implement government agendas and fulfil political goals
Coordination bodies for e-government and ICT; governance structures for cross-level collaboration in e-government and ICT	Develop widely implemented standards to facilitate cross-organisational data exchange

4.2 Decisions about Platforms, Format Standards, Meta Data and Vocabulary

Decisions about technical standards for a specific open data initiative are sometimes made within the realm of a general national framework specifically for open data, such as in Spain. Also, policy-field-specific harmonisation initiatives or domain-specific technical standards play a role. Amongst others, European harmonisation in regard to statistics, which established common definitions, and the INSPIRE directive are examples, where – related or unrelated to open data – standards for measures



and formats have been agreed upon. Also attempts to harmonise data exchange nationally for cross-level e-government projects can serve as a basis. Thus, decisions about technical standards are made in a cluttered environment of various standards. Further complexity is caused by the various IT systems in use in a given jurisdiction. These support different – often proprietary – formats, store data in different structures and vocabularies and rarely support open data by default (Hunnius, Krieger, & Schuppan, 2014).

Spain

Technical standards on open data are negotiated in a cross-level technical working group on ISP. In these working groups, expertise on technical issues is included from associations such as W3C, from universities and a **public enterprise in charge of furthering the information society** in Spain, Red.es. The Spanish W3C chapter, an **applied research centre**, was commissioned to develop and draft the open data-related standards which were subsequently negotiated in the cross-level working groups and finally incorporated into the National Interoperability Framework (NIF). Red.es is attached to the Ministry of Industry, Energy and Tourism. Red.es implemented the Proyecto Aporta and runs the national PSI/open data portal for Spain (datos.gob.es), but is also responsible for a wide range of ICT-related topics. Together with the two involved ministries, it holds regular forums with **business associations** where these express their needs in regard to PSI re-use and also request specific kinds of data. In an annual event for people engaged with open data in Spain they attempt to further an ecosystem around open data. The organisational affiliation to the Ministry of Industry is emblematic for how open data is taken up in Spain at the national level: Open data is perceived and implemented as an industry topic, its main goals being innovation, growth and job creation. Spain is heavily affected by the financial and subsequent economic crisis since 2009. Therefore, in the national debate the economic value of PSI re-use and open data is strongly emphasised. Transparency legislation in regard to open data only followed subsequently in late 2013. Transparency is thus seen as a sub-topic, by some even considered a distraction.

Red.es is generally relatively well-resourced which puts it in a special position within the Spanish administration in the current financial situation. This gives it the means to pursue and push a topic as well as to incentivise others. Thus, although decision rights are diffused in the quasi-federal Spanish political system, influential actors at the national level took up the topic, framed it from a specific perspective on growth and pushed regulation to facilitate open data throughout the Spanish administration. Thus, **coordinating bodies around ICT and e-government** massively shape open data implementation in Spain, by not only putting the topic on the political agenda throughout the country, but also making detailed provisions for how to implement it.

As a result, the National Interoperability Framework (NIF) with its Technical Interoperability Standard for the Reuse of Information Resources provides detailed guidelines, i.a. for data selection, standards

for data formats and meta data schemes to present data in a unique, reliable, persistent manner. Also, in order to be federated to the national portal – which local and regional portals strive for, because it feeds the European portal – local and regional portals have to meet the NIF specifications. Furthermore, the research centre involved in the development of the national standards for open data consults on numerous endeavours and advocates certain technical decisions which thereby become a quasi-standard. Public administrations themselves seem to have little expertise with open data and thus rarely question recommendations or understand possible implications.

“The interoperability framework in Spain is a legal obligation for all public administrations, so they are forced to do that. If they want to federate, they have to follow the interoperability rules in Spain and that is very important.” (interview participant, national level Spain)

However, so far little cross-level harmonisation around data structures and vocabularies can be observed, which activists strive for. Remarkable is the effort of the **network of smart cities** in Spain, RECI, which is especially active in open data. It consists of cities (e.g. Barcelona) who are regarded as more advanced in open data and works to establish standards in coordination with red.es and AENOR, the Spanish **standardisation body**.

Germany

Looking at the development in open government data in Germany, right from the outset the GovData portal was designed as a joint project. Hence, the project partners – the national and state governments – have equal rights regarding the design of GovData. This also means that the federal government as main project leader cannot direct by itself the conditions under which the portal is operated. Furthermore, through the implementation of a public beta version, the federal government decided to publish an unfinished, premature model. Taking this course of action, it initially focussed on a rapid development of the portal’s content regardless of any strict standardisation issues. Also, through less rules and standardisation, the Federal Ministry of the Interior hopes to motivate other institutions, i.a. other ministries at the federal level and their agencies, to participate in the GovData project. Therefore, they kept the barriers for participation as low as possible.

As the GovData Portal is mainly considered to be a cross-level, nationwide project, it is supposed to be a central and consistent gateway to Open Data in Germany. Hence, as various interests are involved, negotiations and also disagreements on the advancement of the portal are inevitable. What complicates the situation even further is that the federal government and the other partners have progressed very differently regarding open data. While some states already run open data portals, others have not yet taken any measures and steps towards an open data policy. Thus, the structures of GovData are solely designed for the operation of the portal and the coordination of Open Government Data in Germany. The federal government, or more precisely a unit within the

Federal Ministry of the Interior, is currently in charge of operating GovData centrally and coordinating the nationwide endeavours around open government data. These tasks include: the formation of a strategy and basic principles, (technical) development, communication and marketing, etc. However, the central unit is not in charge of the allocation or the storage of the data. It solely holds available a meta data catalogue and federates the data from decentralised providers, i.e. it provides links to open data on different portals. This also implies that the providers of open data gather, edit and prepare the data themselves. As a consequence, those who already run open data infrastructures and portals by themselves do this according to their own standards.

This makes the publication of open data on the GovData portal easy, but causes potential ramifications for the use. Especially as GovData functions mainly as gateway to decentralised data, difficulties regarding a harmonised approach arise.

There is not a common *modus operandi* concerning the licensing of the data. Every data provider can decide on its own, which license is applicable. However, in order to attenuate this problem and to enable uniformity on the portal, it was decided to provide another license, which is recommended for use on GovData. This “Datenlizenz Deutschland” is specifically developed for the German legal regime and data by the government and the public administration. Two versions of the “Datenlizenz Deutschland” exist: “Datenlizenz Deutschland – Attribution (BY) – Version 1.0” and “Datenlizenz Deutschland – Attribution (BY) – non-commercial (NC) – Version 1.0”. Whereas the first license is “open” by definition, the latter does not conform to Open Data standards and should only be used in exceptional cases. Regardless of this attempt to foster a harmonisation on GovData, the spectrum of licenses in use is quite broad: Next to the two “Datenlizenz Deutschland”-licenses, the Creative Commons licenses are the most common on GovData. The move to develop a discrete license caused significant uproar among **open data activists** who put pressure on the federal government. Several activists and associations expressed their disappointment and distanced themselves from the federal project, even setting up a web page <http://not-your-govdata.de/>.

A closely related topic concerning the establishment of an Open Data portal is the characteristics of the data sets themselves. Even though the portal’s main goal is the publication of open data, i.e. openly available, usable and machine-readable data (raw data), a rapid development regarding the number of data sets on the portal was given priority. As a consequence, GovData also allows data that is not machine-readable (e.g. PDF-documents) or comes with fees. That way, the government hopes to lower the barriers for a fast progress of the portal with regard to the content. Furthermore, it is argued that the portal should also cater for the needs of those citizens who are not interested in raw data, but in simple document researches. As a consequence of these approaches, there are more than 40 different data formats accessible on GovData, some of them not “open” as per definition.



Even though licenses or data formats are not standardised, data needs to be sufficiently specified meta data in order to be published on the portal. Especially as GovData functions as a central nationwide gateway with a meta data catalogue as its main asset, the description of the data is essential. In this context, a meta data structure based on CKAN was developed. The OGD meta data structure covers, i.a. title, identifier, a description of the data, persons in charge, license, and data resources. Next to the endeavours to spread and test the OGD meta data structure on national level, the federal government also strives for an international harmonisation, i.a. with Austria.

As the federal portal is an entryway to Open Data on various governmental levels, the question of how to technically integrate the data arises. Most of the data sets are imported automatically through so-called harvesters. Basically, there are four different possibilities to harvest information from decentralised portals: JSON-Import, CKAN-CKAN-Harvesting, CSW-ISO19115-Harvesting and CKAN-REST-API. Mainly the first three ways are in use.

Looking at a specific instance of a local open data project, it becomes evident that decisions about licensing, formats etc. are largely made independently with little reference to national practice. This might be preliminary, however, since the GovData portal itself is still in a development phase and a lot of issues remain unresolved. This means that some states and municipalities are still waiting for the normal operation of the GovData portal to begin in order to devise an open government data policy for themselves. One example in this context is the State of Brandenburg. The state government basically decided to await the end of the testing phase of the GovData portal in order to be able to use its infrastructure for its own data. Nevertheless, generally most of the portals created so far on state and municipal level were established independently of the national open data project.

The city of Munich was initially one of the early adopters when its central ICT strategy unit took up open data in 2009. As a part of a citizen's e-participation initiative it also hosted a programming contest for which it made available a limited number of data sets. Because it was one of the first open data attempts in Germany, it received nation-wide media attention and won an e-government award. Also, the city council came on board and all factions filed motions in support of the idea. The project had initially received a far-reaching exemptions from ICT-related rules in the city, to set up a separate infrastructure (server, wiki), even though reservations existed.

“It [drive for open data] comes from the politicians. [...] XYZ as a city presents itself to the citizens, shows itself as a service provider, [...] as attractive and thus modernity always plays a role. E-participation is also a topic - everywhere en vogue - [...] around the elections. A topic that you can catch attention with.” (interview participant, local level Germany)

However, a subsequent attempt to formalise the process of data publication, establish an open data catalogue and involve all city departments could not be implemented. Firstly, **administrative departments** faced the idea of open data with reservations right from the outset. Furthermore, the

initial project had a budget of only about 25k€. When the follow-up project requested an additional 125k€, discussions dragged on. In addition, various extensive legal reservations were raised by **legal experts in the departments and the central legal department**. These legal discussions, around liabilities, third-party copyrights and licenses, subsequently dominated and halted the project.

An additional distraction was caused by a comprehensive IT-reorganisation programme in the city administration to centralise infrastructure and reorganise departmental ICT tasks, what absorbed resources and attention. Nevertheless, the city's **central IT strategy and controlling unit**, STRAC, pressed on and finally, in 2013, received the nod when it included open data as part of a resolution on e-government and open government.

Subsequently, STRAC entrusted a **private technology consultancy** to develop a functional specification and later on set up a platform. Lacking a national framework or nationwide established practice, the project started from scratch when proposing a meta data model and licenses, describing roles and designing processes. These far-reaching specifications were largely decided autonomously within the project and are currently not yet generally agreed upon. So far these seemingly technical decisions do not appear to cause major discussions, as seen elsewhere (see Courmont, 2012). It remains unclear whether the implications for the other actors are fully understood or simply do not affect their interests.

Table 4.2: Stakeholders and their exemplary interests in decisions about open data endeavours

Stakeholder	Exemplary Interests
Organisations in charge of furthering the information society; central IT strategy and controlling units; coordinating bodies around ICT and e-government	Motivate governmental organisations to make open data available; influence character of adoption
Applied research centres	Shape governmental technology policy, showcase technological potentials
Associations outside government, e.g. business associations	Push governments to publish harmonised PSI, ideally for free under minimally restricted terms of use, i.e. most liberal license
Associations of government organisations, e.g. network of smart cities, EURO CITIES	Raise efficiency and innovation among members and develop shared concepts, standards etc.
Standardisation bodies	Develop and enforce widely shared standards
Administrative departments	Conform with government policies, while at the same time retain professional independence

Legal experts in the departments and the central legal department	Develop policies and guidelines that conform with the legal framework and stand the tests in court
Private technology consultancies	Consult public administrations on technology governance and management

4.3 Decisions about Publishing Data Sets

According to Spanish regulation, every public administration at the national level needs a part on its web page where it publishes open data that is also federated to `datos.gob.es`. Furthermore, it forces all organisations to develop an open data strategy and establish a designated unit within each ministry in charge of open data, thereby creating a spread out network of “advocates” for open data. However, these '**open data officers**' do not have any significant decision making power over which data to publish. These decisions remain with the data owners, specific departments within the ministry that generate the data in their regular work.

What can be observed is a conflict within the public administration itself. On the one hand, there are employees whose job is to actively push the topic of open data. On the other hand there are those employees of **operational departments** who ultimately have factual decisive power over the data that is supposed to be provided. Often their intention is to shield valuable, sensitive data. Heimstädt describes this problem as one of unequally distributed benefits that relates to the characteristics of the data (Heimstädt et al., 2014). Controlling the possible usage of the data is on the one hand undertaken to shield politically sensitive data. Also, shielding data has something to do with covering up the poor quality of the data itself.

The situation then arises that *“you can have an open data catalogue with thousands of data sets, it looks pretty good, lots of data are available, but if you start to work with one single data set you see that you just have rubbish.”* (open data advocate and consultant, Spain). Without satisfying neither the initial political idea of open data, nor potential users, this practice nevertheless provides open data officers and strategists some kind of legitimacy, as well as retains control of the data to those public servants who always used to have it.

In Germany, the situation appears to be quite similar. Over all three subsequent open data projects in Munich the decision whether to publish data, which parts to publish and how to publish it is left to the **administrative departments** that generate the data in the cause of fulfilling their public task. These departments have very few incentives to publish data as open data. The strategy units in charge of open data and the political level share the praise for the launch of a platform, while blame and shame are largely left to data contributors, when false data are published. Praise for contributing



significant data presupposes political attention. There never was a strong demand from the political level to publish open data, barely holding out the prospect for praise when making their data available. Rather, they take a number of risks when publishing open data.

The risks involve the above mentioned liabilities, if the data includes copyrighted third-party content as well as possible flaws in the data. Thus numerous legal questions arise, giving special prominence to **legal departments**. Regarding the risks of possible privacy infringements, **data protection officers** are regularly involved in the deliberations. The risks also pose a potential embarrassment for public administrations that generally have a low tolerance of mistakes. Furthermore, fear is prominent that transparency might force politicians' and administration's hands to tackle societal problems that are either hard to solve or require massive financial expenses. These consequences might arise from a clearer picture of a current situation as well as different appreciations of the problem and interpretation of the data.

Also **open data users** have so far not proven or even aroused the perception among data owners that they are capable to deliver meaningful results. The number of actual users who visibly make sophisticated use of open data is perceived as low. When users file requests for specific data sets, for practical reasons they can often not make specific statements what the data will be used for and what the impact will be, for they do not know the content and quality of the data in advance. What remains is the promise of sophisticated and powerful ICT capable of processing, mashing up and analysing large amounts of different kinds of data.

Somehow exceptionally are those parts of public administration that are more used to publishing their data, though not necessarily to the general public or for free. Here ICT systems are ready-made to prepare and share data, e.g. geographic and meteorological data. Also, there is a history of standardisation across administrations and jurisdictions that lead to nation-wide or even European standards, e.g. the INSPIRE directive for geographical data.

However, these administrations often rely on their data as a source of revenue that would need to be compensated. Furthermore, there is staff in those departments that prepares data and handles the distribution and financial transaction processes. These positions might become redundant, when data is completely made available as open data. Thus, efficiency gains are not accepted as an argument. Therefore, benefits – if feasible at all – are unequally distributed.

Nevertheless, open data is perceived as a noble endeavour that lends the appearance of modernity. Thus, politicians pay lip-service to open data without establishing a framework in which it can truly thrive. As a consequence, window-dressing strategies currently best align actor's interests: Administrative departments provide data that is available in a structured format, at a fairly good quality level, not obviously sensitive, requires little maintenance and thus little effort to be

published. This can be termed the “availability approach” to publishing data (Hunnius & Krieger, 2014).

Leaving the decision about a specific dataset to the various ministries creates a spread-out network of data generators. These often have some topical understanding of what the data is about, but not necessarily a profound technical understanding of how open data are used, what requirements for use are or what facilitates their use. Also, the individual person might become some sort of internal advocate for the topic he or she is responsible for, but they mostly lack decision-making powers which is internally further dispersed to **heads of departments**. Thus, under such circumstances the internal open data advocates largely depend on the leadership from the organisation's senior level to encourage and direct departments to make data available, since policies are often discretionary and not mandatory and leave considerable leverage to the agencies (see also Shkabatur, 2012).

“In general, it is the availability of potential data. This is still a point that data are selected based on how easy it can be made available and less based on its usefulness.” (interview participant, local level Germany)

Table 4.3: Stakeholders and their exemplary interests in decisions about publishing data sets

Stakeholders	Exemplary Interests
Data owners – heads of departments	Furthering their professional cause; appearance of modernity, retain independence and influence
Open data officer	Persuade data owners to make open data available and advise on decisions
ICT/open data strategy unit	Maximise the number of data sets on the jurisdiction's portal
Legal department	Develop policies and guidelines that conform with the legal framework and stand the tests in court
Data protection officer	Ensure data privacy while also ensuring administration conforms with FOI legislation
Open data users	Showcase the benefits of open data in order to push governments to publish more open data

5 PROCESSES OF OPEN DATA PRODUCTION

The processes of open data production are the operative day-to-day activities around open data (see above).

5.1 Data Generation, Storage and Management

Data generation “occurs” in the course of an agency or department fulfilling its public task. Data that accrue are regularly processed in ICT systems, however, rarely with the goal of publishing the raw data. Instead, data are often generated without any regard for the data itself and its further use, but with a strong focus on processing the citizen's concerns. Thus, data generation activities are largely uncoupled from the data's extraction and publication. This has an impact on what data is stored and how it is defined, measured etc. Different vocabularies are not simply arbitrary technical determinations, but reflect differences in construct formation. Epistemology is often influenced by perceptions, preferences and interests, what poses future challenges for harmonising data across organisations and jurisdictions. Thus, the kind of data that accrue, how data are defined, what is documented and how it is measured is strongly influenced by professional requirements and conventions. Also, the quality of the data is only relevant if it is a determining factor in fulfilling the public task. Beyond, e.g. purely for documenting purposes, it does not seem to play a role.

The data that public administration accrues in the course of fulfilling its public task is stored in a countless variety when it comes to medium, format, location etc. Even considering only digitised data as potential open data, data storage is still very diverse. From data stored on a local PC to large data bases in computing centres, from a variety of proprietary formats used by different **ICT vendors** to some open formats, the diversity appears barely tameable. The particularities of IT governance in the public sector (Hunnius, Schuppan, et al., 2014) seem to play a role, where **CIOs and ICT strategy units** have difficulties to enforce coherent standards, as well as the role and importance of data management in the public sector in general. At least in the German case, **public ICT service providers** (Hunnius & Schuppan, 2011) that often run applications, store data and consolidate ICT procurement, have – to a varying degree – overtaken some of the roles here. However, a comprehensive overview of all the data within a jurisdiction does not seem to exist and it remains unclear whether it exists within the numerous administrative silos.

In the context of the public management reforms (Pollitt & Bouckaert, 2011) data has gained some prominence for performance measurement; also, so-called smart cities have put a strong emphasis on data and data management. Apart from that, it seems to have kept a low profile. The data are often kept internally within the **department's** or **organisation's** boundaries and thus problems, e.g. with data quality, do not become visible. Efforts to harmonise and share data across public sector organisations have proven pretentious in the past, even regarding internal use (Scholl, Kubicek,

Cimander, & Klischewski, 2012). Even where data consolidation for internal planning processes across departments has progressed significantly, as in the analysed case of the city of Munich, consolidated data do not seem to be used for open data. Instead, separate systems and processes for open data are set up.

Table 5.1: Stakeholders and their exemplary interests in data generation, storage and management

Stakeholder	Exemplary Interests
Data owners – administrative clerks	Reliably collect, document and process all information deemed relevant for the professional task at hand and fulfil any further obligations in a way to satisfy the software's requirements
Professional standardisation bodies	Develop and enforce standards to facilitate professional beliefs and constructs
Administrative departments (data owners)	Maximise departmental autonomy in compliance with jurisdictions ICT and data governance rules
ICT vendors	Develop and sell state-of-the-art ICT products and ideally create lock-in effects
ICT service providers	Ensure performance, security and efficiency in ICT and data management
CIO and ICT strategy unit	Develop and enforce data, architecture etc. standards, coherent program management; facilitate ICT use

5.2 Data Extraction, Preparation and Publication

Data extraction and preparation comprises the activities of extracting data from existing data bases and preparing the data to be published as open data. In general, these activities can be carried out automatically by ICT systems or manually by an **administrative clerk**. Here, the variety of technical systems employed by the different parts of the public administration plays a significant role. Some administrations have developed tools or updated existing applications and added functionalities to extract, clean, prepare and publish data automatically. For this, however, **ICT vendors** need to be commissioned, because most applications appear not to be open data-ready. Instead, they are programmed to keep data within their boundaries protected and secured and only allow its release



through predefined interfaces. Thus, to adjust existing ICT systems causes costs and therefore to open data relies to a large extent on manually extracted data.

Whether data is extracted automatically or manually, the **data owners** are involved not only in decisions about what data is published, which parts of the data and which formats are on offer; they also define the process and assign responsibilities for the tasks. Here, they are assisted by **ICT departments** and **ICT strategy units** who often provide guidance and make recommendations.

In the observed cases where data extraction is done manually, it is **administrative clerks from the administrative departments** themselves who are responsible for extracting and preparing the data, e.g. separating parts that violate privacy rights and adding meta data. However, these persons often have a scarce understanding of open data use and the role formats, structure, vocabulary and meta data play for using open data. Also, these tasks are generally not high on the **departments or organisations executives'** agenda and thus do not justify significant effort. Thus, because of the often labour-intensive way in which open data has been adopted in these cases, resources and priorities arguments are regularly raised.

Finally, the prepared data need to be submitted to the open data platform. For security and performance reasons, the open data base is kept separately from the internal database, so attacks and external use does not reduce internal performance.

According to Spanish regulation, every public administration at the national level needs a part on its web page where it publishes open data that is also federated to datos.gob.es. This ensures that even though data is offered dispersed and only subsequently federated, it can be found and retrieved easily. However, this does not include all autonomous regions and the local level which will nevertheless federate voluntarily in the upcoming months. In addition, federation does not enable the **open data platform provider** or the **ICT/open data unit** that commissions the platform to impose quality standards or format standards. Thus, the central open data catalogue contains data with various formats, structures, vocabularies etc. and in unknown quality, whichever way it is provided by the data owner.

In the German case, no regulation exists so far that mandates anyone to publish on the national platform. Here, even fewer standards exist regarding licensing, as outlined above. Thus, the currently 7.200 data sets by August 2014, are published under twelve different standardised licenses and for approximately 1.000 data sets some not further specified generic license is used. However, some general meta data values are required.

Table 5.2: Stakeholders and their exemplary interests in data extraction, preparation and publication

Stakeholders	Exemplary Interests
Data owners – administrative clerk	Conform with open data-related obligations with little effort as possible, when purpose is little understood and praise by superiors unlikely
Data owners – executive level	Provide as little resources as possible for open data, as long as risks posed remain higher than potential benefits
ICT vendors	Build secure, reliable ICT products; only provide open formats or open data-interfaces when these are mandatory, provide a USP or enhance the product portfolio in another manner
ICT departments	Recommend, train and advise specialist departments on open data standards and processes
Open data platform provider	Provide easy to use, low threshold platform, while ensuring performance and security of running applications
ICT/open data unit	Maximise the amount of open data sets on the central open data portal; subordinate coherent standards

5.3 Feedback Processing and Analytics of Open Data Usage

Feedback and usage analytics can be retrieved automatically by the use of the open data catalogue and the repository or manually by **open data users** making active contact with the public administration. The latter appears to be negligibly low in all observed cases. Apart from institutionalised forums for exchange, as in the case of the Spanish business associations or other events (e.g. 'hackathons'), public administration receives little feedback on its open data in general, its quality or user requirements.

What **open data catalogue providers** regularly analyse and share with data owners are statistics on data requests and retrievals that are processed via their platform. However, they are not aware and not informed about requests made via other catalogues or directly to the repository, so they do not have the full picture of data use.

Also, public administrations – be it data owners, catalogue or **open data repository providers** – rarely know about the actual use and purpose of open data. Due to privacy reasons and on behalf of the **data protection officer**, they are also sometimes not allowed to keep detailed analytics, e.g. IP addresses. Thus, they do not know whether a download is an individual citizen who immediately deletes the data or whether it is a business building an application for thousands of users on top of the data. One reason is that according to the 'open-logic' it is not required to report use of open data. Some users do and are listed by catalogues. In a related case of an administration in the city of Munich that does not provide truly open data, but imposes several hurdles to register, report future use, sign a written user agreement and subsequently provides designated access for every user, fine-grained user analytics are possible.

Table 5.3: Stakeholders and their exemplary interests in feedback processing and usage analytics

Stakeholders	Exemplary Interests
Open data repository providers	Track open data use and share analytics, if results are presentable (i.e. considered positive)
Open data catalogue provider	Track open data use to understand what high-value data sets are
Data protection officer	Ensure privacy rights of data users and therefore minimise usage analytics that disclose personal data
Data owner	Depending on the political relevance of open data, either indifference or share usage statistics that are presentable or shield them, if usage is meagre
Data user	Provide feedback if improvement is foreseeable and personally useful

6 PROCESSES OF OPEN DATA USE AND CONSUMPTION

The processes of open data use and consumption encompass all activities predominantly users of open data undertake that employ open data for any kind of purpose.

6.1 Data Detection

Data detection circumscribes the initial discovery of a data set by a potential user. Here, special attention needs to be paid to the fact that the users interviewed for this report were predominantly small-scale users of open data, often building applications as independent or even free-time users of open data besides their professional occupation. The latter was however related to information technology in all cases. All of the interviewees were technology professionals, programmers in most cases, but also web designers.

The approaches and strategies appear to be vastly different. Whereas one approach can be labelled more data-driven, the other can appropriately be labelled more issue-driven. **Issue-driven users** regularly have a certain interest in mind and know in advance which kind of data they therefore need. They often search via search-terms and keywords in open data catalogues or use general search engines. For them, open data catalogues and portals provide fairly helpful search masks. However, the meta data provided by the catalogues generally does not give them with all the information they need to decide, whether they can use a data set retrieved from the search results. This cannot solely be reduced to meta data quality or a lack of standardised meta data, but also that what is included in common standards is not seen as sufficiently comprehensive and meaningful. Furthermore, they frequently voiced complaints about the scattered portal landscape that seems to them still barely integrated. Thus, they have to conduct similar searches in various catalogues. Another strategy followed by several issue-driven users is to harvest data from governmental websites and subsequently request and negotiate terms of use or even to directly access a specific data set they need which is not made available as open data and thus not listed in any catalogue. This latter strategy even yields success frequently. However, with both strategies – informal requesting and harvesting – questions remain and prove difficult to dissolve about the license and what the data can be used for.

On the other hand, **data-driven users** look for complex, comprehensive, and large datasets, largely without regard for the specific content of the data itself. Often, they do not have a specific purpose in mind. Their presumption is that an interesting data set can be put to a purposeful use. Currently, they feel little supported by the open data portals, since these rarely support search queries that meet data-driven user's needs. Helpful for them seems rather algorithms that analyse the size of a dataset (columns, data points, whether a dataset contains string-data and numeric data or structured and unstructured data), update frequency and whether it is linked or non-linked data. However,

questions remain how to identify relevant, sensitive datasets, because too many datasets are simply published since they are at hand, but of little use (see above).

Overall, there seems to be a relative indifference to meta data standards and even meta data in general. This might be attributable to the scarcity of meta data, the low quality and the lack of content-related meaning – as opposed to formal characteristics of the data set – a lot of the available meta data convey. User interests, especially issue-driven user interests appear to point more strongly to the vocabulary and content of the data, feature that remain largely unharmonised and undescribed as of today.

Table 6.1: Stakeholders and their exemplary interests in data detection

Stakeholder	Exemplary Interests
Issue-driven users	Detect data sets with a specific content or related to a certain topic
Data-driven users	Detect large, complex data sets

6.2 Data Retrieval and Use

The activities to retrieve and use open data include not only how data are technically integrated into applications, e.g. via bulk download, dynamic integration, etc., but also what the wider context of these activities is.

On a technical level, users face difficulties to dynamically retrieve and integrate data where URI point to html-web pages on which data sets in various formats are offered, instead of the specific data sets itself. However, a larger problem seems to pose uncertainties around the terms of use, laid out in licensing terms. Here, a distinction seems necessary between **professional large-scale corporate users of open data** on the one hand and **hobby-users** on the other hand. The latter appear not to be overly concerned with intricate legal questions around open data, but seem content with any standardised license. Quite the opposite, professional users appreciate the problem and raise strong concerns about liabilities and other legal repercussions. This might be attributed to a more thorough understanding of the legal context, with legal departments involved in corporate decision making; another reason seems to be a more complex use of data, e.g. mashing up data from different sources. This amplifies the licensing problem, when sticky licenses and incoherent, sometimes country-specific licenses are used that are incongruent and incompatible.

The use of the data varies considerably and is so far little understood. It always appears to involve extensive work with the data itself, cleaning it, checking its quality. The kind of work depends on numerous aspects, such as the purpose, technologies and on the kinds of data, e.g. whether it is

static data, rarely updated or real-time data. Understanding these activities requires further research which is, however, not within the scope of this deliverable.

More important for the purpose of this research appears to be the context in which open data is used. Users of open data, except occasional lay-users who are out of the scope of this deliverable, appear to be frequently engaged in or at least in contact with **advocacy groups**, such as the Open Knowledge (Foundation), OpenKratio etc. To reveal and illustrate the benefits of open data, these advocacy groups commonly use event-driven approaches (e.g. „hackathons“). For example the Open Knowledge Foundation frequently organises projects like “Stadt Land <Code>” (loosely translated to “City State <Code>”). This project was structured as a competition, where contestants could present their ideas for applications in the field of public transportation, utilities, infrastructure, politics, and civil society in general. The competitors with the best ideas were given the means and resources to realise their idea. That way, the Open Knowledge Foundation demonstrated the importance of civic apps and Open Government Data as its foundation. Another major project by the Open Knowledge Foundation, which is also internationally renowned, is “apps for country XYZ” (cp. Apps4Finland, Apps für Deutschland etc.). Together with two other advocacy groups, the Open Data Network and government 2.0, the Open Knowledge Foundation organised “Apps for Germany” also as a competition for the best civic apps which are based on open government data. In this case, the German Federal Ministry of the Interior even operated as a patron for the project. Also common are “hackdays”, during which over the course of a few days groups of users develop applications for open data. The results often showcase the application of data, sometimes even imaginary, simulated, idealised data, since actual data do not meet the requirements or are not available. However, even when actual open data is used, the sustainability remains questionable, since open data users frequently abandon the endeavour which is thereupon not updated or built upon. Despite the fact that open source platforms to jointly develop and sustain artefacts (e.g. GitHub) are used frequently, transfer, forking or maintenance appear to be rare.

Table 6.2: Stakeholders and their exemplary interests in data retrieval and use

Stakeholder	Exemplary Interests
Professional large-scale corporate users of open data	Collect and use harmonised, standardised, high-quality, reliably provided data sets with no, minimal or foreseeable strings attached
Hobby-users	Collect and use interesting data sets
Advocacy groups	Showcase benefits of open data based on applications

7 STAKEHOLDER REQUIREMENTS

General distinctions between user groups can be drawn based on sectors (private/public) or the degree of organisation (individual/collective/corporate). However, for the questions at stake here, it remains paramount to translate user characteristics and requirements into necessary functionalities of the ODM platform. Therefore, we focus on stakeholder group interests, requirements and understanding of the topic and level of technical expertise in regard to open data.

Table 7.1: Stakeholder Requirements and ODM Potential

Stakeholder	Requirements	ODM-Potential
Policy Makers: Parliament, ministries pushing open data, coordination bodies for e-government and ICT, governance structures for cross-level collaboration in e-government and ICT	Understanding barriers to open data publication and use; understand, develop and enforce widely used standards (formats, structure, licenses etc.)	Benchmark volume and sophistication of the published data as well as its use; highlight coverage of used standards; present usage of open data; metrics per geography (see D2.3)
Commercial User (Associations): corporate advocacy groups, business associations, media outlets	Detect high-value data sets with minimally and transparent strings attached; detect mashable, harmonised data sets on a large scale;	Highlight high-value data sets (e. update frequency of a data set); map mashable content (congruent licenses, harmonised structure and vocabulary); highlight coverage of used standards (esp. licenses); metrics per geography and per data set
Civic Advocacy Groups: civic advocacy groups	Advocate the publication of and detect already published politically sensitive data sets (politico-administrative)	Highlight and compare sensitive data sets to advocate their publication in other locations; map mashable content; metrics per data set and per geography

Government bodies and associations: inter- and supra-national bodies and associations, coordinating bodies around ICT and e-government, public enterprise in charge of furthering the information society, network of smart cities, standardisation bodies	Advocate the publication of high-value data sets; benchmark volume and sophistication of the published data as well as its use to name and shame understand coverage of used standards to align with these; understand what constitutes high-value data sets to advocate their publication	Benchmark volume and sophistication of the published data as well as its use; highlight coverage of used standards; highlight high-value data sets; metrics per geography, catalogue and data set
Data generating and (potentially) providing government bodies	Understand what constitutes a high-value data set in their professional domain; learn about standards in open data in general and their professional domain; understand how open data in their professional domain can be used	Highlight high-value data sets by domain or topic; highlight coverage of used standards (licenses, structure and vocabulary) by domain or topic; highlight applications of open data by domain or topic
Technology providers: Private technology consultancies, ICT vendors, (public) ICT service providers, Open data platform providers; applied research centres	Understand widely adopted technologies and standards to align with these	Highlight coverage of used infrastructure, technology and standards (formats, licenses); metrics per geography and overall

Groups³ for which generally little technical expertise has to be presumed are policy-makers, data generators and some of the support units. Nevertheless these groups are involved in major decisions about open data and shape its conceptualisation and implementation. Policy-makers (parliamentarians, high-level executives) are involved with open data at a rather abstract level. However, their commitment to and interest in the topic in general has a significant impact on how the machinery of government approaches and implements open data. Insightful for policy-makers is to see how their sphere of responsibility (jurisdiction, organisation) compares to others in regards of volume and sophistication of the published data as well as its use. This serves as a basis to benchmark their performance and identify fields of strategic interest. Therefore, for them it is

³ The results of this research have also been reported in Deliverable 3.2 Tool specifications, use cases, mockups and functionalities status report 1.

necessary to see, what data is published by other public sector organisations and how frequently this data is used. Thereby they can get a better understanding of high-value datasets. At the moment administrations often pursue an "availability approach" to publishing data: They publish data that is available in a structured format, at a fairly good quality level and not obviously sensitive, because they lack a profound understanding of what data might be useful. At a more specific level, policy-makers pass laws, issue executive orders or policies about open data that shape how open data is published (e.g. prescribe licenses, formats, meta data standards or even paradigmatic shifts to consider everything open by default) (see Zuiderwijk & Janssen, 2014). However, these decisions are mostly prepared by ministries or other governmental departments, considered here as support units further below.

Another group that approaches open data from a rather thematic and legal perspective are data generators who typically hold the data and often consider themselves as data owners. They generate data in the course of their regular work and are predominantly responsible for the decisions whether and which of this data to publish as open data. Besides information about which data from their subject area is published by other organisations (see above), more detailed thematic and technical aspects are relevant for the decisions they make in terms of open data. Information about data structure, vocabulary and measurement scales could provide guidance for data generators how to publish their data, although they often seem to be unaware of its significance. Here, various European, supra-national and national conventions exist – some codified, others not – in various policy fields which could be built upon, as has been demonstrated with the INSPIRE directive. At a basic level, insights in which meta data schema are used could be helpful. At a far more sophisticated level, patterns in data structures and vocabulary might assist. Furthermore, data generators appear largely unaware of how open data is used and often seem to lack imagination of its possible use. In this respect, successful use cases of open data could prove insightful for them. In addition, the legal perspective is especially significant in the public sector. This is in particular true for data generators and for support units (legal department, data protection officer) who are involved in decisions about which data to publish, with which level of detail and under which license. Thus, such information could assist their decisions about licensing, liabilities and privacy protection. On the whole, data generators are not fully aware of the topic open data, do not initially endorse the idea of publishing data and have not yet integrated open data processes in their routine activities. Therefore, it poses a challenge to even attract this group to information about open data.

IT strategy units, platform providers and private consultancies often have a higher level of technical expertise, although not necessarily in regards to open data and how it is used. They are involved in decisions about portal architecture, publishing processes and to a varying extent can set standards for data published in a catalogue (data format, meta data standards, quality). For these decisions,

information about the spread of platforms (e.g. CKAN), meta data schema and data formats could help them in establishing a state of the art open data portal.

Among the intermediary and end-users, advocacy groups stand out as a group which does not necessarily use open data itself, but gathers and publishes information about open data to further their cause (see Davies, 2013; Open Knowledge Foundation, 2013). Thus, they require a breadth of detailed information about open data, in particular for benchmarking purposes. Advocacy groups in general have a sophisticated technical understanding of open data so it is not necessary to reduce complexity for them in this respect. Quite the contrary, in order to illustrate which catalogue hosts the most exhaustive meta data and points the most comprehensive and sophisticated datasets advocacy groups need to look at technical details. For lobbying efforts, it is necessary to trace back datasets to specific territories, policy fields and organisations. Especially the content-relation (policy field) seems relevant, since in the various institutional arrangements in European countries, different organisations are responsible for and accommodate the same thematic data. With several catalogues by now federating data from numerous organisations, jurisdictions and even countries, this becomes more important for comparisons.

Among the immediate users of open data (esp. application developers, researchers, data journalists) further differentiation appears necessary. Different approaches in data detection which can be termed “data-driven” versus “issue-driven”. Issue-driven users search for datasets in the context of a specific topic, because they have a certain interest and know in advance which data they therefore need. They search directly on an open data platform via search-terms, specific keywords. For these kinds of users, portals/catalogues provide fairly appropriate search masks. Thus far, however, they can only search in a specific catalogue and find results of the data referenced there. Since portals often contain only meta data about data from a specific jurisdiction or even organisation, users might have to search in different catalogues instead of looking into one meta-catalogue. Furthermore, the lack of meta data quality often inhibits or restricts the ability of these users to find relevant datasets. Thus, a meta-catalogue would be even more powerful, if it provided a search mask not only for the meta data in the catalogues, but the datasets themselves which are hosted in the repositories.

Data-driven users on the other hand look for complex, comprehensive, and large datasets, irrespective of their specific topical content. Their assumption basically seems to be that a complex dataset can be put to a purposeful use, even without a prior idea. Until now, they find scant support on existing catalogues to identify relevant, sensitive, high-value datasets. Since a number of datasets are simply published because they are at hand, catalogues are stacked with data of little use for these users and search term queries are of little help. More relevant would be algorithms that analyse the size of a dataset (columns, data points, whether a dataset contains string-data and numeric data or structured and unstructured data), its update frequency or whether it contains linked or non-linked data.



Another distinction seems necessary between part-time/hobby-users and large-scale, often corporate users. The latter appear to put stronger emphasis on legal issues, in particular licensing. This is related to restrictions for commercial use that some data comes with, their deeper understanding of the potential financial and legal risks involved and often a more sophisticated use of open data. For example a more complex use of data might involve mashing-up data from different sources which amplifies the licensing problem: Sticky incoherent country-specific licenses cause ambiguities and potential risks. For such users, not only the licenses used in different portals are of interest, but also possible congruity/compatibility.

A noticeable aspect among users of open data is the different understandings of meta data that become evident: What is sometimes discussed as meta data (publisher, publication date, data format etc.) does not appear to be the meta data many users find the most relevant. Their understanding of meta data relates more strongly to the structure and vocabulary of the data, e.g. which column contains what content, header descriptions, measurement scale, when data was measured, how data was measured. Such information is not only necessary to understand the specific dataset, but also explore scalability and transferability of an application that processes this data. This also explains a relative indifference to meta data standards in general, as long as there is any EU-wide standard.

8 CONSEQUENCES FOR ODM

From the analysis of the empirical data, various functionalities of the ODM platform can be derived. The different functionalities can broadly be distinguished into two groups: One group is targeted towards the analysis and identification of specific datasets and a second group is targeted towards the analysis of catalogues and the entirety of available datasets.

Comparisons/benchmarking

The platform allows comparisons between two and a number of entities (country, city, organisations) in regard to highly aggregated metrics as well as finely differentiated aspects, e.g. number of datasets per country (with or without further qualification) across all available hubs and catalogues on the one hand and on the other hand availability of a specific form of data (e.g. geographic data) per city in a certain format. A sophisticated feature of the platform should recommend suitable objects of comparison, e.g. countries, cities, organisations of comparable size, in a country with similar state structure and level of ICT sophistication.

These functions are specifically aimed at policy-makers and advocacy groups. Therefore, they should assist and allow curating full-fledged reports or exporting results (graphs, figures etc.). Also, a “data set of the week” could be chosen and highlighted, based on transparent criteria, what could create an incentive for data providers and hints for data-driven users. If the analytics in ODM are more sophisticated than of an open data catalogue, operators or open data generators might even be tempted to use analytics of their own data via an ODM reporting tool.

Usage and breadth

Functions unearthing the usage of different infrastructures and artefacts regarding open data comprise e.g., the spread of different available platforms, meta data schemes, data structures and vocabularies, data formats. These need to allow further qualifications, e.g. looking at the spread in a specific policy field, country etc. Such analyses are specifically relevant for platform providers, IT strategy units, technology consultancies and open data infrastructure developers, so an extensive technical knowledge can be assumed. The analysis of artefacts usage not only covers technical, but also legal aspects, e.g. which licenses are commonly used, broken down to jurisdictions, number of datasets or kinds of data. It should be displayed which specific rights a license grants and what use it allows. Therefore, ODM should not only extract licenses used in different portals, but also highlight congruity/compatibility of licenses when combining and mashing up data from different sources. This latter category of legal aspects necessarily has to presume a lower level of technical proficiency.

Identification of high-value datasets

A search for datasets does not necessarily have to start with contextual properties, but can also point potential users to comprehensive, sophisticated and complex datasets. Such datasets could be characterised e.g. by sheer size (columns, data points), complexity (whether a dataset contains string-data and numeric data or structured and unstructured data), their update frequency, whether it is linked or non-linked data and published under a very liberal license. Also, high-value datasets could enable data generators to get a better understanding, which datasets other data generators have published that are extensively used. This necessarily has to include usage statistics in the metric (how often is data requested/downloaded/used; how many apps use a certain dataset, what are the most used applications thereof (built on open data) and which data do they build on) and enable to reduce the search to policy fields, specific topics (e.g. forestry) and organisations. A further function should be able to highlight which overall high-value datasets are not available in specific country (or city or organisation etc.).

Since these features are helpful for data users as well as data generators and the former often have high technological literacy, different level of technical sophistication should be provided.

Search queries

Content-related search queries for specific datasets could be enhanced simply by building on top of search queries existing catalogues already offer. In addition to providing predefined analyses and visualisations of meta data, users would be enabled to run their own search queries. The ODM platform would serve as a front-end to all the portals search. An important restriction of these queries is that they often only search in meta data. Thus, at a more sophisticated level, ODM should provide a search mask not only for the meta data in the portal, but the datasets themselves hosted in the repositories listed in the portals. ODM should also be able to identify patterns in structure and vocabulary of existing datasets and thereby identify cross-portal similarities in datasets, so users see possibilities to transfer and scale applications.



9 REFERENCES

- Arthur, C., & Cross, M. (2006, March 9). Give us back our crown jewels. *The Guardian*. London.
- Auer, S., Bühmann, L., Dirschl, C., Erling, O., Hausenblas, M., Isele, R., ... Williams, H. (2012). Managing the life-cycle of linked data with the LOD2 stack. In P. Cudré-Mauroux, J. Heflin, E. Sirin, T. Tudorache, J. Euzenat, M. Hauswirth, ... E. Blomqvist (Eds.), *International Semantic Web Conference 2* (pp. 1–16). Springer.
- Bannister, F., & Connolly, R. (2011). The Trouble with Transparency: A Critical Review of Openness in e-Government. *Policy & Internet*, 3(1), Article 8. doi:10.2202/1944-2866.1076
- Barry, E., & Bannister, F. (2013). Barriers to Open Data Release: A View from the Top. In *EGPA 2013*.
- Bates, J. (2012). “This is what modern deregulation looks like”: co-optation and contestation in the shaping of the UK’s Open Government Data Initiative. *The Journal of Community Informatics*, 8(2).
- Batley, R. (1994). The consolidation of adjustment: Implications for public administration. *Public Administration and Development*, 14(5), 489–505. doi:10.1002/pad.4230140505
- Blakemore, M., & Craglia, M. (2006). Access to Public-Sector Information in Europe: Policy, Rights, and Obligations. *The Information Society*, 22(1), 13–24. doi:10.1080/01972240500388180
- Cerrillo-i-Martínez, A. (2011). The regulation of diffusion of public sector information via electronic means: Lessons from the Spanish regulation. *Government Information Quarterly*, 28(2), 188–199. doi:10.1016/j.giq.2010.05.009
- Cerrillo-i-Martínez, A. (2012). The Reuse of Public Sector Information in Europe and Its Impact on Transparency. *European Law Journal*, 18(6), 770–792.
- Colomer, J. (1998). The Spanish “State of Autonomies”: Non-Institutional Federalism. *West European Politics*, 21(4), 40–52.
- Conradie, P., & Choenni, S. (2012). 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 - ICEGOV '12* (pp. 5–13). New York, New York, USA: ACM Press. doi:10.1145/2463728.2463731
- Courmont, A. (2012). How to Govern Open Data? The politics of open data portals. In *Paper presented at the IPSA Conference, Madrid, July 8-12, 2012*.
- Dahrendorf, R. (2006). *Homo Sociologicus. Ein Versuch zur Geschichte, Bedeutung und Kritik der Kategorie der sozialen Rolle* (16. Auflag.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Davies, T. (2013). *Open Data Barometer. 2013 Global Report* (p. 45). London.
- Davies, T., & Bawa, Z. (2012). The Promises and Perils of Open Government Data (OGD). *The Journal of Community Informatics*, 8(2).

- Dawes, S. S., Cresswell, A. M., Pardo, T. A. (2009). From “Need to Know” to “Need to Share”: Tangled Problems, Information Boundaries, and the Building of Public Sector Knowledge Networks. *Public Management Review*, 69(3), 392–402.
- Deloitte. (2012). *Open data. Driving growth, ingenuity and innovation*. London.
- Ding, L., Lebo, T., Erickson, J. S., DiFranzo, D., Williams, G. T., Li, X., ... Hendler, J. a. (2011). TWC LOGD: A portal for linked open government data ecosystems. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(3), 325–333. doi:10.1016/j.websem.2011.06.002
- Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *The Academy of Management Review*, 20(1), 65–91.
- Dutton, W. H. (2008). Social Movements Shaping the Internet: The Outcome of an Ecology of Game. In M. S. Elliot & K. L. Kraemer (Eds.), *Computerization Movements and Technology Diffusion. From Mainframes to Ubiquitous Computing* (pp. 499–518). Medford, NJ: Information Today.
- European Parliament, & Council of the European Union. Directive 2003/98/EC of the European Parliament and of the Council (2013).
- Flak, L. S., & Rose, J. (2005). Stakeholder Governance: Adapting Stakeholder Theory to E-Government. *Communications of the Association for Information Systems*, 16, 642–664.
- Garcia, E. G., & Soriano Maldonado, S. L. (2012). Reuse of Public Sector Information in Spain: A solid foundation for a promising future. *Textos Universitaris de Biblioteconomia I Documentació*, (29).
- Geiger, C. P., & von Lucke, J. (2012). Open Government and (Linked) (Open) (Government) (Data). *eJournal of eDemocracy & Open Government*, 4(2), 265–278.
- Gurstein, M. (2014). Open Data: Empowering the empowered or effective data use for everyone. *First Monday*, 16(2).
- Haque, M. (2001). The Diminishing Publicness of Public Service under the Current Mode of Governance. *Public Administration Review*, 61(1), 65–82.
- Harrison, T. M., Pardo, T. a., & Cook, M. (2012). Creating Open Government Ecosystems: A Research and Development Agenda. *Future Internet*, 4(4), 900–928. doi:10.3390/fi4040900
- Hausenblas, M., & Karnstedt, M. (2010). Understanding Linked Open Data as a Web-Scale Database. *2010 Second International Conference on Advances in Databases, Knowledge, and Data Applications*, 56–61. doi:10.1109/DBKDA.2010.23
- Heimstädt, M., Saunderson, F., & Heath, T. (2014). Conceptualizing Open Data ecosystems: A timeline analysis of Open Data development in the UK. In *Proceedings of the International Conference for E-Democracy and Open Government (CeDEM2014), Krems, Austria*. (p. forthcoming).
- Huijboom, N., & van den Broek, T. (2011). Open data: an international comparison of strategies. *European Journal of ePractice*, 12(March/April), 1–13.



- Hunnius, S., & Krieger, B. (2014). The Social Shaping of Open Data through Administrative Processes. In *Proceedings of the 10th International Symposium on Open Collaboration (OpenSym 2014)* (p. forthcoming). ACM Press. doi:<http://dx.doi.org/10.1145/2641580.2641601>
- Hunnius, S., Krieger, B., & Schuppan, T. (2014). Providing, Guarding, Shielding: Open Government Data in Spain and Germany. In *2014 EGPA Annual Conference, 10-12 September 2014 in Speyer, Germany*.
- Hunnius, S., & Schuppan, T. (2011). Mergers of public ICT service providers. In *XVII IRPSM* (pp. 1–20).
- Hunnius, S., Schuppan, T., & Stocksmeier, D. (2014). Institutionelle Gestaltungsmerkmale der staatlichen IT Governance in Deutschland. In T. Gordon, D. Lück-Schneider, J. von Lucke, M. Wimmer, S. Kaiser, & E. Schweighofer (Eds.), *Gemeinsam Electronic Government (ziel)gruppengerecht gestalten und organisieren. Gemeinsame Fachtagung Verwaltungsinformatik (FTVI) und Fachtagung Rechtsinformatik (FTRI) 2014. LNI* (pp. 14–26). Bonn: Köllen.
- Hyland, B., & Wood, D. (2011). The Joy of Data - A Cookbook for Publishing Linked Government Data on the Web. In D. Wood (Ed.), *Linking Government Data* (pp. 3–26). New York: Springer.
- Immonen, A., Palviainen, M., & Ovaska, E. (2014). Towards open data based business: Survey on usage of open data in digital services. *International Journal of Research in Business and Technology*, 4(1), 286–296.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446–456. doi:[10.1016/j.giq.2011.01.004](https://doi.org/10.1016/j.giq.2011.01.004)
- Janssen, K. (2012). Open Government Data and the Right to Information: Opportunities and Obstacles. *The Journal of Community Informatics*, 8(2).
- Janssen, K., & Dumortier, J. (2003). Towards a European Framework for the Re-use of Public Sector Information: a Long and Winding Road. *International Journal of Law and Information Technology*, 11(2), 184–201. doi:[10.1093/ijlit/11.2.184](https://doi.org/10.1093/ijlit/11.2.184)
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268. doi:[10.1080/10580530.2012.716740](https://doi.org/10.1080/10580530.2012.716740)
- Jones, T., & Wicks, A. (1999). Convergent Stakeholder Theory. *Academy of Management Review*, 24(2), 206–221.
- Julien, N. (2012). *Business Opportunities Arising from Open Data Policies*. Imperial College London.
- Kickert, W. J. M., Klijn, E.-H., & Koppenjan, J. F. M. (1997). *Managing complex networks: strategies for the public sector*. Sage.
- King, R. D., Liakata, M., Lu, C., Oliver, S. G., & Soldatova, L. N. (2011). On the formalization and reuse of scientific research. *Journal of the Royal Society, Interface*, 8(63), 1440–8. doi:[10.1098/rsif.2011.0029](https://doi.org/10.1098/rsif.2011.0029)



- Landsbergen Jr., D., & Wolken Jr., G. (2001). Realizing the Promise: Government Information Systems and the Fourth Generation of Information Technology. *Public Administration Review*, 61(2), 206–220.
- Lee, G., & Kwak, Y. H. (2012). An Open Government Maturity Model for social media-based public engagement. *Government Information Quarterly*, 29(4), 492–503.
- Lijphart, A. (1999). *Patterns of Democracy*. Yale University Press.
- Martin, M., Kaltenböck, M., Nagy, H., & Auer, S. (2011). The Open Government Data Stakeholder Survey. *OKCon*.
- Martin, S., Foulonneau, M., Turki, S., & Ihadjadene, M. (2013). Risk Analysis to Overcome Barriers to Open Data. *Electronic Journal of E-Government*, 11(1), 348–359.
- Mayer-Schönberger, V., & Zappia, Z. (2011). Participation and Power: Intermediaries of Open Data. In *1st Berlin Symposium on Internet and Society October 26th – 28th 2011*.
- McKinsey Global Institute. (2013). *Open data: Unlocking innovation and performance with liquid information*.
- Meijer, A. J., & Thaens, M. (2009). Public information strategies: Making government information available to citizens. *Information Polity*, 14(1-2), 31–45. doi:10.3233/IP-2009-0167
- Ministry of the Presidency. Royal Decree 1495/2011, of 24 October 2011 (2011).
- Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *The Academy of Management Review*, 22(4), 853–886. doi:10.2307/259247
- Muñoz-Cañavate, A., & Hípola, P. (2011). Electronic administration in Spain: From its beginnings to the present. *Government Information Quarterly*, 28(1), 74–90. doi:DOI: 10.1016/j.giq.2010.05.008
- Napoli, P. M., & Karaganis, J. (2010). On making public policy with publicly available data: The case of U.S. communications policymaking. *Government Information Quarterly*, 27(4), 384–391.
- Nardi, B. A., & O’Day, V. L. (1999). *Information Ecologies. Using Technology with Heart*. Cambridge, MA: MIT Press.
- O’Riain, S., Curry, E., & Harth, A. (2012). XBRL and open data for global financial ecosystems: A linked data approach. *International Journal of Accounting Information Systems*, 13(2), 141–162. doi:10.1016/j.accinf.2012.02.002
- O’Toole, L. J. (1997). Treating Networks Seriously: Practical and Research-Based Agendas in Public Administration. *Public Administration Review*, 57(1), 45–52.
- Open Knowledge Foundation. (2013). Open Data Index – Published October 28th, 2013. Retrieved July 01, 2014, from <https://index.okfn.org/country>



- Owen, B. B., Cooke, L., & Matthews, G. (2013). The development of UK government policy on citizens' access to public sector information. *Information Polity*, 18(1), 5–19. doi:10.3233/IP-2012-000286
- Pardo, T. A., Cresswell, A. M., Thompson, F., & Zhang, J. (2006). Knowledge sharing in cross-boundary information system development in the public sector. *Information Technology and Management*, 7(4), 293–313. doi:10.1007/s10799-006-0278-6
- Phillips, R., Freeman, R. E., & Wicks, A. C. (2003). What Stakeholder Theory is Not. *Business Ethics Quarterly*, 13(4), 479–502.
- Pollitt, C., & Bouckaert, G. (2011). *Public Management Reform. A comparative analysis* (3rd editio., p. 352pp.). Oxford: Oxford University Press Oxford.
- Proyecto Aporta. (2011). *Characterization Study of the Infomediary Sector* (p. 88).
- Proyecto Aporta. (2012). *Estudio de caracterización del Sector Infomediario en España* (p. 126).
- Püschel, J.-O. (2006). Vom Informationszugang zur Informationsweiterverwendung. *Datenschutz Und Datensicherheit*, 30(8), 481–489.
- Romzek, B. S. (2000). Dynamics of Public Sector Accountability in an Era of Reform. *International Review of Administrative Sciences*, 66(1), 21–44. doi:10.1177/0020852300661004
- Scharpf, F. W. (1988). The Joint-Decision Trap: Lessons from German Federalism and European Integration. *Public Administration*, 66(3), 239–278.
- Scharpf, F. W. (2005). *No Exit from the Joint Decision Trap? Can German Federalism Reform Itself?* Köln.
- Scholl, H. J. (2001). Applying stakeholder theory to e-government: benefits and limits. (B. Schmid, K. Stanoevska-Slabeva, & V. Tschammer, Eds.) *1st IFIP Conference on E-Commerce, E-Business, and E-Government (I3E 2001)*. Zurich, Switzerland: Kluwer.
- Scholl, H. J., Kubicek, H., Cimander, R., & Klischewski, R. (2012). Process integration, information sharing, and system interoperability in government: A comparative case analysis. *Government Information Quarterly*, 29(3), 313–323. doi:10.1016/j.giq.2012.02.009
- Shadbolt, N., & O'Hara, K. (2013). Linked data in government. *IEEE Internet Computing*, (July/August 2013).
- Shkabatur, J. (2012). Transparency With(out) Accountability: Open Government in the United States. *Yale Law & Policy Review*, 31(1), 79–140.
- Tan, C.-W., Pan, S. L., & Lim, E. T. (2005). Managing Stakeholder Interests in e-Government Implementation: Lessons Learned from. *Journal of Global Information Management*, 13(1), 31–53.
- Tennert, J. R., & Schroeder, A. D. (1999). Stakeholder analysis. *60th Annual Meeting of the American Society for Public Administration*. Orlando, FL.



- Van den Broek, T., van Veenstra, A. F., & Folmer, E. (2011). Walking the extra byte: A lifecycle model for linked open data. In E. Folmer, M. Reuvers, & W. Quak (Eds.), *Linked Open Data – Pilot Linked Open Data Nederland* (pp. 95–111). Amersfort: Remwerk.
- Vickery, G., & Wunsch-Vincent, S. (2006). *Digital broadband content: public sector information and content*. Paris.
- Villazón-Terrazas, B., Vilches-Blazquez, L. M., Corcho, O., & Gomez-Perez, A. (2011). Methodological guidelines for publishing government linked data. In *Linking Government Data* (pp. 27–49). New York: Springer.
- Willem, A., & Buelens, M. (2007). Knowledge Sharing in Public Sector Organizations: The Effect of Organizational Characteristics on Interdepartmental Knowledge Sharing. *Journal of Public Administration Research and Theory*, 17(4), 581–606. doi:10.1093/jopart/mul021
- Wirtz, H. (2014). Die Änderung der PSI-Richtlinie - Fort- oder Rückschritt? *Datenschutz Und Datensicherheit*, 38(6), 389–393.
- Yang, T.-M. (2012). The Complexity of Cross-Boundary Information Sharing: An Organizational Perspective on Taiwan E-Government. In *ICEGOV '12, October 22 - 25 2012, Albany, NY, USA* (pp. 143–146).
- Yang, T.-M. (2011). Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors. *Government Information Quarterly*, 28(2), 164–175.
- Yu, H., & Robinson, D. (2012). The New Ambiguity of 'Open Government'. *UCLA Law Review Discourse*, 178, 180–208.
- Zhang, J., Dawes, S. S., & Sarkis, J. (2005). Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing. *Journal of Enterprise Information Management*, 18(5), 548–567. doi:10.1108/17410390510624007
- Zuiderwijk, A., & Janssen, M. (2012). A comparison of open data policies and their implementation in two Dutch ministries. In *Proceedings of the 13th Annual International Conference on Digital Government Research* (pp. 84–89).
- Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17–29. doi:10.1016/j.giq.2013.04.003
- Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of E-Government*, 10(2), 156–172.
- Zuiderwijk, A., Jeffery, K., & Janssen, M. (2012). The potential of metadata for linked open data and its value for users and publishers. *Journal of E-Democracy and Open Government*, 4(2), 222–244.