STATE OF THE ART REVIEW OF OPEN DATA RESEARCH: INSIGHTS FROM EXISTING LITERATURE AND A RESEARCH AGENDA

Mohammad Alamgir Hossain
School of Business, IT, and Logistics
College of Business
RMIT University, Melbourne City Campus
Melbourne, Victoria, Australia
Email: mahripon@yahoo.com

Yogesh K Dwivedi (Corresponding Author)
Swansea University Bay Campus
School of Management
Fabian Way, Crymlyn Burrows
SWANSEA, SA1 8EN
Wales, UK
Tel: +44(0)1792602340
Email: ykdwivedi@gmail.com

Nripendra P. Rana
Swansea University Bay Campus
School of Management
Fabian Way, Crymlyn Burrows
SWANSEA, SA1 8EN
Wales, UK
Email: nrananp@gmail.com

Abstract

With the proliferation of mobile network, mobile devices, and Web of things, many different industries including government departments, private firms, and research communities offer more transparency through releasing data. The resultant effort offers a new paradigm - *open data* - still at infancy stage though. Despite the rising research initiatives explaining its benefits and challenges, and demonstrating policy conception and project details, no systematic survey of extant literature on *open data* is performed yet. Such a study could examine *open data* from a holistic canvas, assess the current status of research and propose future direction. This study conducts a review of the extant literature in order to ascertain the current state of research on *open data* and present an extensive exploration for eleven different types of analyses: contexts, perspectives, level of analysis, research methods, the drivers, benefits, barriers, theory/model development, the most productive journals, authors, and institutions. Also, we present a number of future research agendas. This study also explains the implications to assist researchers, policy makers and journal editors.

Keywords: open data, open government data, literature review, research agenda

1. INTRODUCTION

Our modern society and economy is moving towards a knowledge-based and service-oriented global world (Rohunen et al. 2014). Data is considered as fundamental prerequisite to gain knowledge and to produce services (Shadbolt et al. 2012). Hence, there is argument that data should not be captive but set free. Three recent developments actually catalysed such demand: citizens' sense of ownership in governance and politicians' disposition to decentralise civic services (Cerrillo-i-Martínez 2012; Zuiderwijk et al. 2014a); recent technological developments that has facilitated developments of peoples' computing skills in terms of accessing, storing, manipulating, analysing, linking and distributing data and information (Boulton 2014; Boulton et al. 2011; Rohunen et al. 2014); and the massive growth of mobile network resulting substantial rise of virtual social networks (Huijboom and Van den Broek 2011). Technological advancements using Web 2.0 also made it possible to produce interoperable data (Kulk and Van Loenen 2012); which made some citizens to change their role from passive recipients to active producer and user of data simultaneously. Opposing the theory of diminishing utility, the value of data increases with its reuse (Meijer et al. 2014; Tananbaum 2008); in fact, the products of data are sometimes more attractive than the original data (Serra 2014). Consequently, "there is a growing need for, and growth of open data" (Hrynaszkiewicz 2011, p.31).

Open data is a 'philosophy' or 'strategy' that encourages mostly public organisations to release objective, factual, and non-person-specific data that is generated or collected through the delivery of public services, to anyone, with a possibility of further operation and integration, without any copyright restrictions (Bertot et al. 2014; Bichard and Knight 2012; Hrynaszkiewicz 2011; Kassen 2013). Open data comprise various kinds of data – primary (census data) or secondary (economic trend), real-time (such as traffic or weather data) or offline (government spending), location-based (toxic waste dumps) or generic (regional healthcare costs), reports, maps, satellite photographs, pictures and paintings, the genome, medical data, scientific formula, public sector budgeting, food-safety information, and so forth (Rohunen et al. 2014; Mazumder 2014; Hendler et al. 2012) .

The concept of *open data* received popularity and came into the public domain by President Barrack Obama in 2009, which was followed by the UK government's transparency initiative in 2011 (Meijer et al. 2014). Such calls initiated regional, national, and even municipal agencies to join 'the next big

thing'. But, open data concept is older than the US and UK's current movement; Boulton (2014) claims that, open data principles were established in early years of the European Enlightenment (p. 133). However, technologically, it was established in last few decades, endorsed and proposed by Sir Tim Berners-Lee (Gurstein 2011; Pabón et al. 2013).

Now, many countries mandated public departments and agencies to release data without any further copyright obligation to reuse or distribute. It should be underscored that it is not just about releasing data in the form of machine-readable datasets over Internet but it is essential to release it with the functionality to collate, combine and enhance to provide a broader and/or innovative services (Bichard and Knight 2012). Although most open data initiatives are observed in public sectors, *open data* is not exclusive to 'open government' but to other fields too, including science, culture, economics, or libraries (Uhlir and Schröder 2007). Therefore, any business entity can add value to *open data* and can generate revenue through disseminating the data in a new form (Cerrillo-i-Martínez 2012).

Open data is becoming important in research too. Estermann (2014) claimed that, *open data* movement in academic circles has been started 50 years ago; the first scientific journal (*Philosophical Transactions of the Royal Society*) published in 1965 had the policy of establishing concepts accompanied by the evidence (i.e. data) on which it was based (Boulton 2014). Sharing research data avoids the effort for 'reinventing the wheel', provides evidence that the research methodology was correct and rightly applied, shows accountability of the researcher, and creates the possibility of generation's new findings (Gurstein 2011), not noticed by the initial researcher(s) (Hester 2014). Many journals, especially scientific journals, are in favour of disclosing experimental data [1] (Tananbaum 2008) so that they can be reused, replicated, and verified. Similarly, opening research data is now a requirement to plan data management and release policy of research application (Childs et al. 2014). Some journals ask the consent that data will be provided on request (Hrynaszkiewicz 2011) while few others consider it as a must.

Considering the political, socioeconomic and scholarly significance and possibility of its wider application, the topic has started attracting interest of research community, which is evident through

[1] The raw bits of information from where conclusions of an experiment are derived.

emerging body of knowledge as published papers and active discussion about it during academic conferences. These activities have experienced an exponential growth in recent times in the form of research scholarships and projects.

As this contemporary domain is still at infancy stage for further development of research in this area, it would be useful to explore current state of research on the topic and provide guidance for the further lines of research. Some initial attempts (Zuiderwijk et al. 2014ab) have already been made to review research on this topic. Zuiderwijk et al. (2014a) aimed at contributing to research on open data transparency and policies. The authors discussed definitions, developments, research, challenges and barriers linked to open data transparency and policies. The study concluded the assumptions related to research and policy reports, which indicate that open data can be used as a tool to improve transparency. The research also showed that although open data policies are increasingly developed, they are barely able to guide the publication and use of open data (Zuiderwijk et al. 2014a). A study by Zuiderwijk et al. (2014b) analysed existing publications to explore the relationship between innovation and open data. The analysis of the open data publications indicated that most of the papers in this area are primarily conceptual in nature with the descriptions of the empirical uses of open data or design of technology and systems. The authors also found that limited attention has been given for the theory development (Zuiderwijk et al. 2014b). These two initial review attempts (Zuiderwijk et al., 2014a; Zuiderwijk et al., 2014b) although have already provided important understanding of the development of this topic, they both were rather focussed on specific aspects of the open data concept. A general review that is more holistic in coverage can provide further insights on the current state of research as well as the overall development of the topic, which can form the basis for shaping future research.

This study, therefore, aims to explore and synthesise extant literature on open data by focusing on three high-level agendas (See Table 1). The first agenda is about the intellectual progress/substance in open data – the action or processes the researchers applied in order to acquire knowledge and to understand the topic through experiment and their own and others' experience and thoughts. Then, the second agenda is to understand open data in detail by examining the factors such as the drivers, benefits, impediments, and theoretical development. The final agenda is related to acknowledging the

contribution of the knowledge society by identifying the leading journals, more prolific authors, and institutions (Li and Zhang 2005).

Table 1 Research questions on state-of-the-arts open data

RQ1	What constitutes the intellectual substances?
	RQ1.1 What are the contexts of the studies? RQ1.2 What are the perspectives the studies examined? RQ 1.3 What different levels of analyses the articles covered? RQ 1.4 What are the research methods the articles applied?
RQ2	What progress is made in research to understand open data?
	RQ2.1 What are the drivers of open data initiatives the studies identified? RQ2.2 What are the benefits/opportunities of open data the studies explored? RQ2.3 What are the associated impediments/barriers to adopt open data? RQ2.4 What theoretical models the current studies used/identified?
RQ3	Who are the most active members of the open data knowledge-society?
	RQ3.1 What are the top journals that published the most papers on open data? RQ3.2 Who are the most prolific authors? RQ4.2 What are the most prolific institutions for open data research?

As the main theoretical contribution, this study is one of the first reported comprehensive review works on *open data* research. The practical implications target policymakers to resolve political issues, technologists to work on the technical and technological barriers and researchers to get research agendas while journal editors can generate themes for journal special issues or conference tracks.

The rest of the paper goes along as follows. First, the methodology applied for this study is explained. Then, in the *analyses and results* section (Section 3) the detail examination of the extant literature has been presented. The following section (Section 4) then presents agenda for the future research. Finally, section 5 provides concluding remarks and briefly outlines some of the limitations of this review work.

2. METHODOLOGY

In this study, we performed extensive content analysis as well as descriptive analyses along with explaining the philosophical issues - a frequently utilised approach in information systems (IS) area (Avison et al. 2008; Dwivedi and Kuljis 2009; Dwivedi et al. 2011; Li and Zhang 2005; Weerakkody et al. 2009). The review was conducted by coding relevant elements (such as topics and methods of published articles in the discipline), which were then grouped into more general/high level concept.

In order to select the papers for analyses, we applied a two-stage process. First, a comprehensive

search using a combination of a descriptor, "open data" was realised within the following databases: SCOPUS, Emerald, ProQuest, EBSCOhost (Business Source Complete), and Science Direct. This search produced overwhelming number of records. We considered only the papers that have both keywords (i.e. Open and Data) in the title, in any order. The second stage was more complex. We mainly focused on journal papers believing that journals along with conference proceedings and book-chapters. Afterward, we performed a screening process by skimming the papers, in order to assess its fit with our research objectives. For instance, many papers used both terms but did not mean open data research (e.g. open library data, linguistic linked open data). Moreover, many papers presented studies from the domain beyond IS (e.g. chemistry, biology) – these papers were excluded. Then, the papers that demonstrated hard-core technical issues (e.g. developing algorithms) were excluded. In this process, the editorials, opinions, perspectives too were excluded from analysis. That resulted as a set of 96 relevant papers for analyses.

Selecting the timeline of the papers was a difficult task. The history of *open data* is not too old. Yet, a reasonable number of papers have been found from its inception. In order to reach to an agreed timeline, we decided to include every paper (satisfying our criteria presented earlier) until November 2014. The first paper is reported in 1996, thus resulting 14 years of analysis. It can be considered as representative of the current literature given that the topic and research domain is comparatively new.

The authors had several *Google Hangout* sessions for discussions to determine classification themes. Taking the example of some scholarly papers (Li and Zhang 2005; Weerakkody et al. 2009), we selected the analysis criteria. Then, the articles were examined, evaluated, and coded by the authors individually. Later, the works were compared and amended where discrepancy occurred.

3. ANALYSES AND RESULTS

To present the results of our analysis, the first column in Table 2 ('paper ID') will be used in the subsequent tables as the reference to the related article. From Table 2, it can be derived that overall, two-third of the studies applied empirical methodology. Then, most studies took context from EU region. Furthermore, Open Data is often studied along with other topics such as: open government data (63%), linked open data (13%), big data (8%), PSI (8%), and open systems (5%).

3.1 RQ1: What constitutes the intellectual substances?

3.1.1 Classification of the contexts

Context refers to the setting or environment where a study is conducted that may include organisational, social or marketplace (Zhang and Li 2005). *Open data* studies, from IS background, are conducted mainly in two settings: public and private[2]. It is observed from Table 3 that the both contexts are almost equally investigated with a little more focus on public sector setting.

Within *public* context, studies investigated in overall management issues to policy analysis to application in various service deliveries. Governments of at least 50 countries now are the main proponents of *open data,* realising its value in governance and transparency (Davies et al. 2013). They increasingly promote *open data* initiatives by mandating the departments to release data, and inspire citizens and private firms to reuse such data (Rohunen et al. 2014). Such movements have been initiated by PSI and OGD initiatives in the US and European Union, respectively. Hielkema and Hongisto (2013) presented a case where the municipality of Helsinki successfully arranged open competitions for mobile applications using open data, with further prospect of commercialisation. In contrast, still several government initiatives achieved minimal success where agencies adopted a passive-aggressive attitude toward OGD (Janssen 2011). They "did the minimal required to qualify as a participating agency on Open Data move and shortly stopped cooperating with the Open Data program" (Peled 2011, p.2088).

Similarly, the studies conducted in private contexts examined in research and development (R&D), and how open data increases organisational performance, the barriers toward obtaining and using Open Data, and the ergonomics of the users. From an economic view, Whitmore (2014) found that open data may enhance firm business especially that are military contractors. He evidenced that before any USA military attack, government spending on military contracts increase 'largely'; therefore, firms prepare accordingly when data are open. But Fleisher (2008) demonstrated that marketing managers face several issues while using open data for developing competitive and marketing intelligence.

[2] Science or IT studies in *open data* not necessarily follow contexts in the sense we termed. For instance, explaining the overall technical barriers or offering an architectural solution not necessarily studies a specific context. The papers without contexts are excluded for this analysis.

Table 2 The overall classification of the reviewed papers

Paper ID	Reference	Context/sample	Co-studied topic	Methodology	Paper ID	Reference	Country of context/sample	Co-studied topic	Methodology
1	Cláudio (2013)		OGD	Non Empirical	49	Lassinantti et al. (2014)	Ireland	OGD, PSI	Empirical
2	Claudio and Reinhard (2014)	Brazil	OGD	Empirical	50	Linders (2013)	Finland	OGD	Non Empirical
3	Alexopoulos et al. (2013)	Greece	OGD, PSI	Empirical	51	Lindman (2014)			Empirical
4	Alexopoulos et al. (2014)	Netherlands, Greece		Empirical	52	MacDonald and Martinez (2008)			Non Empirical
5	Andreoli-Versbach and Mueller-Langer (2013)			Empirical	53	Martin et al. (2013)	France, Germany, UK		Empirical
6	Barry and Bannister (2014)	Ireland		Empirical	54	McDonald and Léveillé (2014)		BD	Empirical
7	Bates (2014)	UK	OGD	Empirical	55	McLeod (2012)			Empirical
8	Behkamal et al. (2014)		LOD	Empirical	56	Meijer et al. (2014)	Netherlands		Empirical
9	Bertot et al. (2014)	US	BD	Empirical	57	Murray-Rust (2008)			Non Empirical
10	Bichard and Knight (2012)	UK		Empirical	58	Nugroho et al. (2015)			Empirical
11	Borglund et al. (2014)	Sweden		Empirical	59	O'Hara (2012)	UK		Non Empirical
12	Boulton (2014)			Non Empirical	60	Pabón et al. (2013)		LOD	Non Empirical
13	Boulton et al. (2011)			Non Empirical	61	Peled (2011)	USA	OGD	Empirical
14	Braunschweig et al. (2012)	US, UK, Kenya, Worldbank, UN,		Empirical	62	Reed (2013)			Non Empirical
15	Casellas (2014)	Spain		Empirical	63	Robinson et al. (2012)			Empirical
16	Cerrillo-i-Martínez (2012)	EU		Non Empirical	64	Rohunen et al. (2014)	Finland		Empirical
17	Charalabidis et al. (2014)		OGD	Empirical	65	Rowen et al. (2000)			Empirical
18	Chan (2013)	Singapore	OGD	Empirical	66	Saxby and Hill (2012)	UK		Non Empirical
19	Childs et al. (2014)	UK		Empirical	67	Sayogo and Pardo (2013)			Empirical
20	Conradie and Choenni (2014)	Netherlands	PSI	Empirical	68	Sayogo et al. (2014)	NAFTA region		Empirical
21	Craveiro, Santana, and	Brazil	OGD	Empirical	69	Shadbolt et al. (2012)	UK	OGD	Empirical

	Albuquerque (2013)								
22	Davies (2011)	UK	OGD	Empirical	70	Simón et al. (2014)			Non Empirical
23	Davies et al. (2013)	Developing countries		Non Empirical	71	Stephens (2011)			Empirical
24	Dulong de Rosnay and Janssen (2014)	EU		Non Empirical	72	Streeter et al. (1996)	US, France		Empirical
25	Eckartz, Hofman, and Van Veenstra (2014)	Netherlands		Empirical	73	Tananbaum (2008)			Non Empirical
26	Estermann (2014)	Switzerland		Empirical	74	Thurston (2012)			Non Empirical
27	Ferro and Osella (2012)		PSI	Non Empirical	75	Tolbert and Mossberger (2006)			Empirical
28	Ferro and Osella (2013)		PSI	Non Empirical	76	Tsiavos, Stefaneas, and Karounos (2013)	Greece	PSI	Empirical
29	Fleisher (2008)			Non Empirical	77	Uhlir and Schröder (2007)			Empirical
30	Garbett, Linehan, Kirman, Wardman, and Lawson (2010)	UK		Empirical	78	Veljković et al. (2014)		OGD	Non Empirical
31	Hendler et al. (2012)	US	OGD	Empirical	79	Whitmore (2014)	US	OGD	Empirical
32	Hielkema and Hongisto (2013)	Finland		Empirical	80	Yannoukakou and Araka (2014)			Empirical
33	Huijboom and Van den Broek (2011)	Netherlands		Non Empirical	81	Yoose and Perkins (2013)		LOD	Empirical
34	Hunnius, Krieger, and Schuppan (2014)	Spain, Germany		Empirical	82	Zotti and La Mantia (2014)	EU	LOD, BD	Non Empirical
35	Ivanov, Varga, and Pejić Bach (2014)		OGD	Non Empirical	83	Zuiderwijk and Janssen (2012)	Netherlands		Non Empirical
36	Janssen (2011)	EU	OGD, PSI	Empirical	84	Zuiderwijk et al. (2012c)			Non Empirical
37	Janssen (2012)		OGD	Non Empirical	85	Zuiderwijk et al. (2012c)			Empirical
38	Janssen et al. (2012)		OGD	Empirical	86	Zuiderwijk et al. (2012d)	Netherlands	OGD	Empirical
39	Janssen and Zuiderwijk (2012)	Netherlands		Empirical	87	Zuiderwijk et al. (2012a)			Empirical
40	Jetzek, Avital, and Bjørn-Andersen (2012)		OGD	Non Empirical	88	Zuiderwijk et al. (2012b)		LOD	Empirical
41	Jetzek et al. (2013)	Denmark	OGD	Empirical	89	Zuiderwijk and Janssen (2013)			Non Empirical

42	Kaasenbrood et al. (2015)	Netherlands	OGD	Empirical	90	Zuiderwijk et al. (2013a)			Empirical
43	Kalampokis et al. (2011a)		OGD	Non Empirical	91	Zuiderwijk et al. (2013b)	EU		Non Empirical
44	Kalampokis et al. (2011b)			Non Empirical	92	Zuiderwijk and Janssen (2014a)			Empirical
45	Kassen (2013)	US		Empirical	93	Zuiderwijk and Janssen (2014b)			Empirical
46	Krotoski (2012)			Empirical	94	Zuiderwijk and Janssen (2014c)	Netherlands	OGD	Empirical
47	Kuk and Davies (2011)	UK		Empirical	95	Zuiderwijk et al. (2014c)	Netherlands		Empirical
48	Kulk and Van Loenen (2012)			Non Empirical	96	Zuiderwijk et al. (2014d)			Non Empirical

[Legend: BD: Big Data, LOD: Linked Open Data, OGD: Open Government Directive, PSI: Public Sector Information]

Likewise, adoption of open data in libraries or museums is not substantial despite of several initiatives (Yoose and Perkins 2013). Yet, some studies dealt with overall or holistic nature of Open Data; for instance, Janssen et al. (2012) demonstrated the benefits, barriers, and myths of Open Data and related them with two theories. Some studies examined open data issues from mixed context. For instance MacDonald and Martinez (2008) found that ICTs are transforming academic field offering more collaboration in research, which produce prodigious' data and its use. In order to handle such "data delude", an integrated approach involving researchers, librarians, technologists, publishers, and policymakers is essential (McLeod 2012).

Table 3 Context classification (n=84)

Contexts	Sub-contexts	References	Count (%)
Government			43 (51%)
	State/Municipal initiatives	3, 15, 32, 45	4
	Service departments including council facilities, road and traffic etc. – census	21, 31, 49, 54, 58, 69, 81, 86	8
	Examining legal framework	16, 48, 76	3
	Policy/project analysis	7, 9, 33, 36, 54, 89, 94, 80, 66	9
	Examining infrastructure, innovation strategies	17, 18, 22, 88, 91	5
	Others: overall management, governance, value-creation	34, 35, 41, 44, 56, 61, 78	7
	Understanding/development of the (internal) process/mechanism of data publication and/or use, and add value to data	1, 6, 20, 39, 43, 89, 95	7
Public			26 (31%)
	Possibilities/issues in research and development	5, 13, 19, 67	4
	Realizing productivity and challenge	2, 29, 30, 70	4
	Impacts on business	51, 72, 79	3
	Holistic/overview	24, 38, 93, 85, 96	5
	Library, heritage, NGO, archiving, earth observation	11, 26, 54, 81, 82	5
	Issues - adoption, development	42, 68, 50	3
	Street, traffic data	63, 64	2
Mixed	Government as well as private organisations	14, 25, 37, 41, 55, 84, 90, 92	8 (10%)
General/not defined		4, 8, 27, 47, 53, 77, 74	7 (8%)

3.1.2 Classification of the perspectives

Researchers used different perspectives to explain *open data* movement. Most studies combined different perspectives while a few concentrated on one. For instance, Lassinantti et al. (2014) examined local government *open data* initiatives from institutional, social, technical, and economic perspectives; Janssen et al. (2012) examined from political and social, economic, operational and technical, and legal perspectives; while Barry and Bannister (2014) from the perspective of senior

managers applying technical, social, legal, economic; and MacDonald and Martinez (2008) are a proponent of technological, cultural, and legal perspectives. To analyse the 'perspectives' used in open data literature we adapted Zuiderwijk et al. (2014b)'s seven perspectives. Table 4 illustrate that the studies examined the perspectives almost equally, with more focus on technical or combination of perspectives.

The **technical**[3] perspective explores *open data* from IT and overall technical perspectives. As we excluded purely technical papers from our analysis, the technical issues naturally in our study are mainly behavioural and perception-based. Some studies found that *open data* became a practical philosophy due to technical developments because technology simplified the access to data, analyses of data, and understanding of data using data visualisation. However, technology can be a curse too for *open data* because, by using appropriate technology personal data can be identified by linking data from different datasets (Childs et al. 2014; Floridi 2014). Studies also explained the technical details of open data including metadata, standardisation, interoperability, etc.

Then, papers on **political** perspective deal with the importance and nature of political influence behind open data. Throughout the world, government departments consider open data projects mostly as a political move as was president Obama's "retreat from openness". His OGD declaration instructed U.S. agencies to publish at least three high-value datasets on the Web. However, many agencies satisfied the 'transparency' commitment (sometimes by releasing less quality, useless, unstructured or non-reusable data) but did not participate on the extension; only 5 out of 169 participating accounted for 99.4% of the released dataset on www.data.gov (Peled 2011). The reason for not participating is political too, but by the bureaucrats. Each government-agency and department has own agenda competing over resource, recognition, influence and control, and autonomy over others; they consider their datasets as survival weapon that they would not release in the name of transparency. Hence, agencies would wait till the end before sharing or releasing (important) data. Still, open data is a means for people's empowerment on government (Tolbert and Mossberger 2006).

[3] There can be difference between technical and technological perspectives – we did not initiate or solve this debate. For the time being we assume them they explain open data from same perspective

The **social** perspective examines open data initiative from the societal, community, and cultural angles; it investigates how the benefits of open data can be reached to marginal citizens, and the capacity to benefit them (Davies et al. 2013). Many studies found that most *open data* initiatives aim societal values and benefits. Garbett et al. (2010) found that integrating crime data with social media generate added value to data. Also Bichard and Knight (2012) demonstrated how open data could improve public services. In general, citizens now are more concerned about their rights as well as more responsible to the society than before. They demand that government should disclose crime data, for example, which will then be analysed by journalists or social workers or real estate analysts to examine the trend of crimes in different suburbs of a city. Studies also discussed about potential social challenges too.

Economic perspective explains financial benefits and liabilities, and the promotion and regulation of an economy associated with open data. Studies claim that the size of open data economy is huge, in Europe itself is between €27 billion to €140 billion (Kulk and Van Loenen 2012). Studies actually are obsessed with generating economic benefit from selling innovative applications that use open data. But, *open data* incur costs and also may reduce income (Boulton et al. 2011). Studies proposed several revenue-sharing mechanisms such as: (a) *fee-for–service model*, each participating agency (which will use the data) would pay its share depending on the volume of data sought and transacted; (b) *exchange model*, data itself will be the currency for its exchange; (c) *public and/or private sponsor*, the agency will obtain fund from federal government or private organisations with the promise to share data; and (d) *compensation model*, the departments that sell data as one of their revenue sources, (state/federal) government would compensate such "loss" by providing compensatory fund (Conradie and Choenni 2014; Janssen 2011; Janssen et al. 2012; Peled 2011).

Institutional perspective focuses how institutions may enable open data release and/or use. Opening data is institutionalised in many government departments, universities, and research institutes as a requirement of ensuring budget or funds (Childs et al. 2014; Tananbaum 2008). Similarly, authors to many journals need to confirm that they will provide the experiment data on request (Andreoli-Versbach and Mueller-Langer 2013; Hrynaszkiewicz 2011) while some journals (e.g. *Public Library of Science*) mandate such release (but with a temporary embargo so that the principal

investigators can utilise the fuller range of experiment data) (Tananbaum 2008). Studies are optimistic that such mandatory policy would eventually lead to voluntary data-sharing culture. Accordingly, both government departments as well as research communities started *movement* for special digital depositories of data, which will be accessible to anyone interested to use/challenge the results (Hester 2014; Murray-Rust 2008).

A number of studies used **legal** perspective in order to understand the legal basis of *open data* policies, explore the anomalies and loopholes, and offer solutions. "While many open data initiatives do not have an explicit legal basis" (Dulong de Rosnay and Janssen 2014, p. 5), they follow the existing legal framework which is actually based on "two movements with overlapping perspectives" (Lassinantti et al. 2014, p.18): the Right to Information (RTI) Movement and Open Government Data (OGD) Movement. Integrating these two 'movements', some studies tried to develop a coherent format but still confusion/ misinterpretation exists (Janssen 2011). Moreover, many others consistently urged to develop a legal framework for national, regional, and global use of *open data.* Such framework may focus some of the legal issues: (a) for global and regional datasets who should administer the legal or regulatory issues, (b) who owns data and thus be liable when users are affected using erroneous or obsolete data – data creator or the data provider (Dulong de Rosnay and Janssen 2014)? (c) what reward-punishment model is acceptable in case of (not) acknowledging the originator(s) (Boulton 2014), (d) should data be open, which are collected by private funding but worked on the interest of public (such as Fukushima findings) (Boulton et al. 2011)?

Finally, **operational** perspective explicates the issues for better operation of *open data.* Zuiderwijk, Janssen, Choenni et al. (2014) found that some issues (e.g. privacy, trust, security, standards) are important but makes the data release process complicated, in public setting; Hester (2014) briefly explained the issues of providing raw data for scientific journal publications. But, it is commonly argued that the issue of re-use mechanism is more important and complicated than deposition of data. A common prescription for enhanced reusability of open data is to provide unambiguous and understandable data (Hester 2014) – but the process in not so easy. Nevertheless, Zuiderwijk et al. (2014c) formulated five principles to improve open data release: have a rigorous planning at the

beginning, develop guideline addressing issues such as privacy violation, understand the other associated actors related to data sharing, routinize the opening process, and monitor the use of data. From an operational perspective Zuiderwijk et al. (2014d) proposed the processes, procedures, and stakeholders of an open data ecosystem. Similarly, in earth observation application Zotti and La Mantia (2014) identified that every hour enormous amount of data are produced by satellites that increase the already stored data, and their meaning get changed with changing the context – therefore, data should be processed as soon as the data is produced/received and provide them quickly in standard form so that they can be reused, along with the context (Krotoski 2012).

Table 4 Classification of the perspectives (n=99)

	Brief definition	References	Count (%)
Technical	Technical perspective starts from defining open data then emphasises and explains the technical aspects (e.g. data visualisation, data analysis, optimising metadata) of open data	3, 8, 11, 14, 18, 38, 43, 52, 60, 79, 81, 85, 87, 88, 29, 49, 90, 92	18 (18%)
Social	The *social* perspective examines open data from bigger perspective that focuses on its benefits and challenges to a society through use; it also investigates cross-cultural issues	4, 10, 30, 38, 40, 49, 83, 92	8 (8%)
Institutional/organisational	An institutional perspective examines the ways a institution enables and constrains the publication and adoption of open data; open data as an integral part of the department's mission and routinisation process - not just a discrete or isolated event	5, 24, 26, 38, 54, 72, 86	7 (7%)
Legal/ethical	This perspective demonstrates open data issues from legal point of view: potential conflict of existing laws (copyright, freedom of information, data/information privacy) and advocating open data legislation/ directives and policies etc.	16, 24, 38, 45, 46, 74, 76	7 (7%)
Economic	This perspective advocates that innovation through open data may stimulate economic growth. Economic perspective also analyses open data cost-ownership, cost-benefit analysis and so on	21, 28, 38, 45, 49, 51, 55, 40	8 (8%)
Operational	It focuses on the challenges using *open data* and examines the prerequisites (e.g. standard) for being able to use open data	3, 19, 52, 82, 95, 96	6 (6%)
Political/Policy/strategic	Problems implementing government directives; political issues such as privacy, security, trust; success story	7, 20, 24, 36, 38, 45, 48, 50, 56, 61, 64, 66, 78, 94, 80	15 (15%)
General	General overview, entrepreneur, marketing, business systems (e.g. ecosystems); behavioural	15, 17, 22, 27, 32, 47, 35, 83, 86, 89	10 (10%)
Mixed	Combination of any	1, 6, 12, 23, 25, 31, 33, 34, 39, 41, 42, 44, 63, 67, 68, 69, 77, 84, 91, 93	20 (20%)

3.1.3 Classification for level of analysis

Level of analysis refers to the level at which the research is conducted, data are collected and analysed, and the main issues and discussions are interpreted. Li and Zhang (2005) mentioned four levels of analysis in behavioral IS research: individual, group, organisational, and inter-organisational (societal) – we adapted this approach. Our analysis found that 55% articles analysed open data from organisational and inter-organisational, 17% from individual, 18% examined from abstract level, while the rest 10% used mixed.

Explaining the levels studies explored/used a number of factors. For example, from individual's perspective Meijer et al. (2014) and O'Hara (2012) investigated how open data affect citizen's trust in government, while Kassen (2013) demonstrated how Open Data project foster democratic process through transparency and accountability (Kassen 2013). From organisational level Zuiderwijk and Janssen (2014c) developed a systematic framework to examine *open data* policies and their implementation in seven Dutch government departments. Their study suggests inter-organisational collaboration for better organisational performance. Then, some studies examined open data without at a specific level. For instance, Yoose and Perkins (2013) presented the use and progress of LOD in library from general level, whereas Zotti and La Mantia (2014) focused on standardised linked data to understand earth data better. Finally, 10% articles included more than one level; for instance, the exploratory work of Janssen et al. (2012) examined *open data* from all of the classifications mentioned in Table 5.

Table 5 Level of analysis (n=72)

Level	Brief description	References	Count (%)
Individual	Individual level analysis focuses mainly on the relevant perception, cognition, and individual reactions toward open data. Also it narrates how open data may affect individual productivity or performance or mental satisfaction	4, 5, 17, 25, 30, 48, 56, 63, 64, 65, 70, 88	12 (17%)
Organisational	Analyses organisation-level factors (e.g. top management commitment) in order to enhance organisational performance (e.g. brand image)	1, 2, 6, 11, 15, 26, 29, 39, 42, 50, 51, 54, 55, 68, 72, 79, 83, 86, 95	19 (26%)
Inter-organisational (societal)	Dealing with national, regional or global	7, 10, 12, 14, 16, 18, 22, 32, 33, 34, 36, 40, 45, 49, 76, 78, 61, 66, 80, 89, 94	21 (29%)
Abstract	Dealing at the abstract level - open data quality, structure	3, 8, 27, 35, 41, 44, 52, 77, 81, 82, 90, 91, 93	13 (18%)

Mixed/integrated	Combination of two or more levels	16, 24, 31, 69, 85, 92, 96	7 (10%)

3.1.4 Classification for research methods

Classifying the studies based on research methods utilized in them, first we grouped them into empirical and non-empirical studies (presented in Table 2). Our investigation revealed that empirical method dominates *open data* research. Then, from Table 6, 47% of the studies applied qualitative approach, followed by conceptual papers (38%) leaving only 11% quantitative papers. Finally, classifying the studies according to research method 45% papers used conceptual method followed by other methods (e.g. case study, filed interview) with almost similar trend. Interestingly, no literature review type of paper is reported yet.

Table 6 Classification for research method (n=96)

Classification	Reference	Total count (%)
Research methodology		
Quantitative	5, 8, 14, 17, 18, 21, 26, 67, 72, 75, 78	11 (11%)
Qualitative	2, 3, 6, 7, 9, 10, 15, 19, 20, 22, 25, 30, 31, 32, 33, 34, 35, 36, 38, 39, 41, 42, 45, 46, 47, 48, 49, 50, 51, 53, 55, 56, 58, 61, 68, 69, 79, 83, 84, 85, 86, 88, 91, 92, 93	45 (47%)
Conceptual/theoretical	1, 12, 13, 16, 23, 24, 27, 28, 29, 37, 40, 43, 44, 52, 54, 57,59, 60, 62, 63, 65, 66, 70, 71, 73, 74, 76, 77, 80, 81, 82, 87, 89, 94, 95, 96	36 (38%)
Mixed (qualitative + quantitative)	4, 11, 64, 90	4 (4%)
Research method		
Case study	1, 10, 15, 18, 22, 25, 30, 31, 32, 45, 53, 55, 56, 61, 69, 79	16 (17%)
Workshop and/or field study and/or interview	2, 6, 7, 34, 38, 39, 49, 51, 62, 68, 85, 86, 93	13 (14%)
Conceptual (framework)/ Report review/desk-based	3, 12, 13, 16, 19, 23, 24, 27, 28, 29, 33, 35, 36, 37, 40, 43, 44, 46, 48, 50, 52, 57, 58, 59, 60, 65, 66, 70, 71, 73, 74, 76, 77, 80, 81, 82, 84, 87, 89, 91, 94, 95, 96	43 (45%)
Survey	5, 14, 17, 21, 26, 47, 72, 75	8 (8%)
Mixed	4, 9, 11, 20, 41, 42, 64, 83, 88, 90, 92	11 (11%)
Lab experiment	(Behkamal et al. 2014; Childs et al. 2014; McDonald and Léveillé, 2014; Veljković et al. 2014)	3 (3%)
Secondary data analysis	67, 78	2 (2%)

3.2 RQ2: What progress is made in research to understand open data?

3.2.1 Drivers of open data movement

There are several driving factors for open data movement that include political leadership, institutional pressure, and emergence of digital technologies.

Political leadership: political leaders institutionalised the concept of open data. Although *open data* concept has been discussed within small and discrete communities, political leaders institutionalised it. In fact, a clear difference can be observed - countries moved ahead to opening data where the leaders are its supporters or promoters compared to the other countries (Huijboom and Van den Broek 2011). Those leaders found open data as a tool to ensure transparency and participation of citizens in governance (Janssen 2012). Initiated by strong commitment to be open government precedes the following strategies to understand and implement open data policies:

(a) *Developing guidelines and infrastructure*: the first institutional step for opening data is developing new laws that permit releasing government information to anyone for any use. While some existing laws (such as freedom of information) cover most of it, some new are necessary. Similarly, governments develop guidelines for participating departments. Also, they develop technical standards stimulating access, interoperability, and reuse of data; and monitor the progress (Huijboom and Van den Broek 2011).

(b) *Promoting through learning*: In order to create and disseminate knowledge government departments organise and/or patronise workshops and experience-exchanging programs between departments, and develop knowledge management systems. They also demonstrate the economic side of open data movement using different communication channels (Hielkema and Hongisto 2013; Huijboom and Van den Broek 2011).

Institutional pressure: Open data initiatives are mostly driven by compelling policies that organisations cannot avoid. For example, several governments decided that data related to energy, health, and utility [for example] should be available to public so that they can reuse data by developing innovative applications. Sayogo et al. (2014) found that, recently, governments mandate disclosure of data to several departments and associated private firms to disclose information. Similarly, in order to survive in some markets, disclosing data is a competitive tool and/or requirement; for example, coffee to North America (Sayogo et al. 2014) or livestock to Europe, Japan, and Korea. In research there is a growing pressure to release experiment data, especially that are public-funded (McLeod 2012). Supporting mimetic pressure, Huijboom and Van den Broek (2011) claimed that inspiring examples from inspirational countries is the major driver of open data

movements - UK followed US, and Australia started from UK initiative. Finally, journalists, advocacy groups, and opinion-leaders put pressure and demand to departments and firms to release public data.

Technological development in digital market: More often, governments align policies to respond to environmental change. In last decades, the world experienced tremendous technological developments in digital market especially in computing, telecommunication networks; and availability, usability, and cost of Smartphones. It facilitated development of people's computing skills in terms of accessing, storing, manipulating, analysing, linking and distributing data and information (Boulton 2014; Boulton et al. 2011; Rohunen et al. 2014) and the explosive growth of mobile network, which is followed by the rise of social networks (Huijboom and Van den Broek 2011) with every thinkable apps on the mobile device. Technology improvement (mobile Internet) and technology trend (e.g. mobile apps) drive firms to bring up services that integrate open data (Huijboom and Van den Broek 2011).

3.2.2 Benefits/opportunities/outcomes of Open data

Open data is considered as an opportunity (Estermann 2014). Its potentials are believed to be enormous (Gurstein 2011; Krotoski 2012; Zuiderwijk et al. 2012c) for offering researchers, citizens, companies, and other stakeholders with many advantages (Zuiderwijk et al. 2012b). Actually, the benefits of *open data* are associated with the *scope of openness* (Lindman 2014). Here we present some benefits of open private data, government data, and research data.

Using open data, private firms may reach interact with their customers better (Estermann 2014). An anecdotal evidence from California revealed that *open data* policy can offer three 'great' advantages: (a) it can bring dispersed information so that the users can access data from a single repository, (b) users can use data in different formats and embed into users' website with continuous and automatic update, and data with more visualisation, (c) data are provided as a standard format that inspires third-parties for developing applications (MacGunigal 2014) through visualisation and mash-ups (Hendler et al. 2012). However, Streeter et al. (1996) found that *open data* networks do *marginally* to contribute in profit, and no impact on firm (ordering) efficiency.

In the context of open government, numerous studies claimed that there are many advantages, both to governments and citizens (Hendler et al. 2012). Repetitively, there are four schools of thoughts:

transparency, empowerment, economic growth through innovation, and social value. *First*, almost all studies believe that *open data* can increase transparency of government institutions through disclosing datasets related to government spending, statistics on GDP, export, import and so on (Kalampokis et al. 2011a). In that way public departments become open and accountable to the community (Janssen 2011). Such transparency becomes a tool to fight with government corruption (Linders 2013). *Second*, with open data principles, every citizen can access and use data, which was not true few years back. Opening data increases self-empowerment of citizens and motivates them to participate in the political process (Janssen 2011, 2012) because they now can analyse data and challenge a government. But challenging this thought Gurstein (2011) argue that *open data* will introduce 'data divide' as merely providing the *access* to data is not noteworthy but actual *use* does; a clear divide will be exposed between the groups who 'use' open data and who do not (because of financial and educational resources/skills; access to computing). Apart from citizens, *open data* equips policymakers, social analysts, and advocacy groups preparing better policies through accessing primary data (Arzberger et al. 2004; Janssen et al. 2012). *Third*, *open data* will increase social value: citizens are provided with the datasets related to public facilities (e.g. schools, bus stops, hospitals) (Kalampokis et al. 2011a) that may offer opportunity to enhance quality of social life. For example, the app *Patients Like M*e connects patients with similar disease (symptoms) so that they can share experience or information (e.g. hot spots in cities that trigger asthma attacks) (Hendler et al. 2012). Also, citizens may take a part in policing and law enforcement by criminal recording and investigation-tasks linked with security database (Huijboom and Van den Broek 2011). Similarly, people can analyse a social disorder (e.g. unsafe roads, crime reports, drugs reporting) and monitor what actions have been taken by government and if the situation improves. *Finally,* value-creation through open data can generate wealth through the downstream use of outputs by creating new jobs, developing and selling third-party-owned innovative services (Arzberger et al. 2004; Hendler et al. 2012; Janssen et al. 2012; Lindman 2014). Such applied part of public data offers greater returns from public investment. Additionally, for government, data release through *open data* portals saves costs as it is cheaper than rendering them into reports and applications.

In academic research, *open data* would lead to institutional and community benefits in terms of (a) speeding up research through greater accessibility to scholarly works (hereby cross-consultation and comparison of data/results) and reducing duplicate or redundant research, (b) long-term preservation of research outputs, (c) cost-saving means of collaboration and disseminating research results, (d) validation of experiment results, and (e) reproducibility of data for further processing (Boulton et al. 2011; Leonelli et al. 2013).

3.2.3 Impediments/challenges/risks

In spite of high expectations developed in theory, the benefits of open data (projects) are yet to be realised to reasonable level. Zuiderwijk et al. (2012b) suspect various reasons (for low adoption); some are technical but most are 'soft' or behavioural involving social and organisational issues (McLeod 2012). Open data projects involve different stakeholders (e.g. political leaders, bureaucrats, privacy and transparency advocates) at different stages; integrating their expectations is a vital challenge (Lassinantti et al. 2014).

Significant progress has been observed exploring the barriers of open data adoption. Zuiderwijk et al. (2012b) mentioned about some technical impediments while Zuiderwijk et al. (2012c) focused on both social and technical impediments. The latter study first explored 106 impediments from literature, grouped them into three main categories (data access, data use, and data deposit impediments), then presented 118 socio-technical impediments in 10 categories. Similarly, Barry and Bannister (2014) proposed 20 barriers under six headings: economic, technical, cultural, legal, administrative, and risk related. Martin et al. (2013) identified 49 barriers and grouped them in seven categories. Moreover, Sayogo and Pardo (2013) grouped the barriers into four 'perspectives': technological; social, organisational, and economical; and local contexts and specificity. Recently, Sayogo et al. (2014) developed five challenges to data opening for 'sustainable consumption'; Zotti and La Mantia (2014), however, believe that the problems are related to 4V: Volume, Velocity, Variety, and Veracity. In this section, we present an overall and brief discussion on the barriers assuming it as a synonym of challenges, impediments or risks, as used in literature. For simplicity, we grouped them into four categories.

The **individual** impediments are related to personal: understanding, awareness, and knowledge of and incentive to share open data. Most people have no or little understanding about open data – its potential, use, and problems. On top of that, *open data* use demands knowledge and skills of different types (Conradie and Choenni 2014) including English literacy, computing, statistical, and cognitive. Lack of knowledge develops some other barriers too, such as less use of data, which affects perceived usefulness. Generally, personal barriers can be lowered by (end) user training in data access, use, and visualisation.

Institutional impediments refer to the barriers associated with both organisation and society. Studies found that the main institutional barrier of opening data is related to management and resource. Among the managerial barriers lack of *awareness* and *knowledge* of the leaders as well as risk-averse *leadership* are vital. Next, *top management attitude* is important for setting and implementing effective strategy by assigning appropriate level of authority and resource. For institutions, opening data is expensive as well may reduce income; still, the incentives are not clear. Furthermore, *perceived loss of control* over the data is a vital institutional concern. Therefore, opening data is not a priority to a number of agencies. Next, opening and using data requires an extensive resource-base that includes financial, technical, and human resources. One of the serious consequence of open data is the potential *data-divide* - a privileged segment of people to accessing and using data (Huijboom and Van den Broek 2011). Yet, many studies claim that *open data* would add value to society though the impacts are still 'not so well-defined' (Lassinantti et al. 2014).

The **legal** barriers are related to information privacy and security, and data licensing ownership. Strong tension exists between *open data* policy and *information privacy* (Floridi 2014; Huijboom and Van den Broek 2011). Some datasets may threat individual and organisational privacy, business secret or national security (Janssen et al. 2012). Sometimes, a person may identify another person without prior knowledge or sincere intention - 'motivated intruder'. Privacy is still a concern in spite of using a number of technological solutions (e.g. anonymisation, encryption). Some open data applications favour taxpayers' interest (e.g. doctors who lost license), but removing 'sensitive details' from those applications provides as little value as with no data at all. Consequently, the conflict between privacy, transparency, and accountability may clash. However, Rohunen et al. (2014) believe that privacy is an

important but overly stated concern. The other legal barriers include data ownership, contract/agreement, and copyright and licensing. Generally, the creator of data has the exclusive right to control (use and reproduce) but need to ensure that data collected for one purpose should restrict its use in another application (Hossain and Dwivedi 2014; Janssen et al. 2012). Then, according to Shadbolt et al. (2012) licensing restrictions are the biggest barriers for open data adoption. Still, licensing policies are important because they provide the rules of use and reuse of data, and the cost mechanism.

Many studies (e.g. Janssen et al. 2012) explored **technological** barriers to open data adoption. The barriers include data quality, accessibility, compatibility, credibility, processability, and lack of standards. **First**, open datasets are perceived as complex (Estermann 2014; Huijboom and Van den Broek 2011), same as for open data projects (Davies et al. 2013). Agencies cannot release data before anonymisation, which is a complex process; many departments may not have resources to do such, and thus decide not to disclose. Responding to political pressure, departments may just release unstructured data (Janssen and Zuiderwijk 2012) that cannot be reused (Serra 2014) and consequently possess less **quality of data** (QoD). QoD "continuous to be a major issue" (Whitmore 2014, p.2); low QoD is actually synonymous to no data (Conradie and Choenni 2014). QoD refers to data without incompleteness, obsolesce/invalidity, duplicity/ redundancy, inconsistency, and so on. Comparing open data movements of five countries, Huijboom and Van den Broek (2011) found that the QoD of some governments is too limited to be published. Zuiderwijk and Janssen (2014b) postulated that QoD of a dataset increases with data that are properly documented and easy to interpret and correct. **Accessibility** is a fundamental issue with open data. People cannot access to raw data because many contents are not linked or indexed (which increases difficulty in searching and browsing) (Fleisher 2008; Janssen et al. 2012). More problematically, data are often left scattered across a wide range of sources using non-open (e.g. in text format), and inconsistent terms and standards (Linders 2013) or with *volatility* (keep changing the location of data or even remove it without any indication). For open data, **compatibility** is a serious concern because different parties may use incompatible dataset formats or platforms. Also, many data cannot be used because of language problem (Gurstein 2011; Fleisher 2008). Similarly, ***lack of standards*** (of metadata, for example) between different data-sources

is a serious barrier to open data (Huijboom and Van den Broek 2011; Janssen et al. 2012). Some studies urge to develop and follow a common standard of publishing open data that can be used in different systems (Linders 2013). Other technological barriers include credibility and processibility of data (Fleisher 2008).

The economics of open data is relatively less studied; consequently, the **economic** barriers are still not apparent. One of the greatest barriers to open data is associated *cost*: for opening, updating, and maintaining datasets; and developing/upgrading and maintaining infrastructure. To cover the costs several cost structures are used by public departments: cost-recovery (total cost to produce information divided by the anticipated number of purchasers), charging marginal processing cost, pay per page or per inquiry, and so on. However, imposition of costs may prevent transformative uses who could have offered business growth and revenues (Sunlight Foundation 2010). Furthermore, for many (government) organisations, selling data is a good source of revenue; thus *perceived loss of revenue* or *loss of (extra) income* is another economic barrier to open data adoption.

3.2.4 Theoretical models used/identified by existing studies

Not many studies applied existing theories and models, nor developed from scratch. Among the papers, no single model dominates; 12 theories have been applied by 11 studies.

Applying variables from Technology Acceptance Model (TAM) and IS Success Model, Charalabidis et al. (2014) developed and tested a behavioral model that examine future usage behavior of open data adopters. Jetzek et al. (2012) developed a two-by-two matrix explaining *value creation*, in terms of social and economic values, by OGD initiatives. Later, Jetzek et al. (2013) formulated a *value creation* model with four untested propositions. Estermann (2014) utilised Innovation Diffusion Theory (IDT). From a (pilot) survey using 72 respondents, he explored the risks and opportunities, and expected costs and benefits of *open data*. Moreover, he discussed the results in the light of IDT, yet, how he reached to the results from the survey is unclear. Most importantly, his study explained the generic innovation diffusion attributes without relative weights, and without any contextualised variables unique to open data adoption. Next, applying *Public Value Framework*, from two case studies, Meijer et al. (2014) developed a model of open data and public value where the dependent variable is *trust,* which is positively affected by *transparency,* and negatively by *privacy* and *open data*. Moreover,

open data "may conflict with" *transparency, privacy,* and *security*. However, Tolbert and Mossberger (2006) challenged the positive relationship between (*public) trust* and *confidence in government* with opening government data. Recently, Janssen et al. (2012) explained that open data would develop valuable insight among people who can actually challenge the managers. Zuiderwijk and Janssen (2014c) complimented that, some organisations have a tendency to publicise their data on similar types of websites and in similar ways – Institutional theories call it as *mimetic isomorphism*. Hielkema and Hongisto (2013) study established the effectiveness of *competition* of Porter's competitive forces model. The other models used in open data research are: Social Shaping of Technology (SST) by Lassinantti et al. (2014)**,** Coordination theory by Zuiderwijk and Janssen (2013), Actor-Network theory by Hunnius et al. (2014), theory of *social capital,* theory of *encapsulated trust* by O'Hara (2012). For the other theories in open data see Zuiderwijk et al. (2014b).

3.3 RQ3: Who are the contributing members?

3.3.1 Publishing outlets

The most productive journal publishing open data articles is *the Journal of Theoretical and Applied Electronic Commerce Research* with 8 papers, followed by *Records Management Journal* (5 papers), and *Information Polity* (3 papers). Among these, 87.5% papers applied empirical method of research.

3.3.2 The most prolific authors

In order to find out the most prolific researchers and institutions in *open data* research, we applied a methodology which is followed by prior IS studies (e.g. Li and Zhang 2005); this technique uses three methods: normal rank, adjusted rank, and straight rank. As each method has its limitations, we used all three ranking methods together. Our analysis find that Zuiderwijk, A. and Janssen, M. are the most prolific researchers leaving the others far behind. Such observation reveals the need of more researchers to be involved in open data research.

Table 7 The list of the most prolific authors in open data research

NR	Author	NC	AR	Author	AC	SR	Author	SC
1*	Janssen, M.	21	1	Janssen, M.	7.10	1	Zuiderwijk, A.	16
1*	Zuiderwijk, A.	21	2	Zuiderwijk, A.	7.05	2*	Janssen, M.	2
2	Choenni, S.	6	3*	Janssen, K.	2.50	2*	Janssen, K.	2
3*	Charalabidis, Y.	4	3*	Davies, T.	2.50	2*	Davies, T.	2
3*	Jeffery, K.	4	4	Choenni, S.	1.78	2*	Alexopoulos, C.	2
3*	Meijer, R.	4	5	McLeod, J.	1.50	2*	Albano, C.S.	2
4*	Janssen, K.	3	6	Charalabidis, Y.	1.19	2*	Ferro, E.	2

4*	Davies, T.	3	7	Jeffery, K.	1.16	2*	Jetzek, T.	2
4*	McLeod, J.	3	8	Meijer, R.	0.98	2*	Kalampokis, E.	2
4*	Alexopoulos, C.	3	9	Alexopoulos, C.	0.86	2*	Sayogo, D. S.	2

[Legend: AC: Adjusted Count, AR: Adjusted Rank, NC: Normal Count, SC: Straight Count, SR: Straight Rank]

3.3.3 The most prolific institutions housing open data research

Unlike other similar studies (e.g. Li and Zhang 2005), in this study we did not prepare the list of the institutes that house the most prolific researchers; rather, we identified the most prolific institutes by their own contributing to open data research because we suspect that an institute having 5 researchers producing 5 different papers individually could have been superseded by a university having one researcher producing 6 papers. We applied the same method using normal, adjusted, and straight calculation. In case of the authors with more than one affiliation, the first one we considered. From Table 9 it is depicted that, by any measure, the Delft University of Technology has come up as the top with substantial gap to the next, which urges more institutions to research on open data.

Table 8 The list of the most productive institutions in open data research

NR	Institute	NC	AR	Institute	AC	SR	Institute	SC
1	Delft Uni. of Tech, Netherlands	38	1	Delft Uni. of Tech, Netherlands	14.03	1	Delft Uni. of Tech, Netherlands	17
2	Uni. of the Aegean, Greece	8	2	Uni. of Southampton, UK	3.00	2*	Uni. of Southampton, UK	3
3	Uni. of Macedonia, Greece	7	3	Uni. of the Aegean, Greece	2.66	2*	Uni. of Macedonia, Greece	3
4*	Rotterdam Uni. of Applied Science, Netherlands	6	4*	Uni. of Macedonia, Greece	2.50	2*	Uni. of the Aegean, Greece	3
4*	Copenhagen Business School, Demark	6	4*	Interdisciplinary Centre for Law and ICT	2.5	3*	Interdisciplinary Centre for Law and ICT	2
5*	Uni. of Albany, US	5	5	Rotterdam Uni. of Applied Science, Netherlands	2.19	3*	Uni. of Albany, US	2
5*	Uni. of Maryland, US	5	6	Copenhagen Business School, Demark	2.00	3*	Federal Uni. of Pampa, Brazil	2
6*	Uni. of Zagreb, Croatia	3	7	Uni. of Maryland, US	1.67	3*	Uni. of Maryland, US	2
6*	Luleå University of Technology, Sweden	3	8*	Uni. of Albany, US	1.50	3*	Copenhagen Business School, Demark	2
6*	Interdisciplinary Centre for Law and ICT	3	8*	Federal Uni. of Pampa, Brazil	1.50			

[Legend: AC: Adjusted Count, AR: Adjusted Rank, NC: Normal Count, SC: Straight Count, SR: Straight Rank]

4. RESEARCH GAPS, FUTURE RESEARCH DIRECTIONS AND IMPLICATIONS

In this section, we aim to propose and develop some agenda for open data research. In general, from Table 6, we need more quantitative evidence using positivist and quantitative approach with formal hypotheses and measurable variables. Moreover, Hendler et al. (2012) claimed that the most extensive

use of *open data* is observed by health community; but referring to Table 3, it is surprising that no study used health or education context. Therefore, future research is needed.

Open data behavioral models: Referring to RQ 2.4, we need more investigation on open data adoption and diffusion process and stages, how these stages should be managed. First, existing literature discretely explored/used some antecedents to open data adoption (such as community participation), however, a comprehensive adoption model is to be developed. Second, it is agreed that open data increases government transparency and citizens' empowerment and participation in governance; therefore, it is recommend that citizens use such new powers as more they use, the more possibilities are that OGD will further be released (Shadbolt et al. 2012). This is related to continuance. Third, few studies mentioned that open data policies have to be treated as a regular activity of a department or firm (rather than just releasing some data as part of the commitment) - which IS studies call as routinisation. By doing so, the antecedents and consequents of open data diffusion will be exposed. Studies demonstrate that both public and private organisations are closely involved in open data movement. This may be difficult yet valuable for managers to compare the factors in both settings - Do we need separate models for public and private sector investigations?

Similarly, although a few case studies explained the success stories of government open data, little is known on the success components; a formal guideline indicating how to evaluate the success of *open data* initiatives is yet to be proposed (Whitmore 2014). Nonetheless, Veljković et al. (2014) developed a benchmark in the context of open government. IS success models (e.g. Delone and McLean's) could be a starting point using some already developed constructs (e.g. perceived value (Hendler et al. 2012) or information quality (Zuiderwijk and Janssen 2014b)).

Open data involves chains of activities that include generate/collect, process, store, and disseminate data integrating several value-adding agents including (raw) data providers, linked data provider (or data service provider), data application provider (or application developer) (Lindman 2014), and end-users (Zuiderwijk et al. 2014d). These primary actors are dependent on supportive agents such as regulatory bodies, technology infrastructure, etc. Hence, open data can be understood with a value chain model (Albano 2013). Similarly, future studies can formulate open-data ecosystem (Davies 2011; Sande et al. 2013; Zuiderwijk et al. 2014d) while presenting the detailed nature of the

ecosystem as well the relative responsibilities of the keystone and niche firms (Zuiderwijk et al. 2012c).

To test *open data* behavioral models, the measurement scales of the associated variables need to be developed. The different requirements for open data e-infrastructure, developed by Zuiderwijk et al. (2013a) and Web 2.0 functionalities for data sharing by Alexopoulos et al. (2014) can be contextualised and be used for measuring different dimensions.

'Show me the money': If a department has to release data openly why it should spend too much to collect, store, maintain, and disseminate data? Referring to the discussion on economic perspectives and economic barriers of open data, future research on economic framework is needed for better understanding of the economic model of *open data*. How firms can evaluate open data investments in advance? It is a general perception or expectation that *open data* will contribute to economic growth. However, the economic success factors are yet to be reported. Moreover, data is commodity; hence open data is not necessarily free. Data generation and maintenance costs money - hence, it can be priced – but who should pay and how (Boulton et al. 2011)? What should be the ideal structure of cost sharing of data creation, maintenance, and up-gradation? How should the firms share profit? Moreover, opening data may lessen income (Leonelli et al. 2013) – how to make it up. Furthermore, a creative trade-offs between data release and associated costs could itself be a competitive strategy, and is important to investigate.

Knowing the political game: Opening data may increase transparency but may reduce trust towards a government. Open data often are misinterpreted or abused (to satisfy the interest of a particular party) by the user community. Hence, public agencies sometimes take advantage of some of the related terms used in *open data* policies. For instance, Peled (2011) demonstrated that government bureaucrats, by law, are not bound to publish 'records' but 'data'. A record may contain information about an individual such as personal identifier. But records managers use the terms data, information, and record as interchangeable meaning (Borglund et al. 2014). It may provide a defense to the agency to fragment and manipulate the data. Therefore, a clear political definition of *open data* that may satisfy different relevant stakeholders and context is a burning research agenda.

As discussed in *political perspective* section, the cat-and-mouse game between bureaucrats and

politicians is now an open secret. Political governments *direct* or *mandate* agencies to disclose public data (at least to the minimum); but bureaucrats adhere to their own agenda. In order to qualify as a participant and to escape the scrutiny of politicians; the bureaucrats release less useful or less valuable data (Peled 2011; Zuiderwijk and Janssen 2014b). Therefore, researchers identify the role and effectiveness of directives/mandates entice public department to release data. Future research also may pay attention on how to reduce the gap between political agenda/directive and agency's performance on open data projects, and formulise a reasonable balance between no-data and useless-data.

Moreover, some applications need data from different nations (such as data on public health, food production and safety, meteorology, or terrorism); not to so many proponents are visible yet who examine from multidisciplinary approach and about global openness/sharing, except a few (e.g. Arzberger et al. 2004; Shadbolt et al. 2012). Therefore, future research may define 'global open data' and how are they going to be shared beyond political and cultural boundaries?

Legal and ethical dilemma: Our analysis finds that the current legal frameworks neither protect the interests of the associated parties nor they encourage public data-use. Referring to Table 4, many studies emphasised on the importance of developing a legal framework explaining publishing and using public data (Conradie and Choenni 2014), but they did not identify the components of such legal framework. Such framework would answer to crucial questions such as who is (legally) liable when open data are misused (Kulk and Van Loenen 2012), or what reward-punishment model is acceptable in case of acknowledging (or not) the originator(s) (Boulton 2014).

Open data may conflict with privacy legislations (Boulton et al. 2011; Kulk and Van Loenen 2012): what are the solutions? Prior research found that privacy perceptions and seriousness varies in counties: European countries value privacy more than the US (Zuiderwijk and Janssen 2014b), while in general developed countries are more serious than developing nations (Hossain and Dwivedi 2014). Future research could be performed to investigate the impact of other variables (e.g. reputation/brand of the agency, or data/app provider) to privacy perceptions. Finally, privacy is a multifaceted dimension; it cannot be ensured only with 'explicit consent' (Childs et al. 2014; Rohunen et al. 2014), a holistic view is essential (Hossain and Dwivedi 2014).

⏏ What is the definition of privacy in terms of open data when public interest is associated

with the discloser of personal data (e.g. terrorists, HIV patients, public procurement agents)? What is the right balance of privacy and public benefit? What is the lifetime of private data?

⏏ What responsibilities should governments take in developing and deploying regulations to ensure privacy rights? Differentiate the role in state, national, and global contexts.

⏏ Given that privacy is more an ethical issue than a technological flaw, what social, legal, and behavioural solutions should be implemented at individual, organisational, and societal level?

The current study offers implications mainly to research community: both to editors and individual researchers. Generally, open data practitioners would get an extensive look on the 'so far' progress in open data research and be able to identify papers relevant according to their interest.

Banville and Landry (1989) claimed that most IS studies suffer from "vocational school"-type of research focusing on short-term view of a domain; 'opportunistic' and 'reputation-focused researchers' do not actually focus on long-term theoretical works (p. 57). This study exposed the gaps in open data literature and provides a guideline for the research community. In order to offer long-term contributions in open data research, researchers may consider the findings of this study and conduct future research, while journals may arrange special issues or conferences may arrange focused tracks.

The exploration of the prolific authors and their affiliations are useful to many people including research/doctoral students, research centers, or funding bodies who are passionate about open data research and seeking intellectual or human resources in this topic. It also offers a list of potential collaborators.

6. CONCLUSIONS AND LIMITATIONS

Open data is one step further to knowledge-based society and economy; it ensures seamless flow of data inspiring collaborative decision-making from user side and promoting transparency from provider side. The pressure for opening data now comes from different sources. Consequently, a growing trend is observed on the efforts and investments gone in publishing over a million datasets worldwide.

Neither its benefits are evident nor to declare it as a fad. The literature review conducted by this study is a contribution to open data body-of-knowledge. It has examined a large number of relevant articles from different angles. Moreover, the research agenda it has derived may shape future research in this emerging area.

The current study has some limitations. The main weakness is the exclusion of technical studies and papers from other domains than IS such as biology. Within the IS domain even, the selection is not exhaustive; the papers were selected applying a well-applied methodology. Some other studies could discuss *open data* without using these words in title and were not included for our analysis. The second limitation is: constrained by the research scope, we covered five major databases. Finally, for analysis we emphasised on journal articles to maintain rigor; nevertheless, papers from books or conference proceedings have been used in order to support certain arguments and to elaborate relevant points. This review article promises to play a vital role for advancing and shaping research, practice and policy related to *open data* in future.

REFERENCES

Albano, Cláudio Sonáglio. 2013. *Open government data: a value chain model proposal.* Paper presented at the Proceedings of the 14th Annual International Conference on Digital Government Research.

Albano, Claudio Sonaglio, and Reinhard, Nicolau. 2014. Open government data: facilitating and motivating factors for coping with potential barriers in the Brazilian context. In M. Janssen, H. J. Scholl, M. A. Wimmer and F. Bannister (Eds.), *Electronic Government* (pp. 181-193). Berlin Heidelberg: Springer.

Alexander, Ruth. 2013. Reinhart, Rogoff... and Herndon: The student who caught out the profs. *BBC News Magazine*.

Alexopoulos, Charalampos, Spiliotopoulou, Lefkothea, and Charalabidis, Yannis. 2013. *Open data movement in Greece: a case study on open government data sources.* Paper presented at the Proceedings of the 17th Panhellenic Conference on Informatics.

Alexopoulos, Charalampos, Zuiderwijk, Anneke, Charapabidis, Yannis, Loukis, Euripidis, and Janssen, Marijn. 2014. Designing a second generation of open data platforms: integrating open data and social media. In M. Janssen, H. J. Scholl, M. A. Wimmer and F. Bannister (Eds.), *Electronic Government* (pp. 230-241). Berlin Heidelberg: Springer.

Andreoli-Versbach, Patrick, and Mueller-Langer, Frank. 2013. Open access to data: An ideal professed but not practised. *Research Policy, 43*, 1621-1633.

Arzberger, Peter, Schroeder, Peter, Beaulieu, Anne, Bowker, Geof, Casey, Kathleen, Laaksonen, Leif, . . . Wouters, Paul. 2004. An international framework to promote access to data. *SCIENCE, 303*, 1777-1778.

Avison, David E, Dwivedi, Yogesh K, Fitzgerald, Guy and Powell, Philip 2008. The Beginnings of a New Era: Time to Reflect on 17 Years of the *ISJ. Information Systems Journal*, 18(1), pp 5-21

Barry, Emily, and Bannister, Frank. 2014. Barriers to open data release: A view from the top. *Information Polity, 19*(1), 129-152.

Bates, Jo. 2014. The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom. *Government Information Quarterly, 31*(3), 388-395.

Behkamal, Behshid, Kahani, Mohsen, Bagheri, Ebrahim, and Jeremic, Zoran. 2014. A metrics-driven approach for quality assessment of linked open data. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(2), 64-79.

Berners-Lee, Tim. 2006. Linked data-design issues. from http://www.w3.org/DesignIssues/LinkedData.html

Bertot, John Carlo, Gorham, Ursula, Jaeger, Paul T, Sarin, Lindsay C, and Choi, Heeyoon. 2014. Big data, open government and e-government: Issues, policies and recommendations. *Information Polity, 19*(1), 5-16.

Bichard, Jo-Anne, and Knight, Gail. 2012. Improving public services through open data: public toilets. *Municipal Engineer, 165*(3), 157-165.

Borglund, Erik, Engvall, Tove, McLeod, Julie, and Thurston, Anne. 2014. Open data? Data, information, document or record? *Records Management Journal, 24*(2), 163-180.

Boulton, Geoffrey. 2014. The open data imperative. *Insights: the UKSG Journal, 27*(2), 133-138.

Boulton, Geoffrey, Rawlins, Michael, Vallance, Patrick, and Walport, Mark. 2011. Science as a public enterprise: the case for open data. *The Lancet, 377*(9778), 1633-1635.

Braunschweig, Katrin, Eberius, Julian, Thiele, Maik, and Lehner, Wolfgang. 2012. The state of open data: limits of current open data platforms. *Web Science Track*.

Casellas Serra, Lluís Esteve. 2014. The mapping, selecting and opening of data: the Records Management contribution to the Open Data project in Girona City Council. *Records Management Journal, 24*(2), 87-98.

Cerrillo-i-Martínez, Agustí. 2012. Fundamental interests and open data for re-use. *International Journal of Law and Information Technology, 20*(3), 203-222.

Chan, Calvin ML. 2013. *From open data to open innovation strategies: creating e-Services using open government data.* Paper presented at the 46th Hawaii International Conference on System Sciences (HICSS), Hawaii.

Charalabidis, Yannis, Loukis, Euripides, and Alexopoulos, Charalampos. 2014. *Evaluating second generation open government data infrastructures using value models*. Paper presented at the Proceedings of the 2014 47th Hawaii International Conference on System Sciences.

Childs, Sue, McLeod, Julie, Lomas, Elizabeth, and Cook, Glenda. 2014. Opening research data: issues and opportunities. *Records Management Journal, 24*(2), 142-162.

Conradie, Peter, and Choenni, Sunil. 2014. On the barriers for local government releasing open data. *Government Information Quarterly*, S10-S17.

Craveiro, Gisele S, Santana, Marcelo T, and Albuquerque, João P. 2013. *Assessing open government budgetary data in Brazil.* Paper presented at the ICDS 2013, The Seventh International Conference on Digital Society.

Davies, T, Perini, F, and Alonso, J. 2013. Researching the emerging impacts of open data. *World Wide Web Foundation,[Online], Available: http://www. opendataresearch. org/sites/default/files/posts/Researching% 20the% 20emerging% 20impacts% 20of, 20*.

Davies, T. 2011. Open Data: Infrastructures and ecosystems.

Dulong de Rosnay, Melanie, and Janssen, Katleen. 2014. Legal and institutional challenges for opening data across public sectors: towards common policy solutions. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(3), 1-14.

Dwivedi Yogesh Kumar and Kuljis, Jasna 2008. Profile of IS Research Published in the European Journal of Information Systems. *European Journal of Information Systems*, 17(6), pp 678-693

Dwivedi, Yogesh Kumar, Venkitachalam, Krishna, Sharif, Amir, Al-Karaghouli, Wafi and Weerakkody, Vishanth 2011. Research Trends in Knowledge Management: Analyzing the Past and Predicting the Future. *Information Systems Management,* 28(1), pp 43-56

Eckartz, Silja M, Hofman, Wout J, and Van Veenstra, Anne Fleur. 2014. A decision model for data sharing. In M. Janssen, H. J. Scholl, M. A. Wimmer and F. Bannister (Eds.), *Electronic Government: Lecture Notes in Computer Science* (pp. 253-264). Berlin Heidelberg: Springer.

Estermann, Beat. 2014. Diffusion of open data and crowdsourcing among heritage institutions: results of a pilot survey in Switzerland. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(3), 15-31.

Ferro, E, and Osella, M. 2012. *Business models for PSI re-use: a multidimensional framework.* Paper presented at the The European Commission Workshop "Using Open Data: Policy Modeling, Citizen Empowerment, Data Journalism", Brussels, Belgium.

Ferro, E, and Osella, M. 2013. *Eight business model archetypes for PSI re-use.* Paper presented at the Open Data on the Web''Workshop, Google Campus, London.

Fleisher, Craig S. 2008. Using open source data in developing competitive and marketing intelligence. *European Journal of Marketing, 42*(7/8), 852-866.

Floridi, Luciano. 2014. Open data, data protection, and group privacy. *Philosophy and Technology, 27*(1), 1-3.

Garbett, Andrew, Linehan, Conor, Kirman, Ben, Wardman, Jamie, and Lawson, Shaun. 2010. Using social media to drive public engagement with open data. *Digital Engagement 11, Newcastle, UK*.

Gleditsch, Nils Petter, Metelits, Claire, and Strand, Havard. 2003. Posting your data: Will you be scooped or will you be famous. *International Studies Perspectives, 4*(1), 89-97.

Gurstein, Michael B. 2011. Open data: Empowering the empowered or effective data use for everyone? *First Monday, 16*(2).

Hendler, James, Holm, Jeanne, Musialek, Chris, and Thomas, George. 2012. US government linked open data: semantic. data. gov. *IEEE Intelligent Systems, 27*(3), 25-31.

Hester, JR. 2014. Closing the data gap: Creating an open data environment. *Radiation Physics and Chemistry, 95*, 59-61.

Hielkema, Hendrik, and Hongisto, Patrizia. 2013. Developing the Helsinki smart city: the role of competitions for open data applications. *Journal of the Knowledge Economy, 4*(2), 190-204.

Hossain, Mohammad Alamgir, and Dwivedi, Yogesh K. 2014. What improves citizens' privacy perceptions toward RFID technology? A cross-country investigation using mixed method approach. *International Journal of Information Management, 34*(6), 711-719.

Hrynaszkiewicz, Iain. 2011. The need and drive for open data in biomedical publishing. *Serials: The Journal for the Serials Community, 24*(1), 31-37.

Huijboom, Noor, and Van den Broek, Tijs. 2011. Open data: an international comparison of strategies. *European Journal of ePractice, 12*(1), 1-13.

Hunnius, Sirko, Krieger, Bernhard, and Schuppan, Tino. 2014. *Providing, guarding, shielding: open overnment data in Spain and Germany*. Paper presented at the 2014 EGPA Annual Conference, 10-12 September 2014 in Speyer, Germany.

Ivanov, Marijana, Varga, Mladen, and Pejić Bach, Mirjana. 2014. *Government open data portal: How government strategies should be more open.* Paper presented at the 7th International Conference: An Enterprise Odyssey: Leadership, Innovation and Development for Responsible Economy.

James, Laura. 2013. Defining open data. Retrieved from http://blog.okfn.org/2013/10/03/defining-open-data/

Janssen, Katleen. 2011. The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly, 28*(4), 446-456.

Janssen, Katleen. 2012. Open government data and the right to information: Opportunities and obstacles. *The Journal of Community Informatics, 8*(2).

Janssen, Marijn, Charalabidis, Yannis, and Zuiderwijk, Anneke. 2012. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258-268.

Janssen, Marijn, and Zuiderwijk, Anneke. 2012. *Open data and transformational government.* Paper presented at the tGov Conference, May.

Jetzek, Thorhildur, Avital, Michel, and Bjørn-Andersen, Niels. 2012. *The value of open government data: a strategic analysis framework.* Paper presented at the 2012 Pre-ICIS Workshop.

Jetzek, Thorhildur, Avital, Michel, and Niles, rn-Anderson. 2013. *The generative mechanisms of open government data.* Paper presented at the 21st European Conference on Information Systems (ECIS), Utrecht, Nederlands.

Kaasenbrood, Maaike, Zuiderwijk, Anneke, Janssen, Marijn, Jong, Martin de, and Bharosa, Nitesh. 2015. Exploring the factors influencing the adoption of open government data by private organisations. *International Journal of Public Administration in the Digital Age, x*(y), 1-xx.

Kalampokis, Evangelos, Tambouris, Efthimios, and Tarabanis, Konstantinos. 2011a. A classification scheme for open government data: towards linking decentralised data. *International Journal of Web Engineering and Technology, 6*(3), 266-285.

Kalampokis, Evangelos, Tambouris, Efthimios, and Tarabanis, Konstantinos. 2011b. Open government data: a stage model. In M. Janssen, H. J. Scholl, M. A. Wimmer and Y.-h. Tan (Eds.), *Electronic Government* (pp. 235-246). Berlin Heidelberg: Springer.

Kassen, Maxat. 2013. A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly, 30*(4), 508-513.

Krotoski, Aleksandra K. 2012. Data-driven research: open data opportunities for growing knowledge, and ethical issues that arise. *Insights: the UKSG Journal, 25*(1), 28-32.

Kuk, George, and Davies, Tim. 2011. *The roles of agency and artifacts in assembling open data complementarities.* Paper presented at the 32nd International Conference on Information Systems (ICIS), Shanghai, China.

Kulk, Stefan, and Van Loenen, Bastiaan. 2012. Brave New Open Data World? *International Journal of Spatial Data Infrastructures Research, 7*, 196-206.

Lassinantti, Josefin, Bergvall-Kåreborn, Birgitta, and Ståhlbröst, Anna. 2014. Shaping Local Open Data initiatives: politics and implications. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(2), 17-33.

Leonelli, Sabina, Smirnoff, Nicholas, Moore, Jonathan, Cook, Charis, and Bastow, Ruth. 2013. Making open data work for plant scientists. *Journal of Experimental Botany, 64*(14), 4109-4117.

Li, Na Lina, and Zhang, Ping. 2005. The intellectual development of Human-Computer Interaction research: A critical assessment of the MIS literature (1990-2002). *Journal of the Association for information Systems, 6*(11), 9.

Linders, Dennis. 2013. Towards open development: Leveraging open data to improve the planning and coordination of international aid. *Government Information Quarterly, 30*(4), 426-434.

Lindman, Juho. 2014. Similarities of open data and open source: impacts on business. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(3), 46-70.

MacDonald, Stuart, and Martinez, Luis. 2008. Libraries in the Converging Worlds of Open Data, E-research, and Web 2.0. *Online, 32*(2), 36-40.

MacGunigal, Maria. 2014. Regrets about Open Data? Only that we didn't start sooner. *The Public Manager*.

Martin, Sébastien, Foulonneau, Muriel, Turki, Slim, and Ihadjadene, Madjid. 2013. *Open data: barriers, risks and opportunities.* Paper presented at the Proceedings of the 13th European Conference on eGovernment: ECEG 2013.

McDonald, John, and Léveillé, Valerie. 2014. Whither the retention schedule in the era of big data and open data? *Records Management Journal, 24*(2), 99-121.

McLeod, Julie. 2012. Thoughts on the opportunities for records professionals of the open access, open data agenda. *Records Management Journal, 22*(2), 92-97.

Meijer, Ronald, Conradie, Peter, and Choenni, Sunil. 2014. Reconciling contradictions of open data regarding transparency, privacy, security and trust. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(3), 32-44.

Murray-Rust, Peter. 2008. Open data in science. *Serials Review, 34*(1), 52-64.

Nugroho, R., Zuiderwijk, A., Janssen, M., and de Jong, M. . 2015. A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process and Policy, x*(y), 1-y.

O'Hara, Kieron. 2012. *Transparency, open data and trust in government: shaping the infosphere.* Paper presented at the Proceedings of the 3rd Annual ACM Web Science Conference.

Open Data Institute, ODI. nd. Guides: What makes data open?, from http://theodi.org/guides/what-open-data

Pabón, Gustavo, Gutiérrez, Claudio, Fernández, Javier D., and Martínez-Prieto, Miguel A. 2013. Linked Open Data technologies for publication of census microdata. *Journal of the American Society for Information Science and Technology, 64*(9), 1802–1814.

Peled, Alon. 2011. When transparency and collaboration collide: The USA open data program. *Journal of the American Society for Information Science and Technology, 62*(11), 2085-2094.

Reed, David. 2013. IDM data council open data discussion. *Journal of Direct, Data and Digital Marketing Practice, 15*, 166-169.

Robinson, Ricky, Rittenbruch, Markus, Foth, Marcus, Filonik, Daniel, and Viller, Stephen. 2012. Street computing: towards an integrated open data Application Programming Interface (API) for cities. *Journal of Urban Technology, 19*(2), 1-23.

Rohunen, Anna, Markkula, Jouni, Heikkila, Marikka, and Heikkila, Jukka. 2014. Open traffic data for future service innovation: addressing the privacy challenges of driving data. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(3), 71-89.

Rowen, Lee, Wong, Gane KS, Lane, Robert P, and Hood, Leroy. 2000. Publication rights in the era of open data release policies. *Science, 289*(5486), 1881-1881.

Sande, Miel Vander, Dimou, Anastasia, Colpaert, Pieter, Mannens, Erik, and Walle, Rik Van de. 2013. Linked data as enabler for Open data ecosystems. http://www.w3.org/2013/04/odw/odw13_submission_26.pdf

Saxby, Stephen, and Hill, Chris. 2012. Public sector information and open data: which way forward for the UK? *International Journal of Public Law and Policy, 2*(3), 299-333.

Sayogo, Djoko Sigit, and Pardo, Theresa A. 2013. Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data. *Government Information Quarterly, 30*, S19-S31.

Sayogo, Djoko Sigit, Zhang, Jing, Pardo, Theresa A, Tayi, Giri K, Hrdinova, Jana, Andersen, David F, and Luna-Reyes, Luis Felipe. 2014. Going Beyond Open Data: Challenges and Motivations for Smart Disclosure in Ethical Consumption. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(2), 1-16.

Shadbolt, Nigel, O'Hara, Kieron, Berners-Lee, Tim, Gibbins, Nicholas, Glaser, Hugh, and Hall, Wendy. 2012. Linked open government data: Lessons from data. gov. uk. *IEEE Intelligent Systems, 27*(3), 16-24.

Simón, L Fernando Ramos, Avilés, Rosario Arquero, Botezan, Iuliana, del Valle Gastaminza, Félix, and Serrano, Silvia Cobo. 2014. Open data as universal service. new perspectives in the information profession. *Procedia - Social and Behavioral Sciences, 147*, 126-132.

Stephens, Owen. 2011. Mashups and open data in libraries. *Serials: The Journal for the Serials Community, 24*(3), 245-250.

Streeter, Lynn A, Kraut, Robert E, Lucas Jr, Henry C, and Caby, Laurence. 1996. How open data networks influence business performance and market structure. *Communications of the ACM, 39*(7), 62-73.

Sunlight_Foundation. 2010. Ten principles for opening up govenment information. from http://sunlightfoundation.com/policy/documents/ten-open-data-principles/

Tananbaum, Greg. 2008. Adventures in open data. *Learned Publishing, 21*(2), 154-156.

Thurston, Anne Catherine. 2012. Trustworthy Records and Open Data. *The Journal of Community Informatics, 8*(2), 1-5.

Tolbert, Caroline J, and Mossberger, Karen. 2006. The effects of e-Government on trust and confidence in government. *Public Administration Review, 66*(3), 354-369.

Tsiavos, Prodromos, Stefaneas, Petros, and Karounos, Theodoros. 2013. The transposition of European Union open data/public sector information policies in Greece: a critical analysis. *Policy & Internet, 5*(4), 402-417.

Uhlir, Paul, and Schröder, Peter. 2007. Open data for global science. *Data Science Journal, 6*, 36-53.

Veljković, Nataša, Bogdanović-Dinić, Sanja, and Stoimenov, Leonid. 2014. Benchmarking open government: An open data perspective. *Government Information Quarterly, 31*(2), 278-290.

Weerakkody, Vishanth, Dwivedi, Yogesh K, and Irani, Zahir. 2009. The diffusion and use of institutional theory: a cross-disciplinary longitudinal literature survey. *Journal of Information Technology, 24*(4), 354-368.

Whitmore, Andrew. 2014. Using open government data to predict war: A case study of data and systems challenges. *Government Information Quarterly, xx*(xx), 1-9. doi: http://

Yannoukakou, Aikaterini, and Araka, Iliana. 2014. Access to government information: right to information and open government data synergy. *Procedia-Social and Behavioral Sciences, 147*, 332-340.

Yoose, Becky, and Perkins, Jody. 2013. The Linked Open Data Landscape in Libraries and Beyond. *Journal of Library Metadata, 13*(2-3), 197-211.

Zotti, Massimo, and La Mantia, Claudio. 2014. Open Data from Earth Observation: from Big Data to Linked Open Data, through INSPIRE. *Journal of e-Learning and Knowledge Society, 10*(2), 91-100.

Zuiderwijk, A, Janssen, M, and Choenni, S. 2012. *Open data policies: impediments and challenges.* Paper presented at the 12th European conference on e-government (ECEG 2012). Barcelona, Spain.

Zuiderwijk, Anneke, Gaseó, Mila, Parycek, Peter, and Janssen, Marijn. 2014a. Special issue on transparency and open data policies: guest editors' introduction. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(3), I-IX.

Zuiderwijk, Anneke, Helbig, Natalie, Gil-García, J Ramón, and Janssen, Marijn. 2014b. Review of the State-of-the-Art and an Emerging Research Agend a: Guest Editors' Introduction. *Journal of Theoretical and Applied Electronic Commerce Research, 9*(2), i-xiii. A

Zuiderwijk, Anneke, and Janssen, Marijn. 2012. *A comparison of open data policies and their implementation in two Dutch ministries.* Paper presented at the Proceedings of the 13th Annual International Conference on Digital Government Research.

Zuiderwijk, Anneke, and Janssen, Marijn. 2013. A coordination theory perspective to improve the use of open data in policy-making. In M. A. Wimmer, M. Janssen and H. J. Scholl (Eds.), *Electronic Government* (pp. 38-49. Berlin Heidelberg: Springer.

Zuiderwijk, Anneke, and Janssen, Marijn. 2014a. Barriers and development directions for the publication and usage of open data: a socio-technical view. *Opportunities and Challenges for Public Governance, 4*, 115-135.

Zuiderwijk, Anneke, and Janssen, Marijn. 2014b. *The negative effects of open government data-investigating the dark side of open data.* Paper presented at the Proceedings of the 15th Annual International Conference on Digital Government Research.

Zuiderwijk, Anneke, and Janssen, Marijn. 2014c. Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly, 31*(1), 17-29.

Zuiderwijk, Anneke, Janssen, Marijn, Choenni, Sunil, and Meijer, Ronald. 2014c. Design principles for improving the process of publishing open data. *Transforming Government: People, Process and Policy, 8*(2), 185-204. doi: doi:10.1108/TG-07-2013-0024

Zuiderwijk, Anneke, Janssen, Marijn, Choenni, Sunil, Meijer, Ronald, and Alibaks, R.S. 2012c. Socio-technical impediments of open data. *Electronic Journal of e-Government, 10*(2), 156-172.

Zuiderwijk, Anneke, Janssen, Marijn, and Davis, Chris. 2014d. Innovation with open data: Essential elements of open data ecosystems. *Information Polity, 19*(1), 17-33.

Zuiderwijk, Anneke, Janssen, Marijn, and Jeffery, Keith. 2013a. *Towards an e-infrastructure to support the provision and use of open data.* Paper presented at the Conference for E-Democracy and Open Governement.

Zuiderwijk, Anneke, Janssen, Marijn, Meijer, Ronald, Choenni, Sunil, Charalabidis, Yannis, and Jeffery, Keith. 2012d. Issues and guiding principles for opening governmental judicial research data. In H. J. Scholl, M. Janssen, M. A. Wimmer, C. E. Moe and L. S. Flak (Eds.), *Electronic Government* (pp. 90-101. Berlin Heidelberg: Springer.

Zuiderwijk, Anneke, Janssen, Marijn, and Parnia, Armin. 2013b. *The complementarity of open data infrastructures: an analysis of functionalities.* Paper presented at the Proceedings of the 14th Annual International Conference on Digital Government Research.

Zuiderwijk, Anneke, Jeffery, Keith, and Janssen, Marijn. 2012a. *The necessity of metadata for linked open data and its contribution to policy analyses.* Paper presented at the CeDEM 12 Conference for E-Democracy and Open Government 3-4 May 2012 Danube-University Krems, Austria.

Zuiderwijk, Anneke, Jeffery, Keith, and Janssen, Marijn. 2012b. 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.

About Authors

Dr. Mohammad Alamgir Hossain received Ph.D. in Business Information Systems and Post-grad Certificate in Research Commercialization from Curtin University, Australia; Master in e-Business from Edith Cowan University, Australia; and M.Sc. and B.Sc. in Engineering from Bangladesh University of Engineering & Technology (BUET). He has been working in RMIT University, Australia as a Research Fellow. His major research expertise and interests are: Behavioural Modelling in IS, RFID technology, e-Commerce, m-Commerce, m-Health, Open Data, supply chain management, and ICT in tourism. Dr. Hossain has published papers in Information Technology & People, Information Systems Frontiers, Australasian Journal of Information System, International Journal of Information Management, Journal of Organizational Computing & Electronic Commerce, Journal of Enterprise Information Management, Safety Science, Business Process Management Journal, Total Quality Management & Business Excellence, and International Journal of Information Systems and Supply Chain Management. Moreover, he contributed several book-chapters and presented papers in prestigious conferences including ACIS, PACIS, IFIP, ANZAM, and ANZMAC; and received the best paper award in 2013 IFIP WG8.6. He is a member of the editorial board of Journal of Enterprise Information Management; International Journal of Knowledge-based Organizations; Transforming Government: People, Process and Policy; and International Journal of Technology Diffusion; and is a member of the AIS, ACS, Engineers Australia, ANZAM, and life member of IFIP WG8.6. Mohammad can be reached at mahripon@yahoo.com

Professor Yogesh K Dwivedi is a Professor of Digital and Social Media and Head of Management and Systems Section in the School of Management at Swansea University, UK. He obtained his PhD and MSc in Information Systems from Brunel University, UK. He has co-authored several papers which have appeared in international refereed journals such as CACM, DATA BASE, EJIS, IJPR, ISJ, ISF, JCIS, JIT, JORS, TMR and IMDS. He is Associate Editor of *European Journal of Marketing*, *European Journal of Information Systems* and *Government Information Quarterly*, Assistant Editor of JEIM and TGPPP, Senior Editor of JECR and member of the editorial board/review board of several journals.

Dr Nripendra P. Rana is a Lecturer at the School of Management of Swansea University in the UK. He holds a BSc, an MCA, an MTech, and an MPhil degrees from Indian universities. He obtained his MBA (distinction) and PhD from Swansea University, UK. He is currently working in the area of technology and e-government adoption. He has varied work experience of teaching in the area of computer engineering and applications at undergraduate and postgraduate levels. He also possesses a good experience in the field of software development. He is a life member of computer society of India.