
Being Open About Data

Analysis of the UK open data policies and applicability of open data

ANTTI HALONEN

REPORTS

The Finnish Institute is a London-based private trust. Our mission is to identify emerging issues relevant to contemporary society and to act as catalyst for positive social change through partnerships.

We work with artists, researchers, experts and policy makers in the United Kingdom, Finland and the Republic of Ireland to promote strong networks in the fields of culture and society. We encourage new and unexpected collaborations and support artistic interventions, research, the creative industries, foresight and social innovation in new, socially central areas.

The Reports of the Finnish Institute in London is a series of publications, which publishes research, studies and results of collaborative projects carried out by the institute. The reports provide evidence and ideas for policy-makers and civic society organisations dealing with contemporary social and cultural challenges.

35–36 Eagle Street
London WC1R 4AQ
United Kingdom

T +44(0)20 7404 3309
www.finnish-institute.org.uk

Being Open about Data by the Finnish Institute in London is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License



Table of contents

List of abbreviations	3
Open-data timeline	4
Foreword from the Open Knowledge Foundation	6
Preface	7
Executive Summary	9
1. Introduction – Objectives and methodology	12
1.1 Objectives	12
1.2 Methods	15
2. Open data – What and why?	18
2.1 What is open data?	18
2.2 Why open data?	21
2.2.1 Efficiency	24
2.2.2 Democratic accountability	25
2.2.3 Empowering and public participation	27
2.2.4 Economy and entrepreneurship	28
3. Freedom-of-information continuum	30
3.1 From reactive FOI to proactive online transparency	32
3.2 Tradition of transparency in the UK	34
3.3 Political economy and transparency – entrepreneurial roots of open data and the European PSI Directive 2003	40
4. Institutional and legal contexts of open data	42
4.1 Institutionalisation of open data	42
4.2 From the Power of Information Report to the Public Sector Transparency Board, Local Public Data Panel and Public Data Corporation	45
4.3 Towards Open Licences	53
4.4. Do Licences Work?	58
5. Benefits and how to measure them	64
5.1 Internal benefits	65
5.1.1 Efficiency and new information management	65
5.1.2 Case: UK local-government transparency agenda and data producers' perceptions	66
5.1.3 Transparency	83
5.2 External benefits	85
5.2.1 Entrepreneurship	85
5.2.2 Democratic accountability	88

5.2.3 Case: Open data and cultural heritage	91
5.2.4 Participation	92
5.2.5 Case: Guardian's Data Journalism	97
6. Emerging problems and how to solve them	101
6.1 Confusion on policies	101
6.2 Privacy issues – a non-issue?	103
6.3 Politicised environment and gotcha culture	104
6.4 Data hugging	107
6.5 Data quality and context	111
6.6 Data divide and lack of interest	115
7. Conclusion	118
8. References	124

List of abbreviations

APPSI	Advisory Panel of Public Sector Information
CC	Creative Commons
COINS	Combined Online Information System
CSV	Comma-separated values
DCLG	Department of Communities and Local Government
DFID	Department for International Development
DPA	Data Protection Act
EIR	Environmental Information Regulations
EU	European Union
FOI	Freedom of information
FOIA	Freedom of information Act
GLAM	Galleries, Libraries, Archives and Museums
GP	General Practitioner
ICO	Information Commissioner's Office
ICT	Information and communication technology
LODLAM	Linked Open Data in Libraries, Archives and Museums
NESTA	National Endowment for Science, Technology and the Arts
NGO	Non-governmental Organisation
OGL	Open Government License
OPSI	Office of Public Sector Information
PDC	Public Data Corporation
PDF	Portable Document Format
PM	Prime Minister
PSI	Public Sector Information
RDF	Resource Description Framework
SME	Small and medium sized enterprise
SPARQL	SPARQL Protocol and RDF Query Language
TNA	The National Archives
UCL	University College London
UK	The United Kingdom of Great Britain and Northern Ireland
UKGLF	UK Government Licensing Framework
W3C	World Wide Web Consortium

Open-data timeline¹

Light grey = Labour Prime Minister

Dark grey = Conservative Prime Minister

February 1966	Fulton report recommends reforms in the civil service in order to “get rid of unnecessary secrecy
June 1969	The White Paper “Information and the Public Interest” addresses the issue of secrecy but does not endorse any specific reforms
March 1979	The Callaghan government publishes a Green Paper on open government proposing a non-statutory code on freedom of information. Bill fails when the government falls in the general election in April
July 1984	The Data Protection Act brings subject access rights to personal information held on computerised records
December 1992	Environmental Information Regulations (EIR) come into force, allowing access to environmental information under EU directive
April 1994	Code of Practice on Access to Government Information comes into force across central government
December 1997	The White Paper “Your Right to Know: The Government’s Proposals for a Freedom of Information Act” published
April 2003	Advisory Panel on Public Sector Information is established to advise on and encourage “opportunities in the information industry for greater re-use of public sector information”
November 2003	European Public Sector Information Directive is adopted. States must implement it by July 2005.
January 2005	Freedom of Information Act comes fully into force
July 2005	Re-use of Public Sector Information directive comes into force.
March 2006	Guardian launches “Free Our Data” campaign

¹ Adopted from Davies, Tim: *Open data, democracy and public sector reform. A look at open government data use from data.gov.uk*, <http://practicalparticipation.co.uk/odi/report/>; Tullo, Carol: presentation at APPSI seminar 6 December 2010, <http://www.nationalarchives.gov.uk/documents/meetings/061210Paper1-Carol-Tullo-presentation.pdf>; The Constitution Unit: A history of openness in Britain – a timeline, <http://www.ucl.ac.uk/constitution-unit/research/foi/what-is-foi2>

June 2006	“Measuring EU Public Sector Information Resources Study” estimates mean potential value from PSI re-use across Europe at €27bn.
January 2007	Power of Information Report commissioned.
March 2008	Interim Power of Information report published and Power of Information taskforce established
July 2008	Office of Public Sector Information Data Unlocking Service is launched
January 2009	In US, president Obama issues Memo on Transparency and Open Government as one of his first acts in office
February 2009	Tim Berners-Lee talks to the TED Conference about the need for Raw Data Now from governments
February 2009	Final Power of Information Report published
May 2009	Data.gov launched in the US
December 2009	Putting the Frontline First strategy commits government to greater openness in the release of data. The report also establishes a Local Data Panel to focus in the release of local authority data.
January 2010	Data.gov.uk officially launched and made available to all. London also launches datastore for capital.
April 2010	Ordnance Survey releases significant GeoData as open data following a long “Free our data” campaign by the Guardian, and consistent calls from developers for better access to geodata.
May 2010	Prime Minister David Cameron sends a letter to government departments asking for specific action on opening up government data and establishes a new Public Sector Transparency Board
June 2010	COINS dataset released
September 2010	Open Government License established
November 2010	COINS and central government expenditure data released
January 2011	Deadline for local government expenditure data release
February 2011	Protection of Freedoms Bill suggests several key amendments to the Freedom of Information Act
August 2011	Public consultation on open data policies
October 2011	Chancellor’s Autumn Measures, announcement of Data Strategy Board, Public Data Group and Open Data Institute

Foreword from the Open Knowledge Foundation

Open data has been hailed as one of the most important public policies of our time, and the potential impacts of sharing such data cooperatively are enormous. Its impacts are already visible in our societies today, but there remains a need for further research in order to pinpoint the exact benefits of such information transformations and how they affect us as individuals.

It is because of this fact that our community is so pleased that the Finnish Institute in London has chosen open data as one of its focus points and that they have commissioned this report on British open data policies. It is important to look at issues from a slightly different angle than we have in our own work, and this report contributes to existing debates about transparency and freedom of information in the UK.

Another key element of the report is its emphasis on the importance of accessibility both in terms of open data and access to information. Without equitable access points there cannot be general use - and without use, no general benefits. The report also calls for an increase in user-driven policies, noting that the inherent value of openness as a concept lies in its promotion of co-operation amongst individuals from a variety of diverse backgrounds.

In 2012, it is our hope that the opening up of data will become a larger priority for local, national and international legislation, and that the Open Government Partnership will allow new governments to endorse transparency. This September, hundreds of participants will gather in Helsinki for the first Open Knowledge Festival in order to collaborate and build new projects that explore openness in creative ways.

It is becoming increasingly evident that open data is more than mobile apps - and as Halonen argues, it is crucial to address both how we want digital information landscapes to take shape and how members of civil society can be empowered within such infrastructures. We believe this report provides a significant contribution to this dialogue and look forward to continuing the research and debate it introduces.

Dr Rufus Pollock, Co-founder and Director of the Open Knowledge Foundation

Preface

In April 2010 while working for the Finnish Institute in London, I organised a small morning seminar under the title “Digital Approaches to widening civic participation – cases Open Knowledge and Digital Humanities” with Dr Rufus Pollock, co-founder of the Open Knowledge Foundation, giving a keynote lecture on open data. This was in the immediate aftermath of the launch of data.gov.uk, a now-famous data portal where the UK government planned to release all relevant public datasets for free re-use. Open data was still a very new concept with little literature available on the subject.

Luckily, the Institute understood the importance of the topic. During the following year, it was discussed how we could further facilitate both open knowledge and open data in Finland. Finally, late in the summer of 2011, open knowledge and open data were announced as focus areas of the Institute’s society programme, with this report being commissioned as the first flagship project that would launch the wider programme on knowledge society and its various implications. I am glad to see that the partnerships that started to emerge immediately after that initial seminar in 2010 have recently begun to result in concrete actions.

Open knowledge is a fitting concept for the Institute since the connection between Finland and openness has far deeper historical roots than one might realise. The first ever law on information publicity already had a strongly Finnish touch to it, as the key figures whose thoughts and writings formed the basis of Sweden’s 1766 law on freedom of the press were, in fact, from those parts of Sweden that is nowadays known as Finland. Subsequently, the first 20th century freedom-of-information law was the Finnish law on publicity of government affairs in 1951.

Patriotic historical narratives aside, transparency and openness are increasingly important concepts in a highly networked and interrelated information society. It is not only a question of holding politicians and public officials to account or transferring data seamlessly within various data management systems, it is a question of values and even the core structure of our democracy. Who, in fact, is allowed and able to participate in the public sphere, which is increasingly taking a digital form? Free flow of ideas is considered necessary for new innovations and

solutions for renewing public services, but it is also key to interaction, communication and shaping the contemporary world.

From my perspective, an open society reflects the universal values of intellectual autonomy, equality and trust. At best, these values can lead to a community where each member can think and work as she sees fit and share both ideas and pieces of work without fear of discrimination or judgement because of her thoughts and without fear of someone exploiting her work for petty personal gain. Moreover, she needs to see that her interests are being treated equally in the community.

I hope that this report clarifies the somewhat blurry picture of open data in the contemporary society. It is still such a new concept that many readers may not yet have any overall picture of how data has been applied and what has been the main development process behind 21st century open data policies. I understand that there are also issues that have been left out of this report due to the timescale of the research process. There is still a huge task ahead in studying all the implications of open data and in compiling all the relevant issues into a single literary work. With that in mind, I hope that this report provides the catalyst for some new ideas and projects in the field of open data and open societies.

I wish to thank all my interviewees and everyone who has contributed to this work for their time and kindness in providing their ideas and comments, without which this report would have never materialised. You know who you are.

Holborn, 25 January 2012

Antti Halonen

Executive Summary

This paper presents an analysis of the recent UK open-data policies and draws an argument on how governments can sustainably promote the development and use of open data. Moreover, research contributes to the ongoing discussion on the normative values of openness by presenting a conceptual analysis of open data as an integral part of the freedom-of-information continuum.

The key findings can be listed as follows:

- Key to benefits is the quality of user engagement
- Open data and its objectives should be addressed as a part of the freedom-of-information continuum
- The decision to emphasise the release of expenditure data was not ideal: governments do not know best what kind of data people want to have and should aim at releasing it all
- Leadership, trust and IT knowledge are crucial, not only political leadership but within organisations too
- The social and democratic impacts of open data are still unclear and in future there is a need for sector-specific research

Research was conducted from September 2011 to January 2012, mostly by semi-structured interviews of key experts and analysis of the government policies. Starting point for the research was to primarily address the applicability of open data: how data is being used and what kind of benefits is it possible to identify from the data use. During the research process it became evident, however, that open data as a concept is so diverse that a mere analysis of data use would be insufficient in order to gain a deeper understanding of the phenomenon.

Open data is applied in various ways with lots of small-scale success stories available, mostly in the form of mobile-phone or web applications. These apps and websites – as innovative and useful as they are – are yet not the key issue when addressing the overall value of open data. These services make everyday life of citizens a tiny bit easier, and when accumulated they may result in significant economic benefits. However, the open-data community has also been vocal about the potential positive impacts on democracy. These impacts are significantly harder to identify and need

much more research in order to produce comprehensive and reliable results. In addition, we must realise the difference between transparency and democracy-oriented goals that are usually associated with the freedom-of-information movement and the technology and innovation-oriented goals of the open-data movement.

Key to further benefits, whether economic or democratic, is more education and improved user engagement – of both citizens and public officials. The level of knowledge and understanding of open data is currently rather low, and most data producers don't yet see the potential benefits that lie in open data. Equally, individual citizens are not necessarily capable of using datasets as the threshold for accessing and using raw datasets effectively is, at times, quite high. The best examples, in fact, are those where the data-portal interfaces are made as simple and easy to use as possible by providing relevant context to data and equally where data users are already engaged in public participation, be it within the public sector itself or some organised civic-society movement. Consequently, it is only the data user herself who knows what kind of data would be most useful. A certain service-design approach would be desirable.

There are already examples available where companies have benefited commercially from data and where public-sector organisations have gained efficiency benefits. In the future, it is more important to focus on the normative side of open data and on its potential impacts on democracy. There is a risk of creating a hollow mantra of open data improving the level of democracy without any evidence provided. However, the potential for great improvement in democratic accountability is there, and in some cases there is already sufficient evidence.

All these benefits require the threshold for accessing, understanding and using the data to be as low as possible. In order to achieve this, the data producers must possess a certain level of ICT knowledge to implement the system so that it is both simple enough to use and sophisticated enough to be able to manage information flow comprehensively. In many cases, the ICT and data-management infrastructure is not sufficient, and organisations lack the human resources to renew it so that it matches the requirements of openness. This should not be an excuse for not to release data, however, but a wake-up call for both data providers and open-data community alike.

Finally, it should be made clear that open data is not apolitical initiative. There is a strong political side to it, which dates all the way back to the long development process of governmental transparency in the UK. The initial focus on the release of expenditure data is claimed to be driven by political motives, and in terms of development of sustainable and productive use of data it was not necessarily the right decision to make. The discussion of open data was sidetracked when the focus was on the rift between local and central government and not on how public-sector organisations and civic communities could benefit from data. For many local-government data producers, the whole open-data initiative is equivalent to the £500 expenditure-data agenda and hence they don't necessarily see the wider context and potential benefits that might lie in open data. Bearing this in mind, the open-data community should be wary of arguing too eagerly in favour of open data improving the general access to information. Open data at its current form is mostly a target-driven policy without the reactive pull-factor that is essential for a political right that is freedom of information.

1. Introduction: Objectives and methodology

1.1 Objectives

We live in a data society. The sheer amount of digital information has increased exponentially during the last few years, and the friction in creating and disseminating that information has vanished over the same period. In addition, many of the key infrastructures of society are increasingly in the form of digital data. Whether this is an inconvenient truth or a nerd's dream that has come true is in the eye of the beholder of course, but nevertheless we must face the situation and make the best of it.

How to manage this vast array of information is one of the critical questions of our time. National legislations and public officials are not necessarily up to date in terms of creating feasible frameworks for data management and the capabilities for handling digital data properly. The question also touches on the issue of who is allowed to use and re-use data and who has the best capabilities to make something useful out of it.

Data is said to be the new raw material of the 21st century: a resource that enables unimaginable new services and products and that is just waiting to be exploited.² A big debate is going to take place about the rights for data and how to establish an ecosystem that benefits society in the widest and most equal manner possible. Therefore, we must not be overawed by the potential economic impacts of data, but we must remain aware of the political and social implications, both positive and potentially negative. At the same time, we should not be completely blinded by the hype that takes place around data. It is only natural that we tend to overestimate the importance of new innovations of our time – arguably every generation believes that it is living through a fundamental period of change. Moreover, open data is not as new a concept as we might think: it is claimed to have helped Florence Nightingale in her efforts to revolutionise nursing in the 19th century!³

² Berners-Lee, Tim & Shadbolt, Nigel: *There's gold to be mined from all our data*, <http://eprints.ecs.soton.ac.uk/23090/1/Times%20OpEd%20TBL%20DNRS%20Final.pdf>

³ Ibid.

Talk of data as the “new coal” or even as the “new gold”⁴ is just one example of the hype that has been around open data recently. The number of reports focusing on open data has increased rapidly, with most of them emphasising the economic potential of data. However, the very definition of open data is crucial, as a certain danger looms where the concepts “big data” and “open data” become mixed. The large international consulting agencies Deloitte and McKinsey have both endorsed data as a medicine for growth^{5,6}, and the individual researchers Graham Vickery⁷ and Heli Koski⁸ have reviewed and estimated open data’s impact on economic growth respectively. With regards to the social side of open data, one of the leading British think tanks, Demos, has very recently launched a research programme on open societies, where open data has been chosen as one of the most important factors.⁹

What this report is not?

This report is not a hands-on guide on how to open up public datasets. There is plenty of guidance available, written by people more competent in technical issues than the author of this report is.¹⁰ Nor is this an analysis of the quality of ontologies or other technical aspects of published data or a review of open-data business models or advice for enterprises on how to exploit open data commercially.

This report is not a purely ethnographic narrative of the open-data community, either, despite the fact that a large share of this report is devoted to interviews and reviews of case studies. This partly ethnographic approach has enabled the analysis of cultural and human factors behind the social movement of open data.

4 Kroes, Neelie: *Opening Remarks, Press Conference on Open Data Strategy*, Brussels, 12 December 2011, <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/872&type=HTML>

5 *Unlocking Growth – How open data creates new opportunities for the UK*, Deloitte, 2011, http://www.deloitte.com/view/en_GB/uk/market-insights/deloitte-analytics/2d376cdde7e14310VgnVCM1000001a56f00aRCRD.htm

6 *Big data: The next frontier for innovation, competition and productivity*, McKinsey Global Institute, 2011, http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation

7 Vickery, Graham: *Review of recent studies on PSI re-use and related market developments*. Information Economics, Paris, 2011.

http://www.epsiplatform.eu/news/news/review_of_recent_psi_re_use_studies_published

8 Koski, Heli: *Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth?* The Research Institute of the Finnish Economy, September 2011,

<http://www.etla.fi/eng/julkaisuhaku.php?type=details&id=1853>

9 Leadbitter, Charles: *The Civic Long Tail*, September 2011,

<http://www.demos.co.uk/publications/theciviclontail>

10 See, for example, Poikola, Antti; Kola, Petri; Hintikka, Kari: *Public Data: an introduction to opening information resources*, <http://www.scribd.com/doc/57392397/Public-Data>

However, what this report aims for is to present a factual and timely picture of the status of open data in the UK and to form an argument on the development process of those policies that have intended to support open data. This would not have been possible without a certain historical perspective or a critical approach in analysing the implementation of public policies and the initial goals of the open-data movement itself. The report aims to get under the skin of the whole concept of open data: what it is about and why it is important.

Moreover, this report makes a contribution to the normative discussion on openness in our contemporary society and whether open data can play a crucial role in it and under what terms. In the end, I hope to have raised some topical questions for future discussion on open data and given some potential answers to some of them.

The time scale during which this report has been compiled was not generous enough for a comprehensive longitudinal approach on the applicability of data. Also, during the research process it became evident that the crucial issues that need further clarification are not necessarily the practical examples of how open data is being used for business purposes but the whole epistemological nature of the open-data movement and its connection to society and democracy at large. There are plenty of good examples of how data is being used in innovative ways, and at this time I feel there is no immediate need to present them more than is necessary to support the main argument on the benefits of open data. However, this report presents, on one hand, a framework that can be applied in future research and, on the other, an ethnographic review of the open-data environment and some of the most critical questions it has faced so far and will inevitably face in the future.

I hope that this report gives insight on what the discussion on open data has really been about and what we can expect to happen in the near future. The whole question of transparency and openness has so far been relatively vague, and in order to do further research on the subject it is essential to clarify the concept a bit.

The primary objective of this report is thus to present an accurate picture of the development and status of open data in the UK and also to review the recent open-data policies and how they have supported the applicability of data.

The United Kingdom has so far been seen as a pioneering nation in the scope of open data due to its new, progressive data-licensing and re-use policies.¹¹ Moreover, open-data policies in the UK have been in place for a relatively long time, which makes it a unique research subject on the European scale. The research therefore aims to take advantage of the UK experiences and form conclusions that may be applied in other countries, particularly in Finland.

The secondary objective is to facilitate the uptake of open data and formulation of best practices. We hope to be a catalyst for further analysis and discussion on open data in Finland, and also elsewhere.

In short, the study aims to answer the following questions:

- How do the existing open-data policies support the applicability of open data and what kind of benefits have already been achieved?
- How does open data fit the development of freedom-of-information legislation?
- What are the most important social implications of open data and how should they be addressed in the future?

1.2. Methods

Since open data is still such a new phenomenon – despite Nightingale and her efforts – at the moment, it is more important to generate a certain depth of understanding, rather than breadth. This can be achieved mainly via in-depth interviews of various key players in the field of open data – both critical and supportive – and also by selecting a few application areas of open data for case-study purposes.¹²

Before explaining the methods in more detail, a few words about the author are needed. Despite aiming for objectivity, it is highly questionable whether author can ever completely achieve that goal. I have argued for more transparency and more

¹¹ Schellong, Alexander & Stepanets, Ekaterina: *Unchartered waters – the State of Open Data in Europe*, CSC Public Sector Study Series 1/2011, http://assets1.csc.com/de/downloads/CSC_policy_paper_series_01_2011_unchartered_waters_state_of_open_data_europe_English_2.pdf

¹² Denzin, Norman K. & Lincoln, Yvonna S (edit.): *Strategies of Qualitative Inquiry*, Sage Publications, Thousand Oaks, 1998.

effective use of public information already before I was commissioned to conduct this research, and this set up must be taken into account when reading this report.

Research data is gathered mainly by a responsive, semi-structured interviewing method.¹³ Another data-collection method is participant observation, where the researcher follows the discussion in a string of situations instead of just isolated interviews. This method provides an extensive basis of information for interpretation and analytic use.¹⁴ Participant observation is a particularly useful method in the case where the subject is in constant change and new ideas and debates emerge continuously, as is the case with open data. Events attended for observational purposes include several conferences, seminars and other discussion events.¹⁵ Some of the research data was already gathered before the start of this particular project. This data includes the survey that was conducted for all English local authorities on their perceptions of the recent transparency agenda and that was used as a part of the author's MRes dissertation last year. This data is mainly used in the case study on expenditure data in Chapter 5.

The interviewees were selected based on their relevance for the overall topic. The presupposition is that at this stage of open-data development, the best source of information is those people who are involved in either the policy-making process or in shaping the general discussion in some other way, hence the decision to interview people with different backgrounds and different approaches to open data. This approach explicitly excludes quantitative methods and is more narrative in nature, but provides a way for describing successful cases where open data has been applied. Moreover, with interviews I have tried to open up some of the arguments behind the open-data movement to a reader, who does not necessarily have any background knowledge on the topic.

Interviewees have their own opinions and approaches to open data and therefore they are given freedom to express views frankly. For this reason, I have only quoted them with a vague definition of their job titles or descriptions. All interviewees are experts of their subjects and this certain degree of anonymity protects their views from

¹³ Rubin, Herbert J. & Rubin, Irene S.: *Qualitative Interviewing: The Art of Hearing Data*, Sage Publications Ltd, Thousands Oaks, 2005

¹⁴ Becker, Howard S. & Geer, Blanche: *Participant observation and interviewing*, in Seale, Clive (ed.): *Social Research Methods – a reader*, Routledge, New York, 2004

¹⁵ List of events attended [at](#) the end of this report

unnecessary comparison of the opinion with the name. With anonymity provided, the arguments have room to speak for themselves.

With interviews I have also aimed at speaking with a voice of the open-data community itself, thus eliminating the approach of a researcher as the sole intellectual authority. As is argued by Kalela¹⁶, for instance, the researcher should foremost act as an interpreter between two different cultures, those of the subject and the audience. In order to achieve in this, she should primarily be able to engage in a dialogue with the subject. However, researcher can never fully disintegrate herself from the subject and the society she is part of, therefore a purely objectivist approach is not desirable for even a contemporary topic such as open data.

The study is intrinsically cross-disciplined, combining elements from qualitative social sciences, historical studies, media studies and information studies. Qualitative literature and policy analysis provide a means for assessing how open-data policies have been driven, what the historical framework of open-data policies is and what kind of research there already is on the topic.

I will put forward my argument by first briefly reviewing the existing research on open data: what it is all about and why it is such an important topic. I summarise the results of this literature review in chapter 2. In chapter 3 I present a short narrative of the development of transparency in the UK and how open data fits this continuum. Before continuing to the benefits of open data and emerging problems (chapters 5 and 6), I analyse the exiting legal and political frameworks in order to provide an institutional context for open data (chapter 4). This is a crucial part of the research in order to present comprehensive conclusions on the overall societal impacts of open data.

¹⁶ Kalela, Jorma: *Making History. The Historian and Uses of the Past*. Palgrave Macmillan, 2012

2. Open data – What and why?

2.1. What is open data?

The term “open data” usually refers to non-personal data that is accessible to all and can be freely used, re-used and distributed by anyone. Re-use of data is made possible by releasing data in machine-readable format and under such a licence that typically allows both commercial and non-commercial use. Typical examples are datasets that are created by public authorities, but open-data principles may be applied to private datasets, as well.

There are several widely used definitions for open data. The Open Knowledge Foundation, an influential British NGO advocating governmental transparency in the digital age, sums up the definition of open data as follows:

“Open data is data that can be freely used, re-used and redistributed by anyone – subject only, at most, to the requirement to attribute and share-alike.”¹⁷

The UK government Public Sector Transparency Board, however, has given the following definition for public data:

“Public data is the objective, factual, non-personal data on which public services run and are assessed, and on which policy decisions are based, or which is collected or generated in the course of public service delivery.”¹⁸

Despite these definitions, conceptually the open-data movement derives from a wider epistemological background. At minimum, three different aspects of openness can be listed, as follows:

- 1) Technological openness
- 2) Non-proprietary openness
- 3) Legal openness

¹⁷ See <http://opendefinition.org>

¹⁸ New Public Sector Transparency Board and Public Data Transparency Principles, <http://data.gov.uk/blog/new-public-sector-transparency-board-and-public-data-transparency-principles>

Firstly, in technical terms, open-data discussion is typically taking place around concepts of machine-readability, the semantic web and linked open data, which further emphasises the connectivity of data and possibilities to create so-called smart applications. The essential issue in terms of technological openness is thus the question of machine-readability.

The basic principles of linked open data are typically listed in a five-star model as follows:

- * Data is available on the web (in whatever format), but with an open licence
- ** Data is available as machine-readable structured data (e.g. in Excel, instead of an image scan of a table)
- *** As in two stars plus non-proprietary format (e.g. CSV instead of Excel)
- **** All the above plus use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at things created by others
- ***** All the above, plus linking your data to other people's data to provide context ¹⁹

Non-proprietary openness, or openness as an ideology, derives from the demand for interoperability and inclusivity, which also draws from the creative value of sharing and utilising common resources. Eric Raymond argues in his influential essay *Cathedral and bazaar* that the creative process of hackers and software developers mainly derives from the combination of gaining a personal reputation by creating “cool” new features and the virtually seamless cooperation that takes place between developers.²⁰ An intimate connection between the open-data movement and the earlier open-access movement is clearly visible.

In terms of government data, the argument goes that data that is created by public funds should be free for all to use and re-use. Data is thus seen as a common resource that does not suffer from scarcity. It is seen as morally wrong to restrict the use of data to just a certain group of people. The international “right to data” initiative derives from this setup and can be seen as a tail of the long-standing freedom-of-information movement.

¹⁹ Five-star model adapted from <http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/>

²⁰ Raymond, Eric: *The Cathedral and the Bazaar*, 2000, <http://catb.org/~esr/writings/homesteading/cathedral-bazaar/cathedral-bazaar.ps>

The idea of legal openness is linked to the licensing of data. In Chapter 4, I will put forward an argument on the necessity of the institutional basis of open data, in which licensing is a major determinant. In short, data must be licensed under such a licence that recognises the user's right to exploit data in a variety of ways, including commercially.

All these aspects of openness are effectively compiled in an eight-point list, which has been adopted by open-data advocates and governments alike²¹:

1. Data must be **complete**: all public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.
2. Data must be **primary**: data is collected at source, with the finest possible level of granularity, not in aggregate or modified norms.
3. Data must be **timely**: data is made available as quickly as possible to preserve the value of the data.
4. Data must be **accessible**: data is available to the widest range of users for the widest range of purposes.
5. Data must be **machine-processable**: data is reasonably structured to allow automated processing.
6. Access must be **non-discriminatory**: data is available to anyone, with no requirement of registration.
7. Data formats must be **non-proprietary**: data is available in a format over which no entity has exclusive control.
8. Data must be **licence-free**: data is not subject to any copyright, patent, trademark or trade-secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Paul Clarke has identified four broad types of datasets that are published by public bodies²²:

21 Open Government Data: eight principles of open government data, <http://www.opengovdata.org/home/8principles>; explaining remarks adapted from Schellong, Alexander & Stepanets, Ekaterina: *Unchartered waters – the State of Open Data in Europe*, CSC Public Sector Study Series 1/2011,

http://assets1.csc.com/de/downloads/CSC_policy_paper_series_01_2011_unchartered_waters_state_of_open_data_europe_English_2.pdf

22 Clarke, Paul: *There's data, and there's data*, <http://paulclarke.com/honestlyreal/2010/06/theres-data-and-theres-data/>

-
1. **Historical data:** information on what has happened in the past
 2. **Planning data:** what is projected to happen
 3. **Infrastructural data:** static information on society: postcodes, opening hours, organisation structures, etc.
 4. **Operational data:** public-transport timetables and other real-time information

These four types of definition are useful in terms of analysing the applicability of data – for whom different types of data are most useful and why. It also helps in further categorisation of datasets. One of my main arguments is the importance of conducting sector-specific research, for which these types of definitions would probably be useful.

There is an ever-increasing list of applications and web services that are created by applying open data: from live tube timetables to services that indicate the level of crime in the area where a user is walking. These apps make the user's everyday life more convenient and are increasingly important in their financial value. More information on different apps and services can easily be found online, and so this report does not concentrate on these examples.

2.2. Why open data?

Open data is intrinsically a combination of various different things and thus is associated with different objectives and benefits for different groups of people. There is not any single pattern of goals but various interrelated application areas, which together form a compilation of objectives and potential benefits.

Moreover, as Rufus Pollock, for instance, argues, open data is primarily a means to an end, not an end in itself.²³ The objectives of open data are thus related to goals set – be it in the field of economics, democracy or public services – but data openness *per se* does not necessarily have significant value. The question of the intrinsic value of openness is a substantial philosophical issue, and it will not be tackled here in detail.

²³ Pollock, Rufus: *Open Data: a means to an end, not an end in itself*, Open Knowledge Foundation Blog 15 September 2011, <http://blog.okfn.org/2011/09/15/open-data-a-means-to-an-end-not-an-end-in-itself/>

At the moment, it is difficult to predict which particular application area will rise into an outstanding position in the future. It seems that new potential application areas are popping up rapidly. Nevertheless, it is still very early days, and it looks like the research so far has been mostly a combination of evidence and expectations, as Jonathan Gray, for instance, has pointed out.²⁴ Moreover, history shows that predictions on the applicability of various social and technological innovations have often been rather inaccurate.

Research literature, however, suggests that there are a few emerging areas where open data could potentially be most beneficial. These objectives are expected to release both internal and external benefits.

Internal objectives are those that would enhance an organisation's own work in some particular way or form. These benefits include better efficiency, increased internal understanding of the organisation's work and objectives and releasing of resources into more productive tasks.

On the other hand, the open-data movement is also expected to provide value for society at large. These *external* objectives arise mostly from supporting entrepreneurship, empowering citizens, and democratic accountability and participation. In short, open data is seen as a tool for a more democratic society, an improved economy and empowered citizens.

The joint feature of both sets of objectives is the perception that open data has provided opportunities for the general public to use public information in a way that was not possible before.²⁵ In comparison to reactive freedom of information, the open-data movement – or two-way online transparency, as it is sometimes called – is seen as enabling a participative *writing society* instead of a *reading society*, where citizens are theoretically able to receive information but not to re-use it in creative ways.

24 Gray, Jonathan: <http://jwyg.okfn.org/2011/12/06/interview-for-university-of-southampton-open-data-study/#comments>

25 See, for instance: Davies, Alysia & Lithwick, Dara: *Government 2.0 and Access to Information: Recent Developments in Proactive Disclosure and Open Data in the United States and Other Countries*, Library of Parliament of Canada, 15 April 2010, <http://www2.parl.gc.ca/Content/LOP/ResearchPublications/2010-15-e.pdf>

In the UK, the central government has particularly emphasised two key areas that open data could benefit: accountability and entrepreneurship.²⁶ In addition, there are reports that suggest wide-scale benefits that open data could provide in public service provision.²⁷ These objectives play a key role in the recent Cabinet Office white paper on the future of public services.²⁸ More on the current government discourse is analysed in Chapters 3 and 4.

The emphasis of the UK central government is in accordance with the findings of Noor Huijboom and Tijs Van den Broek, who have compared open-data policies of five different countries where open-data strategies have already been implemented: the United Kingdom, United States, Spain, Denmark and Australia. They conclude that the following three primary motivations can be pointed out for publishing government data (Figure 1):

- 1) Increasing democratic control and participation
- 2) Fostering service and product innovation
- 3) Strengthening law enforcement²⁹

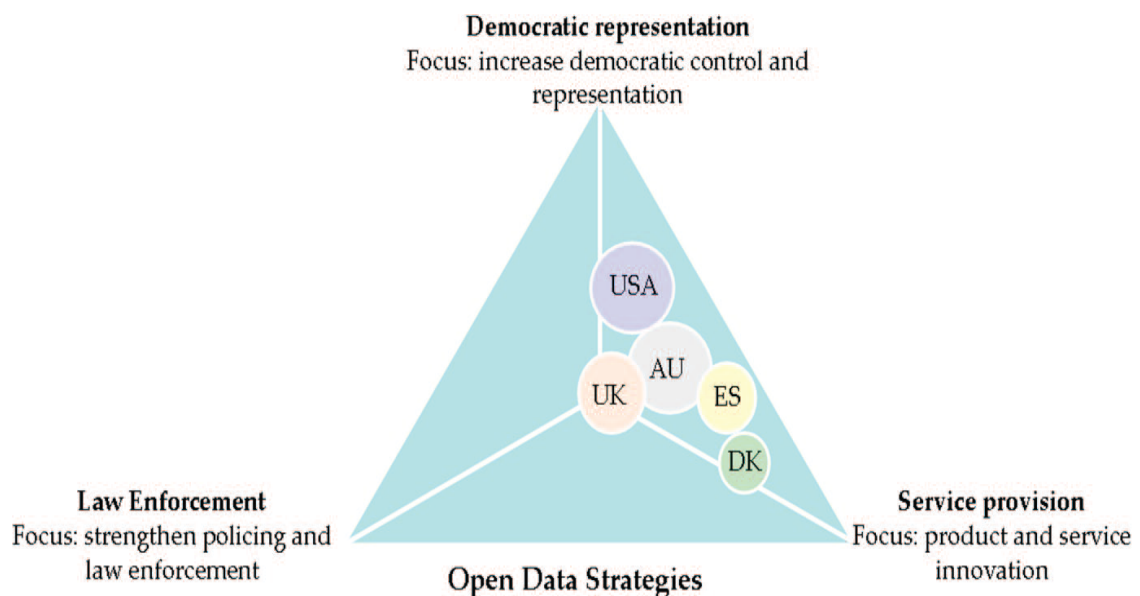


Figure 1: overview of open data strategies (Huijboom & Van den Broek, p.5)

26 See, for instance: Transparency Board Minutes, 11 January 2011, <http://data.gov.uk/blog/transparency-board-minutes-11th-january-2011-0>

27 2020 Public Services Trust: *Online or In-line: The future of information and communication technology in public services*, March 2010, http://clients.squareeye.net/uploads/2020/documents/online_or_inline.pdf

28 Cabinet Office White Paper: *Open Public Services*, July 2011, <http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf>

29 Huijboom, Noor & Van den Broek, Tijs: *Open data: an international comparison of strategies*, European Journal of ePractice, No. 12, March/April 2011, http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012_1.pdf

According to Huijboom and Van den Broek, all five countries clearly prioritise either the participative or service-innovation aspect.³⁰

The next few pages are dedicated to identifying the core objectives that can be singled out from the research literature and open-data discussion.

2.2.1 Efficiency

A common expectation throughout the open-data community is that by opening up their datasets to public scrutiny, organisations can expect to gain significant efficiency savings.³¹ The reasoning behind this argument derives from the idea that through scrutiny, enabled by transparency, it is easier for vigilant citizens to identify wasteful behaviour in public-sector organisations. Another argument states that by opening up expenditure information, an organisation's employees themselves gain better understanding of the organisation's financial situation and can thus act accordingly.

The UK government has relished the emergence of “armchair auditors”, that is a group of interested citizens who can easily hold public officials to account by scrutinising datasets of public spending. Community Secretary Eric Pickles (Con) has stated the following:

"The public should be able to see where their money goes and what it delivers. The swift and simple changes we are calling for today will unleash an army of armchair auditors and quite rightly make those charged with doling out the pennies stop and think twice about whether they are getting value for money."

32

There is already some research to back up the expectations of improved efficiency. Publishing of expenditure data in the state of California, for example, has reportedly

30 Ibid.

31 See e.g. PM David Cameron's letter to Government departments, 31 May 2010, <http://www.number10.gov.uk/news/letter-to-government-departments-on-opening-up-data/>

32 Secretary of Communities and Local Government Eric Pickles: New era of transparency will bring about a revolution in town hall openness and accountability, 4 June 2010, <http://www.communities.gov.uk/news/corporate/1606882>

led to efficiency savings of \$20m.³³ In the UK, there are reports of data transparency enhancing the performance of those NHS organisations that have opened up their data to public scrutiny.³⁴ In Canada, open data is claimed to have helped in revealing one of the biggest tax frauds in the country's history.³⁵

Another line in the efficiency argument is related to the possible savings that could be made by managing freedom-of-information requests more efficiently. According to David Eaves, the costs of complying with freedom-of-information requests in Canada alone have increased rapidly during the last few years, and the number is constantly growing.³⁶ The number of requests has, in fact, been growing throughout the English-speaking post-industrialised world, and the average cost per request is estimated to vary between £200 and £800, depending on the country.³⁷ With proactive publishing of open data, however, citizens are expected to find the information they want without the need to submit a freedom-of-information request.

2.2.2 Democratic accountability

In a wider sense, I regard democratic accountability as including both transparency *per se* and government accountability for using public funds and making political decisions as representatives of the wider demos. However, later in this report, I have separated transparency and accountability when addressing the benefits achieved.

Literature suggests that there are varying views on whether technologically more sophisticated ICT-environment and e-government initiatives have led to a more transparent government. Cory Armstrong has linked the development of e-government and online transparency in the USA to the level of professionalism and ease of information accessibility on websites. Not only is the type of information placed online important in terms of determining the level of transparency, but also where and how it is placed online. She also argues that the level of online

33 UN e-Government Survey 2010, <http://unpan1.un.org/intradoc/groups/public/documents/UN-DPADM/UNPAN038853.pdf>

34 See, for instance, <http://www.ucl.ac.uk/constitution-unit/events/tim-kelsey-presentation>

35 Eaves, David: Case study: *How open data saved Canada \$3.2 billion*, <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>

36 Eaves, David: *Access to Information is Fatally Broken... You Just Don't Know it Yet*, <http://eaves.ca/2011/03/30/access-to-information-is-fatally-broken-you-just-dont-know-it-yet/>

37 Colquhoun, Anna: *The Cost of Freedom of Information*, The Constitution Unit, December 2010, <http://www.ucl.ac.uk/constitution-unit/research/foi/countries/cost-of-foi.pdf>

transparency was greater in communities with a higher proportion of Republicans.³⁸ In terms of the political discourse that takes place around transparency and other political implications, it might be interesting to assess whether similar trends occur in a UK context, but unfortunately it is outside the scope of this study.

Armstrong's findings that professionalism and accessibility enhance online transparency are countered by Vicente Pina, Lourdes Torres and Sonia Royo, who argue that at the level of the European Union, ICT initiatives have not yet had any dramatic impacts on government transparency.³⁹ In a sense, Pina et al are members of the same continuum as Darrell West, who initially argued that e-government has not met its main objectives of transforming service delivery and public trust in government⁴⁰, and perhaps even Frank Bannister and Regina Connolly, whose conclusion is that expectations of e-government-powered transparency might have been too high so far.⁴¹

I argue that open-data initiatives are intrinsically related to this wider context of e-government and online transparency, and therefore the basis for this research has consisted of a literature review of relevant studies on e-government. In Chapter 4, I will consider further the historical and conceptual development of open data and transparency.

The accountability arguments usually state that while accountability is one of the most important aspects of democracy, it is very difficult to have it without true openness in government.⁴² Open data is seen as a means for maximising citizens' potential to scrutinise government and spot wasteful spending, hence improving accountability.⁴³

38 Armstrong, Cory L.: *Providing a clearer view: An examination of transparency on local government websites*, Government Information Quarterly, Volume 28, Issue 1, January 2011, pp. 11-16.

39 Pina, Vicente; Torres, Lourdes; Royo, Sonia : *Is e-government leading to more accountable and transparent local governments?* An overall view, Financial Accountability & Management, Volume 26, Issue 1, February 2010, pp. 3-20.

40 West, Darrell M.: *E-Government and the Transformation of Service Delivery and Citizen Attitudes*, Public Administration Review, Volume 64, Issue 1, February 2004, pp.15-27.

41 Bannister, Frank & Connolly, Regina: *The Trouble with Transparency: A Critical Review of Openness in e-Government*, Policy & Internet, Volume 3, Issue 1, Article 8, 2011, <http://www.psocommons.org/policyandinternet/vol3/iss1/art8/>

42 See, for example, Chapman, Richard; Hunt, Michael: *Open Government in a Theoretical and Practical Context*, Ashgate Publishing, Aldershot, 2006, p. 140.

43 See, for example, Pickles, 2010.

Justin Longo argues that accountability has, in fact, been the leading objective of the UK open-data initiative and that this has its roots in the expenses' scandal and the use of crowdsourcing to “unearth previously hidden examples of misappropriation of funds”.⁴⁴ Moreover, accountability is hoped to increase public trust in government, which has been one of the key objectives of reactive freedom-of-information laws.⁴⁵

2.2.3 Empowering and public participation

Enhancing public participation has been one of the most widely used objectives of both freedom of information (FOI) and e-Government initiatives, although it has not necessarily been completely achieved.⁴⁶ The share of citizens of the total population who have made FOI requests so far has been relatively low in the UK, and any significant connection between FOI and increasing participation has been extremely difficult to prove.⁴⁷ In terms of e-Government, the connection is not much clearer, with several research findings indicating conflicting results.⁴⁸ Raymond La Raja, for one, concludes that increased transparency has resulted in a decreasing level of participation, since people are not as willing to sign petitions if they believe that the petition and the list of people who have signed it will appear online.⁴⁹

However, open data, with its more interactive nature, is hoped to enable the creative re-use of information, which would enhance the opportunities for meaningful public participation. Governments who open up public data hope that enthusiastic citizens will create innovative platforms and services for public use. It is hoped that the threshold for joining the public discussion will be effectively demolished.

In the UK, citizens are encouraged to join in establishing public services and, in this development, it is argued that access to public data is crucial.⁵⁰ This kind of

44 Longo, Justin: *#OpenData: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse?*, Public Policy and Governance Review, Volume 2, No. 2, Spring 2011, pp. 38-52.

45 See, [for example](#), Hazell, Robert; Worthy, Ben; Glover, Mark: *The Impact of the Freedom of Information Act on Central Government in the UK – Does FOI work?*, Palgrave MacMillan, 2010

46 Hazell et al., 2010.

47 Ibid.

48 Andersen, KN; Henriksen, H; Medaglia, R; Danziger, JN; Sannarnes, MK; Møyfrid K; and, Enemærke, M. 2010. 'Fads and Facts of E-Government: A Review of Impacts of E-Government (2003-2009)'. International Journal of Public Administration Vol. 33 (11) 564-579

49 La Raja, Raymond: *Does Transparency of Political Activity Have a Chilling Effect on Participation?* Meetings of the Midwest Political Science Association, Chicago, March 31 – April 3, 2011.

50 Cabinet Office: *Open Public Services White Paper*, July 2011, <http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf>

empowering is also claimed to increase public trust and help citizens to value and understand public services better when they are involved in the actual provision phase.⁵¹

2.2.4 Economy and entrepreneurship

Rufus Pollock has estimated that the opening up of government datasets would create an economy with an annual worth exceeding £6bn in the UK alone.⁵² Other calculations estimate the value of open data on a global level to be up to £20bn annually⁵³ or even €28bn in the EU27 market, with rapid growth.⁵⁴

These calculations are based mainly on the expectations of a network of flourishing enterprises, which could exploit datasets commercially by creating new products and services built on public sector information. Also, enterprises and public-sector organisations are expected to reduce costs in accessing and using public information.⁵⁵

The entrepreneurial approach to open data has been particularly popular among the current UK government, whose Chancellor, George Osborne, has described the movement as an opportunity to “maximise our business opportunities at hand”. It is a logical part of the continuum, where the possible market potential of public-sector information has been emphasised since the 1990’s.⁵⁶

In terms of data use for entrepreneurial purposes, it is argued that crucial issues are charging, licensing and data format. In order to exploit data for entrepreneurial benefits, a licence, for example, must allow the use of data for commercial purposes. In the case of UK government data, this is mostly the case. Open licences are considered further in Chapter 4.

51 Ibid.

52 Pollock, Rufus: *The Economics of Public Sector Information*, University of Cambridge, 2008, <http://econpapers.repec.org/paper/camcamdae/0920.htm>

53 Speech by the Chancellor of the Exchequer George Osborne at Google Zeitgeist 16 May 2011, http://www.hm-treasury.gov.uk/press_48_11.htm

54 Vickery, 2011

55 ibid.

56 More in Chapter 3.

According to Pollock, the benefits of open data derive mainly from the marginal-cost release of data.⁵⁷ Marginal-cost release enables the small and medium-sized enterprises to exploit data, which they couldn't necessarily do if there were charges associated with the data. The loss of sales profit would then be replaced with government subsidies. Moreover, Koski has argued that the charging regime could smother small and medium sized enterprises and give a disproportionate edge to bigger companies that could afford to purchase data regardless of the fees involved.⁵⁸

Despite all the aforementioned areas of society where open data could potentially prove beneficial, arguably the most intriguing aspect of data is the fact that no-one really knows all the possible applications for the data. There would be a myriad of opportunities if all relevant data were in an appropriate open and linked format and released under open licence. The best we can really do is to expect the unexpected. Some of the different application areas are presented in this paper as case studies, in order to illustrate the diversity of data-powered civic action.

⁵⁷ Pollock, 2008.

⁵⁸ Koski, 2011.

3. Freedom-of-information continuum

Despite a certain technological emphasis in open-data discussion, an understanding of the open-data movement – as arguably is the case with all contemporary social and political phenomena – would remain hollow if the concept were not placed in a proper historical context.

Despite the boom that began in the latter half of the first decade of the 2000's, the open-data movement did not originate in a vacuum.

The historical set-up for open data includes a development process where post-industrialist democracies have gradually granted access to more and more governmental information. The reason for this development is subject to academic discussion, and the area is too wide to cover comprehensively in this research. However, I will illustrate some of the main arguments given, in order to clarify partly why open data is so important today as it is.

Firstly, this development has its roots partly in the emergence of new information technology, which has effectively lowered the threshold for producing, storing and disseminating information, and partly in social, political and economic structures, which practically depend on swift transfer of information. This idea has an intimate connection with the wider concept of the information society, as is defined by Manuel Castells, for instance. Governments throughout the post-industrialised world aim at creating digital platforms for public services, and open data has emerged as an integral part in this development process.

Secondly, the impact of the capitalist sector on public policy has been significant throughout the 20th century. Venture capitalists have always had a certain need for public information, and this relationship is evident in contemporary discussion on open data. The power of the private sector can not be overlooked in terms of the emergence of the current open-data paradigm, however altruistic some of the motivations behind open-access and open-source movements have been. The development of European Directive 2003/98/EC on the re-use of public-sector

information is a good example of emphasis on the enhancement of the market value of public-sector information, as Katleen Janssen has argued.⁵⁹

Thirdly, the emergence of post-materialist values has arguably played a role in the development of contemporary information regimes. Post-war generations have increasingly valued issues like freedom and self-expression. These are concepts that don't necessarily provide any physical security, which means they differ greatly from the values that were regarded as most important in pre-war societies.⁶⁰ One reason for the emergence of new values is the highly improved level of education, which also partly explains the growing demand for government information. Citizens demand more from their governments and don't accept secrecy as a legitimate structure of governance.

Finally, the emergence of concepts such as New Public Management, targeted transparency and drive for increased efficiency savings has further encouraged governments to become more open in terms of the amount of information they choose to publish. Tero Erkkilä, for one, argues the transparency discourse has been gradually reframed to focus mostly on the information market and government performance instead of democratic accountability.⁶¹ Proactive release of information therefore aims at meeting certain targets, in most cases public-sector efficiency. This development of targeted transparency has been particularly visible in terms of arguments concerning open data, as Archon Fung and David Weil have argued, for instance.⁶² As long as all public data is not released proactively, the decision of what data is published is arguably – at least partly – based on the goals that the data provider wants to achieve with the release.

These factors have effectively resulted in a development where information regimes based on bureaucratic secrecy have been replaced initially by reactive freedom-of-information laws and later with participative online-transparency schemes where

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

60 See, for example, Inglehart, Ronald & Abramson, Paul R.: *Measuring Postmaterialism*, American Political Science Review, Vol. 93, No. 3, September 1999

61 Erkkilä, Tero: *Reinventing Nordic Openness: Transparency and State Information in Finland*, University of Helsinki, 2010.

62 Fung, Archon & Weil, David: *Open Government and Open Society*, in Lathrop, Daniel & Ruma, Laurel (ed.): *Open Government: Collaboration, Transparency, and Participation in Practice*, O'Reilly Media, Sebastopol, 2010

citizens are expected to actively use public data and create information and services of their own. In the following chapters, I will further illustrate this process.

3.1 From reactive FOI to proactive online transparency

A significant increase in the number of FOI legislations around the industrialised and democratised world illustrates the ongoing trend of pursuit for government transparency: globally, the total number of freedom-of-information laws has grown from the inaugural Swedish law of 1766 to 11 in the mid-1980's, then to 59 at the end of 2004⁶³, and then to 86 by the year 2008.⁶⁴ According to some scholars, governmental transparency has already gained a “quasi-religious significance in the debate over government and institutional design”.⁶⁵

During the first decade of the 21st century, several countries – including the UK – introduced their plans to launch schemes for proactively publishing information online in an open - that is openly accessible and machine-readable - format⁶⁶, thus creating opportunities for the general public to re-use public information in a way that had not been possible before.⁶⁷ This can be seen as an integral part of widening the concept of transparency. There have been arguments that a fully functional transparency regime should provide both access to information and a platform to contribute to society by means of re-using public information, thus creating a two-way “writing society” instead of a mere “reading society”.⁶⁸

Jon Gant and Nicole Turner-Lee have introduced a three-generation model of evolving government transparency. Initial right-to-know policies were primarily introduced in order to prevent arbitrary governance. The second generation provided

63 Bertot, John Carlo, Jaeger, Paul T. & Grimes, Justin M: *Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for services*, Government Information Quarterly 27 (2010), p.264.; Roberts, Alasdair: *Blacked out: Government secrecy in the information age*, Cambridge University Press, New York, 2006

64 Vleugels, Roger: *Overview of all 86 FOI countries (2008)*, <http://www.statewatch.org/news/2008/sep/foi-overview-86-countries-sep-2008.pdf>.

65 Hazell et al., 2010.

66 More on open formats, see for instance: Open Knowledge Definition, <http://www.opendefinition.org/okd/>

67 See, for example. Davies, Alysia & Lithwick, Dara: *Government 2.0 and Access to Information: Recent Developments in Proactive Disclosure and Open Data in the United States and Other Countries*, Library of Parliament of Canada, 15 April 2010, <http://www2.parl.gc.ca/Content/LOP/ResearchPublications/2010-15-e.pdf>

68 See, for example. Tim Berners-Lee's speech on the next Web in February 2009. http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html

more targeted transparency policies, which mandate baseline levels of information disclosure by the government.⁶⁹

Finally, Gant and Turner-Lee present the emergence of collaborative transparency policies, which aim to leverage ICT in order to provide a platform where government may interact with citizens. According to Gant and Turner-Lee, this approach is user-centred and two way, with government playing a facilitating role to communicate information in real-time and in machine-readable formats.⁷⁰

Tom McClean argues that the institutional and historical context has had a major role in the development process of freedom-of-information regimes worldwide: different countries have followed different routes in their transparency policies.⁷¹ It is, however, unclear whether similar development patterns can be applied when analysing the emergence of open-data regimes.

“I’d be reluctant to suggest that freedom-of-information regimes and open-data regimes are similar in that regard. I think there are some really important differences between them. Perhaps the most important difference is that freedom of information is almost always seen as a political policy, or political institution. It is resisted by politicians and bureaucrats because of the political implications of losing control over access to specific kinds of information. And, it is put to use by civil-society organisations, and sometimes by journalists and the general public, because of its role in transforming the structures of power.”⁷²

Although the development process, initially for freedom of information and later for open data, is not identical by any means, they have something in common, and that is a certain institutional and historical basis. However, it is unclear how the open-data development will vary in different countries. Development of open data is more interrelated by nature, and the social and political reality in which open-data policies

69 Gant, Jon; Turner-Lee, Nicol: *Government Transparency: Six Strategies for More Open and Participatory Government*, The Aspen Institute, February 2011, p.17, http://www.knightcomm.org/wp-content/uploads/2011/02/Government_Transparency_Six_Strategies.pdf

70 Ibid.

71 Interview, an academic/researcher

72 Interview, an academic/researcher

are created is fundamentally different than the one in the 1950's and 60's when freedom-of-information policies started to fully emerge.

“The way the politics of open data – the rhetoric – in different countries continues to develop will be very different. That is because the significance of information and the release of information vary from country to country.”⁷³

Despite the different approaches to the development of transparency, some key observations can be made. Research literature suggests that there is a certain change of paradigm emerging in global freedom-of-information discourse. The concept of transparency is evolving from reactive freedom of information to participative transparency, which is increasingly taking the form of technologically sophisticated e-government platforms.

Also, it is suggested that the open-data community has emphasised more issues like data management, citizen engagement and public participation, while the traditional transparency discourse has been focused firstly on democratic accountability and perhaps later more on the financial efficiency of governance.

3.2 Tradition of transparency in the UK

The development process for freedom of information and transparency in the UK has been spread over several decades, with a string of smaller-scale reforms that have taken place before the introduction of the UK Freedom of Information Act (FOIA 2005) and the current drive for participative online transparency.⁷⁴

It is difficult to explain why the UK has emerged as the leading European nation in the field of open data, despite its somewhat dubious reputation when it comes to governmental openness and transparency. There are countries where freedom-of-information legislation has been in place for decades before the UK, but which have so far lacked the drive for open data. The issue is outside the scope of this study, but I will present some of the explanations given and present some of the major milestones in the development process of transparency in the UK.

⁷³ Interview, an academic/researcher

⁷⁴ See, [for example](#), Hazell et al., 2010

FOIA was officially introduced under the Tony Blair administration in 2000 and it came into effect in 2005. It is argued that transparency was, in fact, an integral part of New Labour's approach to public administration and governance.⁷⁵ Moreover, the Labour party had had some plans for installing freedom of information law already under the Jim Callaghan government in the 1970's, but this plan was scrapped when the Callaghan government fell down in the 1979 general election.⁷⁶

Blair (Lab) clearly had been a vocal supporter of freedom of information during his years in opposition. As a leader of the Opposition, Blair had stated that,

“Information is power and any government's attitude about sharing information with the people actually says a great deal about how it views power itself and how it views the relationship between itself and the people who elected it”

and also that,

“If a government is genuine about wanting a partnership with the people who it is governing, then the act of government itself must be seen in some sense as a shared responsibility and the government has to empower the people and give them a say in how that politics is conducted.”⁷⁷

The human mind is volatile though, as became evident when, in 2010, Blair described the establishment of FOIA as his biggest mistake as Prime Minister and describes himself with the now-famous words *“You idiot. You naïve, foolish, irresponsible nincompoop.”*⁷⁸ Blair's disgust towards freedom of information essentially derives from his fears of transparency having a chilling effect on governance: he argues that policymakers can not deliberate freely if there is a risk of their views being released into public. Similar concerns have been expressed by e.g. Sir Gus O'Donnell, the

⁷⁵ Interview, an academic/researcher

⁷⁶ See, for example, <http://www.ucl.ac.uk/constitution-unit/research/foi/what-is-foi2>

⁷⁷ Speech by Tony Blair at the Campaign for Freedom of Information's annual Awards ceremony, 25 March 1996, <http://www.cfoi.org.uk/blairawards.html>

⁷⁸ See, for instance: *BBC Open Secrets: Why Tony Blair thinks he is an idiot.* http://www.bbc.co.uk/blogs/opensecrets/2010/09/why_tony_blair_thinks_he_was_a.html

newly retired head of civil service,⁷⁹ but so far research has not found comprehensive evidence for significant “chilling” due to freedom of information.⁸⁰

The idea of freedom of information was already present in the public discussion way before the Blair government. Blair’s predecessor as Prime Minister, John Major (Con), had already introduced a non-statutory code (Code of Access 1994), whose aim was to “liberalise Britain’s strict information regime”.⁸¹ The Cabinet Office white paper from 1993 states that:

“Government believes people should have freedom to make their own choices. Information is a condition of choice and provides a measure of quality.”⁸²

Fast forward a mere 18 years and the current coalition programme reads:

“Government believes that we need to throw open the doors of public bodies, to enable the public to hold politicians and public bodies to account. We also recognise that this will help to deliver better value for money in public spending, and help us achieve our aim of cutting the record deficit.”⁸³

These two examples from the very recent past illustrate the rhetorical pursuit for open government that has already taken place in the political culture of the UK before the current drive for open data.

The governmental drive for open government has, in fact, roots in the Fulton Committee report from 1968 on the Civil Service, which highlighted the problematic approach to secrecy in public administration:⁸⁴

“The Committee considered that there was not enough contact between the Civil Service and the rest of the community. This was partly because Civil

79 See, for instance: <http://www.bbc.co.uk/news/uk-politics-16229867>

80 Hazell et al., 2010

81 Worthy, Ben: *John Major’s Information Revolution? The Code of Access ten years on*, Open Government: a Journal of Freedom of Information, Volume 3, Issue 1, 2007, <http://www.opengovjournal.org/article/view/994/792>

82 Cabinet Office White paper 1993. *Open Government*. <http://www.official-documents.gov.uk/document/cm22/2290/2290.pdf>

83 The Coalition: *Our programme for government*, http://www.cabinetoffice.gov.uk/sites/default/files/resources/coalition_programme_for_government.pdf

84 Home Office, Committee on the Civil Service, Her Majesty’s Stationery Office, London 1968, <http://www.civilservant.org.uk/fultonreport.shtml>

Servants were expected to spend their entire working lives in the Service, and partly because the administrative process was surrounded by too much secrecy. The Committee therefore recommended greater openness in Government, less anonymity for officials, and greater mobility of staff into and out of the service.”⁸⁵

Although FOIA itself was established relatively late in the UK, the debate on access to information dates back all the way to the Official Secrets Act from the late 19th Century, that is several decades earlier than the Fulton report. Moreover, according to McClean, the competitive two-party system combined with centralisation of authorities has encouraged politicians to criticise the secrecy of incumbents. The opposition party, in particular, has always been keen to highlight the misbehaviour of the ruling party.⁸⁶

It could be argued that since the Fulton report each UK government has promised more transparency ⁸⁷, and in this continuum the current transparency agenda is by no means revolutionary. However, on the other hand, while promising more transparency the UK government has simultaneously remained highly cautious about some aspects of FOI, for example regarding the secrecy that takes place around the Royals and the Cabinet.⁸⁸

Despite the highly political debate on openness and the fact that Labour governments established both FOIA and the later data.gov.uk data portal, the idea of access to information can not be intrinsically linked to any particular political party. This is significant since the debate on open data has taken a highly politicised tone at times, as I will argue later in this report.

During the general election campaign in 2010, the Conservative party launched their plan for public-policy reforms, the Big Society, a scheme where citizens are expected to take more responsibility for provision of public services and for auditing how

85 *ibid.*

86 Interview, an academic/researcher

87 Interview, an academic/researcher

88 [A good source of more information on the parliamentary debate about openness is the website TheyWorkForYou](http://www.theyworkforyou.com/), where it is possible to find transcripts of Commons debates and oral answers to questions. <http://www.theyworkforyou.com/>

public funds are being spent. Citizen-driven auditing has been labelled as “armchair auditing”, which would be enabled by proactive release of public expenditure data.⁸⁹

According to leading Conservatives, the Big Society would revolutionise the way public services are provided, and in this revolution open government data would indeed play a crucial role.⁹⁰ Arguably, this is the first time that transparency or freedom of information has been incorporated as an integral part of the discussion on the provision of public services.

Given the historical continuum briefly explained above, it remains intriguingly unclear how sincere the Conservative party has been in its recent commitments to open data and transparency. Freedom of information enjoys popular support after the expenses’ scandal and FOI’s crucial role in it, and one of the core rules of FOI indeed is the fact that it can not be prevented any more.⁹¹

There are arguments that stress the Conservatives’ ideological wish to cut the size of the government and link open data to this goal, mostly by reducing waste and increasing efficiency.⁹² There is arguably a certain implicit assumption that the government is inefficient by nature and more transparency is the best cure. However, it remains unclear whether this assumption can in fact lead to more trustful public.

“The Cameron government is doing two things. Firstly, it is signalling to the electorate that we are trustworthy, we are giving you all the information you’ll need to judge us. We accept that it is legitimate for you to judge us and hold us to account. Look how open we are – we are trustworthy.”⁹³

It is argued that the current government is even going back to older Conservative policies that date back to Major.

“Using the proactive release of aggregated data on public-sector performance is a way of encouraging public sector to be more efficient and

| 89 Transparency Board Minutes, 11 January 2011, <http://data.gov.uk/blog/transparency-board-minutes-11th-january-2011-0>,

| 90 See, *for instance*: Speech by the Chancellor of Exchequer, George Osborne, at Google Zeitgeist, 16 May 2011, http://www.hm-treasury.gov.uk/press_48_11.htm.

| 91 Hazell et al., 2010

| 92 Interview, an academic/researcher

| 93 Interview, an academic/researcher

*effective. So there's a sort of regulation of public sector through transparency rather than authority going on."*⁹⁴

One must realise the conceptual difference between reactive freedom of information and proactive open-data transparency. In the case of open data, the government has so far been able to decide themselves what kind of data to publish: the integral pull-factor of reactive freedom of information has yet not been there. Moreover, there is still a certain lack of standardisation of procedures.

*"Although some might say that the UK is quite advanced in its policy position regarding open data, I would actually say – from a local-government perspective anyway – we're still at a relatively immature stage. Once data has been published, it's usually published in a variety of different formats, which doesn't always render data analysable. Consequently, the benefits are still a bit elusive."*⁹⁵

Open-data policies can thus be regarded as comprehensive transparency policies only when certain data standards are in place and the general public can request data they want themselves. It is a dangerous scenario, if the reactive freedom of information is replaced by open-data transparency without the possibility to request for data, because it is highly unlikely that the government is able to open up all data on a short notice. Not all politicians and civil servants are supporters of FOIA, and even the PM David Cameron has recently stated, that freedom of information is about "money that goes in and results that come out", not about "endless discovery processes".⁹⁶ It is a striking evidence on the importance of the debate that is going on about the future of FOI.

In early 2011, the government introduced a Protection of Freedoms Bill, where FOIA was amended in a way that would oblige public authorities to release electronic information in a machine-readable format if so requested. If the bill goes through in Parliament, then the UK would arguably have one of the most progressive open-data policies in the world, given that all relevant datasets were released under open-data definitions.

⁹⁴ Interview, an academic/researcher

⁹⁵ Interview, an academic/researcher

⁹⁶ PM David Cameron, HOC Liaison Committee, evidence from the Prime Minister 6 March 2012, <http://www.parliamentlive.tv/Main/Player.aspx?meetingId=10438>

3.3 Political economy and transparency – entrepreneurial roots of open data and the European PSI Directive 2003

At the European level, the start of the current drive for opening up government data was the establishment of the European Directive 2003/98/EC on the re-use of public-sector information.⁹⁷ The directive does not oblige member states to create new data but to make available the data they have created while fulfilling their own public tasks.

Janssen argues that the principal reason behind the PSI directive was the European Commission's concern about the inability of the EU to compete with the USA in the information market.⁹⁸ Moreover, the two objectives of the directive included making public-sector data available to third parties at low prices and ensuring a level playing field between public bodies that operate in the information market in competition with the private information industry.⁹⁹

The entrepreneurial roots in the UK go deeper than the 2003 PSI directive. Saxby argues that the UK government already had a certain competitiveness agenda in place in the mid-1990's, when it was considering relaxing the Crown copyright regulations in order to boost the private sector's abilities to exploit public-sector information commercially.¹⁰⁰

The European Commission continued to emphasise economy in December 2011 – quite understandably under the economic circumstances – when the Vice President Neelie Kroes announced the new European Open Data Strategy:

“Just as oil was likened to black gold, data takes on a new importance and value in the digital age. [...] We calculate that public-sector information already generates €32 billion of economic activity each year. This package would more than double that – to around €70 billion. That is not coffee

⁹⁷ European Directive 2003/98/EC of 17 November 2003 on the re-use of public sector information. Official Journal of the European Union 345/90,

http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive/psi_directive_en.pdf

⁹⁸ Janssen 2011; Saxby, Stephen: *The development of UK Government policy towards the commercialization of official information*, International Journal of Law and Information Technology, Issue 3, 1996, pp. 199-233.

⁹⁹ Ibid.

¹⁰⁰ Saxby, 1996

*money. That is, indeed, a badly needed boost to our economy.*¹⁰¹

Tero Erkkilä has undertaken research on the development of transparency discourse and found that the focus has changed from the initial democratic accountability to a more efficiency-oriented one, where government performance and financial competence have become the main goals of governmental information policies.¹⁰² With more and more uncertainty over the economy, the economy and efficiency arguments are likely to acquire an even more dominant position in the near future.

Open data clearly is an integral part of the wider freedom-of-information continuum, and should be addressed as such. The discussion should not be let to be driven by arguments of the potential economic benefits of open data, as important as they are. Access to information is a political right and further work must be done in order to improve that.

The development of open data has indeed followed some interesting routes, and it continues to take shape. We are still in the very early days of open data, and it is difficult to draw a conclusive narrative on the development process. However, in the following chapter, I will illustrate some of the contemporary issues that are taking place at the policy level.

¹⁰¹ Kroes, Neelie: *Opening Remarks, Press Conference on Open Data Strategy*, Brussels, 12 December 2011, <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/872&type=HTML>

¹⁰² Erkkilä, 2010

4. Institutional and legal contexts of open data

4.1 Institutionalisation of open data

In terms of open data emerging as a formidable public-policy reform, a crucial issue proves to be the inclusion of leading public-sector institutions in the formulation process of data policies from the beginning. Without a certain institutional push, there is a significant threat that the impact of new policy ideas will remain weak. In many countries there is a vibrant open-data community, but it is highly questionable whether this creative energy of individuals – however strong it may be – is enough to overcome obstacles that may be in the way of open data becoming a major factor in a wider social context.

I read the transparency agenda and the wider drive for open data in the UK as a public-policy reform, and it should be addressed accordingly. Eric Patashnik argues that the most efficient reforms are those that are able to change the political structure effectively and alter the landscape of the interest groups.¹⁰³ There are three principal factors that help reforms persist. Firstly, reform should break or at least undermine the existing framework of interest groups who would like to keep the status quo. Secondly, reforms that are able to eliminate the losing-side actors are likely to be more durable. Thirdly, there has to be positive feedback and a change in the beliefs and behaviour models of political elites.¹⁰⁴ In this chapter, I will analyse how the transparency agenda fits into this framework of institutional change and whether that can be regarded as something especially typical for open data.

In the UK, the initiative for open data was launched in 2009 at the highest possible level, when the then Prime Minister, Gordon Brown (Lab), announced his government's vision of opening up public datasets. The initiative was carried by his successor, David Cameron (Con), after the General Election in 2010 and the establishment of the current coalition government between the Conservatives and Liberal Democrats.

¹⁰³ Patashnik, Eric: *Reforms at Risk: What Happens After Major Policy Changes Are Enacted*. Princeton University Press, 2008

¹⁰⁴ Ibid.

The coalition speedily announced their ambitions for more open data and transparency. As with Barack Obama after his inauguration in 2009, one of the Coalition's very first announcements included the endorsement of open data and transparency in the government and other public-sector organisations. In PM Cameron's letter to all departments in May 2010¹⁰⁵, he explained the government's initial transparency commitments as follows:

Central government spending transparency

- Historic COINS expenditure data to be published online in June 2010.
- All new central government ICT contracts to be published online from July 2010.
- All new central government tender documents for contracts over £10,000 to be published on a single website from September 2010, with this information to be made available to the public free of charge.
- New items of central government spending over £25,000 to be published online from November 2010.
- All new central government contracts to be published in full from January 2011.
- Full information on all DFID international development projects over £500 to be published online from January 2011, including financial information and project documentation.

Local-government spending transparency

- New items of local-government spending over £500 to be published on a council-by-council basis from January 2011.
- New local-government contracts and tender documents for expenditure over £500 to be published in full from January 2011.

Other key government datasets

- Crime data to be published at a level that allows the public to see what is happening on their streets from January 2011.
- Names, grades, job titles and annual pay rates for most Senior Civil Servants with salaries above £150,000 to be published in June 2010.

¹⁰⁵ Letter from PM David Cameron to Government departments on opening up data, <http://www.number10.gov.uk/news/letter-to-government-departments-on-opening-up-data/>

-
- Names, grades, job titles and annual pay rates for most Senior Civil Servants and non departmental public body officials with salaries higher than the lowest permissible in Pay Band 1 of the Senior Civil Service pay scale to be published from September 2010.
 - Organograms for central government departments and agencies that include all staff positions to be published in a common format from October 2010.¹⁰⁶

In July 2011, the PM expanded this list with an announcement that the major health, education, transport and crime data is to be opened up.¹⁰⁷ According to him, the transparency agenda had now effectively entered “phase two”, where instead of “core information” government would be publishing services-related information.¹⁰⁸ In addition, the quality of data was emphasised, with the promise of plain English descriptions to be included in expenditure data, an action plan for improving the quality and comparability of data to be created by every department, and unique reference indicators to be introduced in order to enable the public to track the interaction between companies and government bodies more easily.¹⁰⁹

Another major data-release announcement was made in November 2011, when in the Chancellor’s Autumn Statement the following actions were declared for implementation¹¹⁰:

1. Linking primary and secondary healthcare datasets
2. Publishing prescription data
3. Releasing train and bus data
4. Releasing rail-fares data
5. Releasing aviation-services data
6. Releasing highways and traffic data
7. Releasing the anonymised fit-note data
8. Designing the Universal Credit ICT system and releasing aggregate-benefits data

106 *ibid.*

107 Letter from PM David Cameron to Cabinet Ministers on Transparency and Open Data, 7 July 2011, <http://www.number10.gov.uk/news/letter-to-cabinet-ministers-on-transparency-and-open-data/>

108 *ibid.*

109 *ibid.*

110 Cabinet Office: *Further Detail on Open Data Measures in the Autumn Statement 2011*, http://www.cabinetoffice.gov.uk/sites/default/files/resources/Further_detail_on_Open_Data_measures_in_the_Autumn_Statement_2011.pdf

9. Linking welfare datasets to other governmental and commercial datasets
10. Establishing the Open Data Institute
11. Releasing reference data
12. Ensuring access to personal GP records

The Autumn Statement announcement was immediately described as an example of the “sustained ongoing release of data” that gives “confidence to the open-data community and business that the process is here to stay”¹¹¹. The statement did indeed consolidate the tradition of economic arguments in the UK transparency discussion. The Open Data Institute, for example, is expected to “help business exploit the opportunities created by release of public data”.¹¹² Other key aspects of transparency seemed to be neglected, however. Tim Davies has pointed out the lack of any mention of fostering data skills in the statement.¹¹³ Moreover, there are no explicit references to enhancing the citizen’s right to access government information, which would increase the level of democratic participation or corporate accountability. The vague use of the definition of open data has attracted concerns as well. It is questionable whether private healthcare data can be considered to be a part of the open-data discussion at all. There are fears that this could lead to the government selling something as open data when it clearly does not fit the original definition of ‘open’.¹¹⁴

4.2 From the Power of Information Report to the Public Sector Transparency Board, Local Public Data Panel and Public Data Corporation

A widely accepted perception is that the initial phase in the development of open data in the UK started after the publication of the Power of Information Report in 2007. In the report, Tom Steinberg and Ed Mayo put forward suggestions to the government on how to exploit the increasing value of public-sector information.¹¹⁵

¹¹¹ Guardian, 28 November 2011, Secret Whitehall statistics to be published for the first time, <http://www.guardian.co.uk/technology/2011/nov/28/secret-whitehall-statistics-published-data?newsfeed=true>

¹¹² Cabinet Office: *Further Detail on Open Data Measures in the Autumn Statement 2011*, http://www.cabinetoffice.gov.uk/sites/default/files/resources/Further_detail_on_Open_Data_measures_in_the_Autumn_Statement_2011.pdf

¹¹³ Davies, Tim: *Evaluating the Autumn Statement Open Data Measures*, <http://www.timdavies.org.uk/2011/12/02/3090/>

¹¹⁴ Ibid.

¹¹⁵ Mayo, Ed & Steinberg, Tom: *The Power of Information*, http://webarchive.nationalarchives.gov.uk/20100413152047/http://www.cabinetoffice.gov.uk/media/cabinetoffice/strategy/assets/power_information.pdf

After the kick-start that the PoI report facilitated, an institutional basis for the open-data movement took shape when the Local Public Data Panel and Public Sector Transparency Board were set up under the Cabinet Office.¹¹⁶

The initial task of the Transparency Board was to “drive forward the Government’s transparency agenda, making it a core part of all government business and ensuring that all Whitehall departments meet the new tight deadlines set for releasing key public datasets”.¹¹⁷ The task was to be approached by “listening to what the public wants and then driving through the opening up of the most needed datasets”.¹¹⁸

In the summer of 2011, after PM Cameron had announced another round of the opening up of data, it was proposed that the Transparency Board members sit on the departmental open-data panels, which would further guide the development of the open-data strategy.¹¹⁹

At a local level, however, the institutional approach started to take shape earlier. In order to “highlight the importance of releasing public data”, the Local Public Data Panel was already established in late 2009, that is during the Brown era.¹²⁰ Key aims of the Panel included “to identify local public data and their potential uses that are likely to have the greatest impact on empowering citizens or improving local service delivery, and develop an agreed approach to their content and format”.¹²¹

According to the panel, the main data local authorities possess are as follows:

- Democratic data;
- Accounts;
- Service data;
- Performance data;
- Operational data, and policy and strategy data.¹²²

¹¹⁶ <http://data.gov.uk/blog/new-public-sector-transparency-board-and-public-data-transparency-principles>

¹¹⁷ *ibid.*

¹¹⁸ *ibid.*

¹¹⁹ Transparency Board Minutes, 8 June 2011, <http://data.gov.uk/blog/transparency-board-minutes-8th-june-2011>

¹²⁰ Data.gov.uk blog: *Local Public Data Panel – first meeting*, 26 February 2010, <http://data.gov.uk/blog/local-public-data-panel-first-meeting>

¹²¹ Local Public Data Panel: *Background paper to inform discussion on what data local authorities could make available*, 19 January 2010, <http://data.gov.uk/blog/local-public-data-panel-first-meeting>

¹²² *Ibid.*

Based on this list, the panel initially started to work by asking what data would be most helpful for citizens and the community and thus be prioritised in the opening process.¹²³ Another visible point of concern was the data guidance for local councils, which was thus a focus area for development throughout 2010.¹²⁴

In the autumn of 2010, the Local Public Data Panel considered that raising awareness of local data and associated issues was its priority area for action.¹²⁵ In relation to this, various projects were initiated during late 2010 and early 2011, the most visible of which has probably been NESTA's 'Make it Local' project. The project's aim was to "show how local authorities can work with digital agencies to unlock their data and provide really useful web-based services for their citizens".¹²⁶ The winning projects included 'Sutton Open Library' online service, which gave residents instant access to library-collection information and further developed into the 'Sutton Bookshare' project, which gives users an opportunity to register their own books in the system and lend them as a common resource.¹²⁷

In January 2011, the Government announced their plans to establish the Public Data Corporation (PDC)¹²⁸, whose primary duty would be to bring together government bodies and data into one organisation and also to:

- "Provide a more consistent approach towards access to and accessibility of public sector information, making more data free at the point of use where this is appropriate and consistent with ensuring value for taxpayers' money;
- Create a centre of excellence driving further efficiencies in the public sector;
- Facilitate or create a vehicle that can attract private investment"¹²⁹

The minister responsible for the PDC, Francis Maude, expected a full statement of the PDC policies to be published in the autumn of 2011. The statement would cover the

¹²³ Ibid.

¹²⁴ Local Public Data Panel – sixth meeting, 16 December 2010, minutes <http://data.gov.uk/blog/local-public-data-panel-%E2%80%93-sixth-meeting-16-december-2010>

¹²⁵ Local Public Data Panel – fifth meeting, 21 October 2010, minutes <http://data.gov.uk/blog/local-public-data-panel-%E2%80%93-fifth-meeting-21-october-2010>

¹²⁶ http://www.nesta.org.uk/areas_of_work/public_services_lab/make_it_local

¹²⁷ See

http://www.nesta.org.uk/areas_of_work/public_services_lab/make_it_local/assets/features/sutton_bookshare

¹²⁸ Cabinet Office blog: *Public Data Corporation to free up public data and drive innovation*, 12 January 2011, <http://www.cabinetoffice.gov.uk/news/public-data-corporation-free-public-data-and-drive-innovation>

¹²⁹ Transparency Board Minutes, 2 March 2011, <http://data.gov.uk/blog/transparency-board-minutes-2nd-march-2011>

issues of what data and agencies would be included in the PDC, how the data would be made available, under what licence and who would oversee and regulate the corporation.¹³⁰ According to Maude, issues included, “how and to what extent information would be made available for the customers” and whether data “form part of [the] organisation’s ‘public task’ or whether they are a by-product in the organisation’s business model.”¹³¹

In order to highlight the importance of public engagement, the Public Data Corporation launched a public consultation on data policies in March 2011. The purpose of the consultation was to determine data-users’ opinions of current open-data policies and expectations on how the PDC should operate.

One of the key issues initially identified was charging for data. Three possible models for the future charging regime were introduced in the PDC consultation paper¹³²:

1 Status Quo plus a commitment to more data for free: bodies within the PDC would continue to work within the existing legal and policy framework, but with a commitment to make more data available for free re-use

2 Harmonisation & simplification: some data would be made available for free; for all PDC information within the public task, there would be a single price for a particular unit of PDC information, and this price would apply to all uses of the information

3 Freemium model: data would be free at the basic level, while there would be charges for advanced features, functionality or related products¹³³

The plans to establish the PDC raised several questions on the overall feasibility and justification of the scheme. Chris Taggart, an influential blogger, a member of the Local Public Data Panel and founder of the open local-government data web-portal Openly Local, argued that the PDC might, in fact, cool down open-data development.¹³⁴ The primary fear was that the PDC might further enshrine charging practices for relevant data, which is claimed to be intrinsically against the open-data

¹³⁰ UK Parliament, Commons_ Written Answers, HC Deb, 12 July 2011, c268W

¹³¹ UK Parliament, Commons_ Written Answers, HC Deb, 4 May 2011, c791W

¹³² Cabinet Office: A Consultation on Data Policy for a Public Data Corporation.
<http://www.cabinetoffice.gov.uk/resource-library/making-open-data-real-public-consultation>

¹³³ Ibid.

¹³⁴ Taggart, Chris: *What’s that coming over the hill? Is it the Public Data Corporation?*

<http://countculture.wordpress.com/2011/01/14/whats-that-coming-over-the-hill-is-it-the-public-data-corporation/>

principles.¹³⁵ Moreover, charging can arguably encourage big corporations to use their size advantage over SME's and private individuals, who cannot necessarily afford to exploit datasets for which there is a charge.¹³⁶

Taggart has not been the only critic of the PDC. Several open-data advocates have raised their concerns on the plan. Socitm, the association for ICT and related professionals in the public sector, has stated that the establishment of the PDC was driven by the interests of a few institutions and private investors, all of whom have vested interests in charging for data.¹³⁷

It is felt that governmental bodies responsible for the PDC have not accepted Pollock's initial economical argument of releasing data at marginal cost.¹³⁸ It is claimed that the fee regime will suffocate the SME sector, which would potentially flourish due to open data. The reluctant attitude of the civil service and the vested interests of those bodies that provide data are seen as some of the biggest obstacles.

“I don't see the point of the PDC. There is no problem, so why should we create one. Keep it simple – if one public agency wants to charge and one to keep it free, you will end up in a bureaucratic mess and it will stall the development. What problem does the PDC solve? There is no problem and therefore no need for the PDC.”¹³⁹

“My first worry is that it's not clear what the purpose of the PDC is. In consultation, there is a mention of four different purposes that are contradictory and do not match with the ambitions of the 'making open data real project'”¹⁴⁰

The Local Public Data Panel has emphasised the importance of the structure of governance, an issue that is arguably neglected in the PDC consultation paper itself. It is considered very suspicious that the membership and terms of reference of the

¹³⁵ See, for instance: Open Definitions, <http://opendefinitions.org>

¹³⁶ See, for instance: Koski 2011

¹³⁷ “Socitm slams plan for Public Data Corporation”, Guardian, 27 October 2011, <http://www.guardian.co.uk/government-computing-network/2011/oct/27/socitm-public-data-corporation>

¹³⁸ Interview, a local government civil servant

¹³⁹ Interview, a local government civil servant

¹⁴⁰ Interview, an open data advocate

transition board were not included in the consultation paper and that the initial discussions on the PDC have been made behind closed doors.¹⁴¹

Furthermore, the Panel raised two fundamental questions for immediate and full clarification:

1. What is the scope of any potential charging and licensing policy
2. What charging and licensing policies should be applied to datasets outside the public task¹⁴²

The Panel recommends that the board of the PDC should include a majority from outside central government and trading funds. An open dialogue with various stakeholders is encouraged in order to “give civil servants and ministers access to a vast range and depth of knowledge, understanding and ideas from within the wider public sector, non-governmental organisations, SMEs, developers and community activists”.¹⁴³

In the consultation process, the PDC did get a relatively negative response. Respondents argued heavily against charging and also against the governance model that was initially suggested.¹⁴⁴

“Why it is needed and what does it try to solve? Secondly, whoever makes the decisions on licensing policies and charging policies, those decisions must not be made behind closed doors as they are made now. That’s ironic in terms of transparency. Clearly the PDC has had lots of impact from various different interest groups.”¹⁴⁵

“There is a possibility that trading funds would be allowed to sell more data than they are selling now and to charge for core reference data, which would

¹⁴¹ Local Public Data Panel response to the Public Data Corporation consultation, <http://data.gov.uk/blog/local-public-data-panel-response-to-the-public-data-corporation-consultation>

¹⁴² Local Public Data Panel response to the Public Data Corporation consultation, <http://data.gov.uk/blog/local-public-data-panel-response-to-the-public-data-corporation-consultation>.

¹⁴³ Local Public Data Panel response to the Public Data Corporation consultation, <http://data.gov.uk/blog/local-public-data-panel-response-to-the-public-data-corporation-consultation>

¹⁴⁴ Interview, a central government civil servant

¹⁴⁵ Interview, a local government civil servant

be regressive in terms of public ownership of data. I'm worried that whatever framework will be selected, it will be far from ideal.”¹⁴⁶

Despite the negative feedback for the PDC, some form of institutional framework is badly needed. Public sector organisations need something to lean on, and even if the PDC is not an ideal solution, some framework needs to be developed. In terms of the very definition of openness, some degree of standardization is already crucial.

“It is quite important that the government is clear what it means by open, and because the terminology is emerging we need to be consistent. For example, does open mean ‘free’? Which bodies should be within our scope, and to what extent should the private sector fall within our grasp? Are we going through an open-society model where not only public-sector organisations should be transparent but any kind of organisations?”¹⁴⁷

“In practice we definitely need a framework, because bureaucrats worry intensely about their responsibilities, the damages they might cause and possibly facing the court. If we give them a framework for that, and that is controlled by central government, then we can possibly move to next stage. If you don't have that framework, it won't happen. Ideally, we wouldn't need that framework and everybody would just do it [release the data] and nobody would give a damn, but that just won't happen.”¹⁴⁸

After the Chancellor's autumn statement, it seemed that the plans for the PDC had disappeared, but two new initiatives were introduced: the establishment of the Public Data Group and the Data Strategy Board.¹⁴⁹ The announcement raised immediately fears that due to a bad response the name of the PDC had been changed but the substance remained similar.

Debate and the development of the PDC are important and interesting also in terms of the definition of openness, and it may possibly be reflected in the wider international discussion on open data. The UK is often looked up to as a model of an open-data –oriented country, and therefore any potential setbacks for open-data

¹⁴⁶ Interview, an open data advocate

¹⁴⁷ Interview, a central government civil servant

¹⁴⁸ Interview, a local government civil servant

¹⁴⁹ Chancellor's Autumn Statement 2011, http://cdn.hm-treasury.gov.uk/autumn_statement.pdf

development are far from desirable. In fact, the very definition of openness and open data proved to be one of the outstanding issues that emerged in the open-data consultation.¹⁵⁰

Despite the high level of political support, not all public authorities have greeted open-data policies with joy. Interviews suggest that in the initial phase of implementing open-data policies the crucial thing has been the enthusiasm of a few individual civil servants, who have been vocal supporters of the agenda. This has effectively led to joint action between grass-root activists and public authorities.

“Persuading officials is a good idea. That kind of change is hard to achieve by legislation.”¹⁵¹

The open-data movement is no exception to the rule of institutional support. However vibrant and active the grass-root movement is, open-data policies will not become an integral part of the political mainstream without a certain institutional push.

It seems that despite the general perception of the UK as a forward-looking open-data country, there are certain obstacles in the way of open-data revolution.

The support for open data is not as extensive as it could be, and at the organisational level it still is up to individual open-data minded officials to push the agenda forward. However, in terms of increasing the institutionalisation, it should be noted that creating public bodies just for the sake of it rarely leads to the results desired. The open-data community has observed the development of the Public Data Corporation with a critical eye, and the reason for its establishment has not been clear.

Moreover, there are lots of vested interests, and arguably the big corporations are those who could benefit most from the possible fee regime. It is highly questionable whether the establishment of the PDC would result in the creative destruction of existing policy frameworks and interest-group structures that is essential in

¹⁵⁰ Interview, a central government civil servant

¹⁵¹ Interview, a central government civil servant

Patashnik's model of sustainable public-policy reform. In the case of the PDC, it is argued that governance would merely consist of the existing interest groups with various vested interests at stake. In other words, another layer of bureaucracy based on existing power structures.

The institutional development seems to have headed towards enhancing the capabilities of established companies to exploit public data. The main driving force has evidently been the economic benefits of data release, and this has at times led to a situation where definitions of open data and big data have become mixed up. It is argued that the government has tried to sell something as open data when it clearly is not open, and any possibilities of this kind of development in other countries should be observed with care.

4.3 Towards Open Licences

The conceptual definition of 'openness' is an intrinsic aspect of open data, as is already argued in the preceding chapters.

Data can thus be in linked and machine-readable format, yet not truly open. In order to be open, data needs to be licensed in a way that recognises the user's right to access and use data freely. In the USA, public-sector information is not copyrighted at all and is thus in the public domain, but in the European context there are relatively complex licensing and copyright legislations and regulations in place.

There are arguments in favour of discarding licences altogether, but in the current European information framework that approach does not sound sustainable. Without a proper licence, data users may get caught in a legal battle over the rights concerning data, and on a wider scale, a lack of licensing makes the emergence of an open-data ecosystem practically impossible, since there would be no clear rules on who can use data and how.

“If you, in fact, have no rights – it is not just a question of whether copyright is bad or copyright is good – when you start talking about data and start releasing stuff, and find yourself in a situation where you have no rights for

your data, no rights whatsoever, someone else can put their name on your work, or change your stuff and put your name on it.”¹⁵²

The question of licensing is no exception in the institution-led trend of open-data development in the UK. In this case, the principal player has been the National Archives, more precisely the UK Government Licensing Framework (UKGLF).

The initial development phase of the Open Government License (OGL) happened about the same time as the release of the influential Power of Information Report¹⁵³, which ensured that public officials were relatively interested. The report had already made certain recommendations for more transparent and user-friendly systems, and the Government decided to do something based on these recommendations.

“The Open Government Licence process started in 2008-2009 when we started to think about online licences where people didn’t have to register and think what they can use and what not and whether they have to pay a fee. Instead, can’t we just be more enabling, that would cost government less and be easier for developers. This came around the same time as the Power of Information report, which was important.”¹⁵⁴

In 2010, after a series of consultations, the National Archives established the Open Government Licence, which officially encourages the user to:

- Copy, publish, distribute and transmit the information;
- Adapt the information;
- Exploit the information commercially, for example, by combining it with other information, or by including it in your own product or application.¹⁵⁵

¹⁵² Interview, an academic/researcher

¹⁵³ Mayo, Ed & Steinberg, Tom: *The Power of Information*, http://webarchive.nationalarchives.gov.uk/20100413152047/http://www.cabinetoffice.gov.uk/media/cabinetoffice/strategy/assets/power_information.pdf

¹⁵⁴ Interview, a central government civil servant

¹⁵⁵ The National Archives: *Open Government Licence for public sector information*, <http://www.nationalarchives.gov.uk/doc/open-government-licence/>,

In terms of public data use, the OGL imposes certain demands on the user, namely to:

- Acknowledge the source of the Information by including any attribution statement specified by the Information Provider(s) and, where possible, provide a link to this licence;
- Ensure that you do not use the Information in a way that suggests any official status or that the Information Provider endorses you or your use of the Information;
- Ensure that you do not mislead others or misrepresent the Information or its source;
- Ensure that your use of the Information does not breach the Data Protection Act 1998 or the Privacy and Electronic Communications (EC Directive) Regulations 2003.¹⁵⁶

Exemptions to the licence include:

- Personal data in the Information;
- Information that has neither been published nor disclosed under information-access legislation (including the Freedom of Information Acts for the UK and Scotland) by or with the consent of the Information Provider;
- Departmental or public-sector organisation logos, crests and the Royal Arms, except where they form an integral part of a document or dataset;
- Military insignia;
- Third-party rights the Information Provider is not authorised to license;
- Information subject to other intellectual property rights, including patents, trademarks and design rights; and
- Identity documents such as the British Passport¹⁵⁷

Moreover, according to Francis Maude, certain data where “charges are required to ensure the quality and sustainability of data” would be exempt from the OGL.¹⁵⁸

¹⁵⁶ *Ibid.*

¹⁵⁷ *Ibid.*

¹⁵⁸ UK Parliament Commons, Written Answers, HC Deb, 4 May 2011, c791W

In technical and legal terms, the OGL is relatively similar to the widely used Creative Commons CC-BY licence, and questions have arisen why there was a need for a completely new licence in the first place. One of the main reasons for a separate licence was “the need to account for database rights, which are particular to the EU”.¹⁵⁹ Database rights were therefore apparently the principal reason why the Creative Commons licence was not adopted by the UK government.

“Initially we thought of adopting Creative Commons for government information. We were aware that New Zealand had done that and also Australia. I think there were a couple of issues why we couldn’t: the database-rights issue and the fact that there wasn’t a CC licence to cover the whole of the UK. There was a question of consistency and a question of implementation of this policy across government. We wanted something simpler and shorter than the legal deeds that were behind CC.”¹⁶⁰

In terms of the principles of licensing, UKGLF had highlighted a few outstanding issues, namely:

- Simplicity of expression - the terms should be expressed in such a way that everyone can understand them easily;
- Non-exclusivity - so that access can be provided to a range of users on fair and equal terms;
- Fairness of terms;
- Non-discrimination – terms are extended fairly to all for similar uses;
- The need for acknowledgment and attribution;
- The need for transparency by publishing standard licence terms¹⁶¹

The development of the Open Government Licence dates back to the inaugural EC Directive on Public Sector Information Re-use 2003, which establishes “a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public-sector bodies of the Member States”.¹⁶²

¹⁵⁹ Open Knowledge Foundation mailing list discussion, <http://lists.okfn.org/pipermail/open-government/2010-October/000264.html>

¹⁶⁰ Interview, a central government civil servant

¹⁶¹ UK Government [Licence Framework](http://www.nationalarchives.gov.uk/information-management/government-licensing/the-framework.htm), <http://www.nationalarchives.gov.uk/information-management/government-licensing/the-framework.htm>

¹⁶² European Directive 2003/98/EC of 17 November 2003 on the re-use of public sector information. Official Journal of the European Union 345/90,

The Advisory Panel on Public Sector Information¹⁶³, however, had identified some potential obstacles in the way of optimal re-use policies.

Firstly, according to the Directive, the UK public-sector bodies were not obliged to make all potentially beneficial PSI available. According to the APPSI, this could potentially have led to inconvenient requests for PSI to be “simply parked”.¹⁶⁴

Secondly, the APPSI questioned the vague definition of the Public Task. Under the then existing regulations, information created by any organisation under its Public Task should have been subject to PSI re-use policies.¹⁶⁵ According to the APPSI, however, publication could have potentially been avoided due to unclear definition: there simply was no clarity on what constituted the publication of information collected by the government.¹⁶⁶

Thirdly, exclusive intellectual property rights licensing arrangements between public-sector bodies and third parties that took place were deemed unwise.¹⁶⁷ According to the UKGLF, licensors are, in fact, obliged to identify any parts of the information for which rights are owned by a third party.¹⁶⁸

Finally, the APPSI had concerns about the generally low awareness of the benefits of PSI re-use and therefore – coincidentally in accordance with the Local Public Data Panel – emphasised the need for raising awareness.¹⁶⁹

With all these issues in mind, the APPSI argued in 2009 that the main obstacles were no longer technological but rather cultural, institutional, financial and policy-based ones.¹⁷⁰ In order to overcome these obstacles, three complementary elements were considered essential:

http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive/psi_directive_en.pdf

¹⁶³ Board of experts, which was set up under the National Archives.

¹⁶⁴ Letter from the APPSI to Minister Michael Wills MP, 21 July 2008,

<http://www.nationalarchives.gov.uk/documents/appsi-minister-letter-ec-submission-07-2008.pdf>

¹⁶⁵ Ibid.

¹⁶⁶ Letter from the APPSI to Minister Lord McNally, 6 July 2010,

<http://www.nationalarchives.gov.uk/documents/APPSI-letter-to-Lord-McNally.pdf>

¹⁶⁷ Letter from the APPSI to Minister Michael Wills MP, 21 July 2008,

<http://www.nationalarchives.gov.uk/documents/appsi-minister-letter-ec-submission-07-2008.pdf>

¹⁶⁸ UK Government Licensing Framework, Licensing principles and practices,

<http://www.nationalarchives.gov.uk/information-management/government-licensing/licensing-principles.htm>

¹⁶⁹ Letter from APPSI to Minister Michael Wills MP, 21 July 2008,

<http://www.nationalarchives.gov.uk/documents/appsi-minister-letter-ec-submission-07-2008.pdf>

¹⁷⁰ APPSI: *Exploiting governments information assets for public good*,

-
- Celebrate and extend existing successes
 - Gather additional PSI and make it immediately available
 - Follow a longer term, prioritised information-harvesting strategy¹⁷¹

In addition to these three elements, the APPSI explicitly urged the government to take other actions that could enhance the re-use of PSI, simplifying the licensing being arguably one of the most important.¹⁷²

It should be noted that, apparently, the approach the APPSI adopted was not aimed at maximising the amount of information available, but at maximising utility and minimising ad-hoc procedures and “unnecessary effort”. In fact, according to the APPSI, this approach could very well have resulted in less information being collected.¹⁷³

To conclude on the development of PSI licensing in the UK, a strong institutional basis and a well-formulated long-term information strategy are considered to be essential. There has been some overlap in terms of the responsibilities of the different public authorities, and partly as a result of this further simplification of duties has been suggested.

4.4. Do Licences Work?

The Open Government Licence has received a relatively good response from the open-data advocates so far. It is considered simple enough and liberal in its re-use clauses. Some questions remain unanswered, however, and should be taken into consideration when facilitating the use of open data.

<http://www.nationalarchives.gov.uk/documents/Exploiting-government-information-assets-for-public-good.pdf>

¹⁷¹ APPSI: *Exploiting governments information assets for public good*,

<http://www.nationalarchives.gov.uk/documents/Exploiting-government-information-assets-for-public-good.pdf>

¹⁷² *ibid.*

¹⁷³ Letter from APPSI to Minister Lord McNally, 6 July 2010,

<http://www.nationalarchives.gov.uk/documents/APPSI-letter-to-Lord-McNally.pdf>

One outstanding indicator of whether a licence is a success or not is the sheer number of public-sector institutions that have used the licence. By January 2011, over sixty local authorities had adopted the licence, and the number has kept rising since.¹⁷⁴ Moreover, according to officials from the National Archives, there has been constantly growing pressure among public-sector data producers to adopt the OGL. Openness is already seen to be the acknowledged standard.¹⁷⁵

“Licensing seems to be very successful, especially among local government. It is not mandatory but recommended and encouraged. There is lots of pressure to use it – websites and campaigns like OpenlyLocal have pressurised authorities to use it. Openness is already more or less the acknowledged standard among the government. We were concerned whether the licence would appeal to organisations outside central government, so we wanted it to be something that all public-sector agencies would like to use. It seems to have worked well.”¹⁷⁶

Despite the mostly positive response for the licence, the issue of machine-readability is a concern that should not be underestimated. Although OGL is marked up using the Creative Commons vocabulary, there are arguments that Creative Commons itself has taken the demands of machine-readability into consideration better than the OGL has.¹⁷⁷

Without a sufficient level of machine-readability in the licence, the data provider faces the threat of excluding itself from the wider Internet community. Interviews suggest that one of the principal reasons for the success of Creative Commons has been – in addition to its high number of top-level IPR lawyers and their work – its emphasis from the beginning on machine-readability.¹⁷⁸

“One of the great advantages of using a licence like Creative Commons’ is that they are tried and tested.”¹⁷⁹

¹⁷⁴ European Public Sector Information Platform, Topic Report 22: *Simplifying PSI re-use in the United Kingdom: the UK government Licensing Framework and the Open Government Licence*.

¹⁷⁵ Interview, a central government civil servant

¹⁷⁶ Interview, a central government civil servant

¹⁷⁷ Interview, an academic/researcher

¹⁷⁸ Interview, an academic/researcher

¹⁷⁹ Interview, an academic/researcher

The question that many governments are currently facing is whether or not to develop and adopt a licence of their own. In addition to the UK, there are a couple of countries with their own licences and some – mostly outside Europe – have adopted CC licences for open data.

“We drew a pretty tough line that unless someone gave a really good reason to develop a new licence, we would stick with the CC licence.”¹⁸⁰

When deciding on the licensing framework, it is highly recommended that governments firstly identify their principal clients: who is it hoped will use and re-use public information. After this identification process, governments can act accordingly. In terms of applicability of data, it would be advisable to copyright information under an open licence that minimises the risk of attribute stacking and thus lowers the threshold of re-using data freely. With open data, there is arguably the danger of a situation where huge numbers of datasets with different licensing models are gathered and analysed, thus creating a whole new piece of original and copyrightable work. Licensing that work is effectively almost impossible given the jungle of different licences according to which the original datasets were copyrighted.

“The issue with attribution is that it would be likely to be impossible to attribute the whole myriad of contributors to datasets.”¹⁸¹

A comprehensive licence for open-data purposes should thus satisfy three key demands:

- Machine-readability
- Human-readability
- Lawyer-readability

Naomi Korn and Charles Oppenheim also claim that the ideal licence types for open data are those with minimal restrictions, not even an attribute clause.¹⁸² Their review of different models of licensing open data is presented in Table 1.¹⁸³

¹⁸⁰ Interview, an academic/researcher

¹⁸¹ Interview, an academic/researcher

¹⁸² Korn, Naomi & Oppenheim, Charles: Licensing Open Data: A Practical Guide,

Table 1: Different types of open-data licences

Licence type	Who can use the resource and under what terms?	Can the licensed data be modified?	Suitability for data, datasets and databases
CC Attribution (CC-BY)	Anyone	YES, but you must attribute	Not specifically geared towards data, datasets and databases, but can be used with minimal amounts of data (to avoid attribution stacking) and as long as only an “insubstantial” amount of any databases or datasets are re-used
CC Attribution Share Alike (BY-SA)	Anyone	YES, but you must attribute, and if you use or re-use the data, etc., you must use the CC BY-SA end-user licence for onward licensing.	As above. The Share Alike requirement can impact negatively on interoperability of data and prevent linked open data.
CC Attribution Non-Commercial (BY-NC)	Anyone – for non-commercial purposes only	YES, but you must attribute.	As above. Although the NC restriction does not pose immediate problems, the ambiguity surrounding what constitutes non-commercial may be problematic. There may also be interoperability problems with linking to data licensed under more permissive terms.
CC Attribution No.Derivatives (BY-ND)	Anyone	NO, and you must attribute.	As above. Re-use and re-purposing of data, datasets and databases are not permitted.
CC Attribution Non-Commercial Share Alike (BY-NC-SA)	Anyone – for non-commercial purposes only	YES, but you must attribute and if you use or re-use the data etc, you must use the CC BY-SA end-user licence for onward licensing.	As above. The Share Alike requirement can impact negatively on interoperability of data and prevent linked open data. Although the NC restriction does not pose immediate problems, the ambiguity surrounding

http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf

¹⁸³ Ibid.

			what constitutes non-commercial may be problematic. There may also be interoperability problems with linking to data licensed under more permissive terms.
CC Attribution Non-Commercial No Derivatives (BY-NC-ND)	Anyone – for non-commercial purposes only	NO, and you must attribute.	As above. Re-use and re-purposing of data, datasets and databases are not permitted. Although the NC restriction does not pose immediate problems, the ambiguity surrounding what constitutes non-commercial may be problematic. There may also be interoperability problems with linking to data licensed under more permissive terms.
Creative Commons Zero	Anyone	YES, with no restrictions whatsoever.	Ideal.
Open Data Commons Open Database Licence	Anyone	YES, but you must attribute any public use of the database, or works produced from the database, in the manner specified in the ODbL. For any use or redistribution of the database, or works produced from it, you must make clear to others the licensing situation for the database and keep intact any notices on the original database. Share-Alike: If you publicly use any adapted version of this database, or works produced from an adapted database, you must also offer that adapted database under the ODbL.	Ideal – although there may be some attribution requirements, leading to possible attribution stacking and also interoperability issues associated with the Share Alike requirement.
Open Data Commons Attribution Licence	Anyone (applies to data, datasets and	Yes – but you must attribute any public use	Ideal – although there may be some

	databases)	of the database, or works produced from the database, in the manner specified in the ODbL. For any use or redistribution of the database, or works produced from it, you must draw third parties' attention to the original licence for the database and keep intact any notices on the original database.	attribution requirements, leading to possible attribution stacking.
Public Domain and Dedication Licence	Anyone (applies to databases)	YES, with no restrictions whatsoever.	Ideal.
Open Government Licence	Anyone (applies to content, data, databases and source code)	YES, but you must attribute	Can be used with minimal amounts of data (to avoid attribution stacking).

5. Benefits and how to measure them

“Generally speaking the benefits are still waiting to be realised.”¹⁸⁴

Throughout this report, I’ve argued in favour of primarily identifying non-economic benefits for open data. Arguably, the economic benefits are relatively easy to find evidence for, and as important as they are, in order to generate wider value from open data, more evidence on the social and political implications needs to be found. This is hard to achieve without bold ways of analysing the impact using non-economic or non-quantitative means.

However, there are varying views as to whether even the economic impacts have been that clear yet. Huijboom and Van den Broek argue that the precise economic impact of open-data policies remains largely unclear and calculations differ substantially.¹⁸⁵

Moreover, there is currently even less evidence on the social and democratic effects of open data. It seems that the causal relationship between open data and democratic participation is far from clear.¹⁸⁶ If it turns out that opening up of government datasets has not produced the estimated economic impacts, it falls solely to intrinsic arguments to justify the purpose of the initiative.

“I’ve been saying to people in the government, already since the previous government, that actually the democratic arguments are much more important.”¹⁸⁷

In the first chapter, I reviewed some of the potential areas where open data has been claimed as particularly beneficial. Drawing from the research literature, it was possible to categorise these areas into two primary groups, internal and external benefits. In this chapter, I will analyse both areas and argue whether current UK policies have supported these areas and, moreover, how open data has been applied in terms of achieving these objectives.

¹⁸⁴ Interview, an open data advocate

¹⁸⁵ Huijboom & Van den Broek, 2011.

¹⁸⁶ Huijboom & Van den Broek, 2011; McClean, 2011.

¹⁸⁷ Interview, a local government civil servant

5.1. Internal benefits

5.1.1 Efficiency and new information management

This paper argues that the most important potential internal impact that could be achieved with open data is the one concerning the attitudes of how public-sector information management, and governance at large, should be run. Public-sector organisations are run by public funds, and therefore it is justified for them to be as efficient and well established as possible.

Typically, arguments that are given by politicians – especially by those that are fiscally on the right – concentrate on public spending and eliminating potential waste in it.¹⁸⁸ Already, the establishment of the FOIA was expected to reduce this waste and lead to more efficient decision-making, although later research has suggested that this is not necessarily the case.¹⁸⁹ Open data, on the other hand, is expected to further enhance the development of better decision-making by giving opportunities to an even wider element of the public to scrutinise public spending effectively.¹⁹⁰

Despite the increased amount of data available to citizens, it is questionable whether open data has resulted in massive efficiency savings in the UK so far. Despite the promising initial research results from the USA, it is still a bit too early to draw conclusions on possible efficiency savings in the UK public sector. It may well be that in the long run the release of data can lead to significantly enhanced efficiency, but certain obstacles have seemingly emerged, namely the operability of current IT-systems, the lack of context in data, data literacy and negative attitudes among public officials, who should be the ones implementing the scheme.¹⁹¹ Moreover, the initial research results indicate that there has so far been little sign of the emergence or effectiveness of armchair auditors.¹⁹²

¹⁸⁸ See, for instance: Osborne 2011, Pickles 2010.

¹⁸⁹ Hazell et al., 2010.

¹⁹⁰ See, for instance: <http://openspending.org>

¹⁹¹ Halonen, Antti: *Data without context? A study on the UK local government transparency agenda and its impacts on information management*, Unpublished MRes dissertation, University College London, 2011.

¹⁹² Worthy, Ben et al. *Town Hall Transparency? The Impact of Freedom of Information on Local Government in England*, The Constitution Unit, December 2011. <http://www.ucl.ac.uk/constitution-unit/research/foi/foi-and-local-government/town-hall-transparency.pdf>

The release of expenditure data in particular is hoped to increase public-sector efficiency. Some of these impacts are assessed in a case study on the English local-government transparency agenda.

5.1.2 Case: UK local-government transparency agenda and data producers' perceptions

The current UK government initially emphasised the release of expenditure data in particular. There are likely to be various reasons for this decision – mostly political – and some tentative conclusions on the initial impacts can already be made. The case study is based on separate research conducted by the author of this study.

As mentioned earlier in this study, in the summer 2010 the UK government released a statement, which obliged all local councils to publish their expenditure data of public spending over £500. During the year 2011, all English local councils were surveyed, and their perceptions on the publishing of open data were analysed with some interesting results. By the time of doing the survey, all but one council had obeyed to the regulation, thus giving a good starting point for the analysis.¹⁹³

The survey was initially sent to councils' freedom-of-information departments, and where the FOI team was not responsible for publishing expenditure data, they were asked to forward it to the appropriate sector. Of 356 councils, 112 responded (n=112) within the time limit given, thus making the invitee response rate 31%. Due to the relatively low sample, the statistical results should be taken with a measure of caution, but nevertheless there are some initial conclusions to be drawn. In-depth interviews and open-ended survey responses provided more quality for analysis.

The survey reveals that there is no clear countrywide pattern in place in terms of implementing the scheme.¹⁹⁴ Job descriptions and responsibilities vary from council to council, and this partly illustrates the need for clearer guidance on implementation.

When respondents were asked whether they think that the transparency initiative has been successful so far, the most popular answer was “in theory yes, but in practice no”. In total, roughly 38% of the respondents perceived the initiative as unsuccessful, and 36% either very successful or somewhat successful. 23.6% answered “don't know”, which was

¹⁹³ Halonen, 2011

¹⁹⁴ Halonen, 2011

due to the respondents' failure to understand the question wording¹⁹⁵ and the perception that it is still quite early to analyse the overall success of the initiative.

“It depends on how you define successful. While we have produced the data, there appears to be minimal interest in it.”

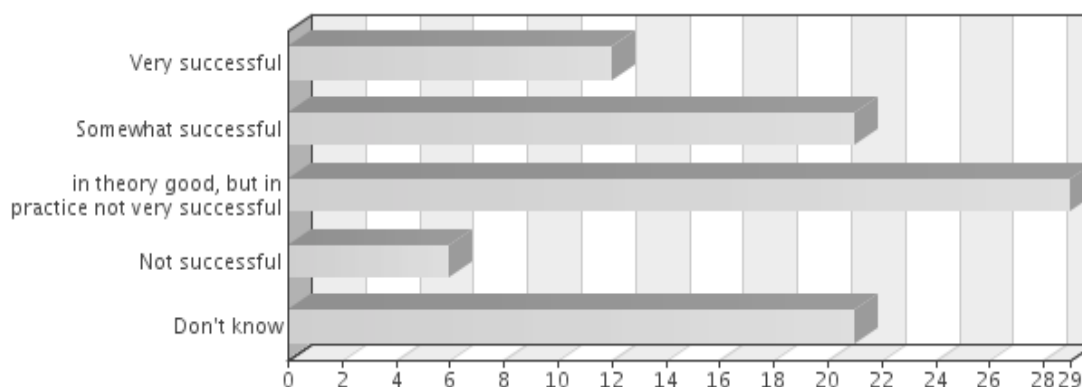
“Difficult to know, the government’s assertion is that publishing this data would reduce FOI requests. There is certainly no evidence that this is the case.”

“It is too early to tell yet. To date, only three months’ data have been published, with very few requests for information.”

“The initiative has been a success in that all but one council has implemented it. But there is little benefit arising from the initiative – it’s just window-dressing.”

“The predicted surge in armchair auditors has not happened (thank goodness)”

In your council, do you think that the recent government initiative to publish online local expenditure data for spending over £500 has in general been:



As noted already on several occasions, some of the main arguments for proactive publishing and open data are often related to efficiency and accountability. Out of the survey respondents, however, only one explicitly mentioned efficiency (“reducing unnecessary spend”) as one of the most important positive impacts of the transparency initiative. Transparency *per se* was seen as the most important objective, and based on

¹⁹⁵ The word “successful” was criticised as it left too much scope for the respondents' own interpretation of the question's purpose. Some wondered for whom the initiative is “successful”: to council workers, citizens, politics or the central government. Ibid.

this survey it is not possible to draw conclusions about whether respondents class efficiency as an intrinsic part of transparency as a whole. One answer option, however, was intrinsically related to efficiency, namely “better decision-making”. It is interesting that this particular option was the least popular, thus perhaps undermining the argument for publishing expenditure data to increase government efficiency. Based on qualitative interviews, the main reason for the lack of trust on the efficiency-improving impact of the agenda is its early stage: scepticism is seen as typical considering where the councils are now in implementing the transparency agenda, and the situation may change if the government further increases proactive publishing.¹⁹⁶

“While I support initiatives to be more transparent, I don’t think that publishing spending data we’ve been asked to produce does any of these. It’s not comparable with other authorities, no-one knows what it’s spent on, so it has no qualitative value.”

“To the best of my knowledge, very few people are even interested in that level of detail, so I feel it has been a large waste of time and money.”

“Although it could help to reduce wasteful spend as people will not want this to be public, it does seem that people focus on the pennies rather than the pounds. Also, the data is only useful in context – i.e. why the money was spent. So people try to put their own interpretation on this, which is often incorrect, but gets into the press and gives a completely false impression.”

“The open-data movement should get public bodies to improve the thinking behind what data is collected and how it can be presented – they should think of data consumption in the first place.”¹⁹⁷

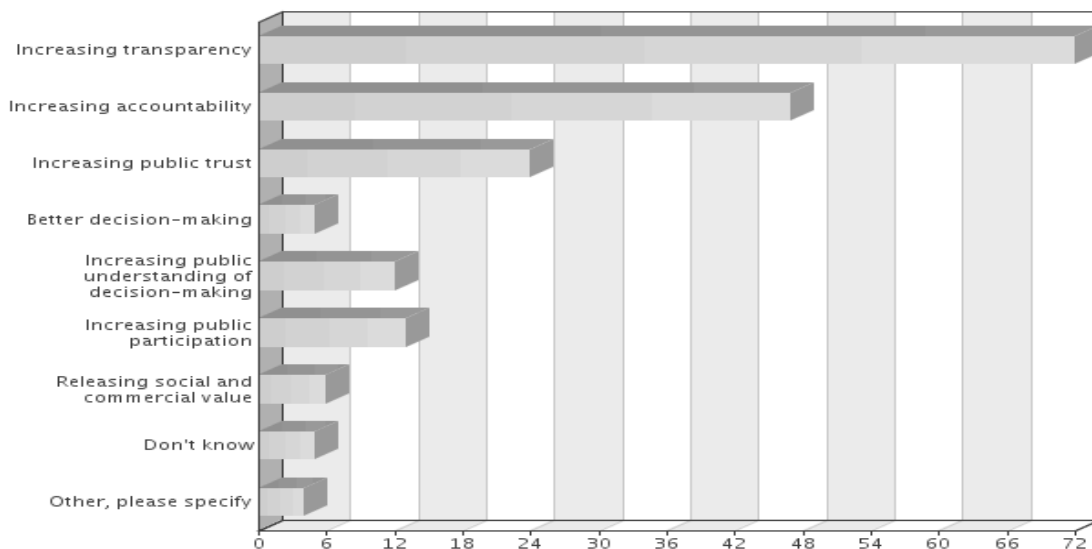
However, in order to present strong empirical claims on whether expenditure data has, in fact, increased efficiency, much more sophisticated empirical methods of measuring organisational efficiency and return of investment would be needed, and that would require research of its own, preferably in the field of administrative science or organisational economics.

¹⁹⁶ Halonen_2011

¹⁹⁷ Halonen_2011

In your opinion, which of the following objectives do you think would be most likely to be accomplished by publishing expenditure data?

Please pick a maximum of three choices.



Furthermore, despite the enhancing impact on transparency, respondents were in general quite sceptical towards other positive impacts of the initiative. Given that empowerment of citizens and enhancing public participation has been one of the primary objectives the UK government has set for the transparency agenda, the level of respondents' confidence in its accomplishment is relatively low, only ~15%. Also, council officers don't see much commercial value in publishing expenditure datasets (~7%).

In general, the findings on respondents' views on objectives seem to indicate a certain level of discrepancy between the goals of the agenda set by the government and how officials perceive the real outcome of the agenda. Only 6% of the respondents felt that increasing the quality of decision-making would be one of the three most important objectives of the scheme, while enhancing efficiency clearly has been one of the major arguments for the whole transparency agenda. One possible reason for this is the highly politicised setup of the agenda. The debate has been relatively political, and this could be reflected in the views of some councils.

When asked about the possible ill-effects, quite a few respondents questioned the value of the whole transparency agenda. Interestingly, the majority of respondents felt it would primarily result in an unnecessary workload (~56%), and many wondered whether the benefits would ever exceed the costs. In addition to fears of misinterpretation, respondents expressed their concerns about the usability of data and the confusion it may cause among the public. As one officer put it, “data is meaningless beyond providing just a figure against a company name”, and two others commented, “I believe the information is generally confusing to the average member of the public” and “[causes] confusion and raises questions due to a lack of detail”.

Moreover, many respondents shared a concern on how the data is used. There are worries that data is mainly used for the purposes of “lazy journalism” or drawing false conclusions on council spending, which effectively decreases public trust in local government. As one respondent put it: “The biggest risk is misinterpretation, leading to unwarranted loss of public trust”. Another one raised a seemingly similar concern: “Third-party users have the opportunity to challenge, in a destructive manner, the decision-making of local authorities, i.e. the media sensationalises aspects of information obtained”. Negative press coverage has, however, not significantly decreased the public trust in local government. According to a survey conducted by the Local Government Association and LGinsight, the level of trust has remained relatively high, with 67% of respondents being satisfied with the job their local council is doing.¹⁹⁸

Despite the concerns regarding data use and misinformation, overall the biggest concern by far was officials’ perception of the impact on personal workload. More than half of the respondents chose “increase of workload” as one of the most significant ill-effects of the scheme.

Qualitative interviews shed some light on the reasons behind respondents’ tendency to be relatively cautious. The transparency agenda is still at an early stage, and thus it is not an easy task to analyse its impacts so far. “It’s very early days, you’d think that you have more important things to do, but it may take many years to do it [to implement the agenda properly], and then we’ll have to tip our hats to people who had the foresight and government who had the courage”.¹⁹⁹

¹⁹⁸ LGA & Lginsight: *What people really think about their council and what should you do about it*, <http://www.lgcomms.org.uk/asset/513/Poll%20Headlines.pdf>,

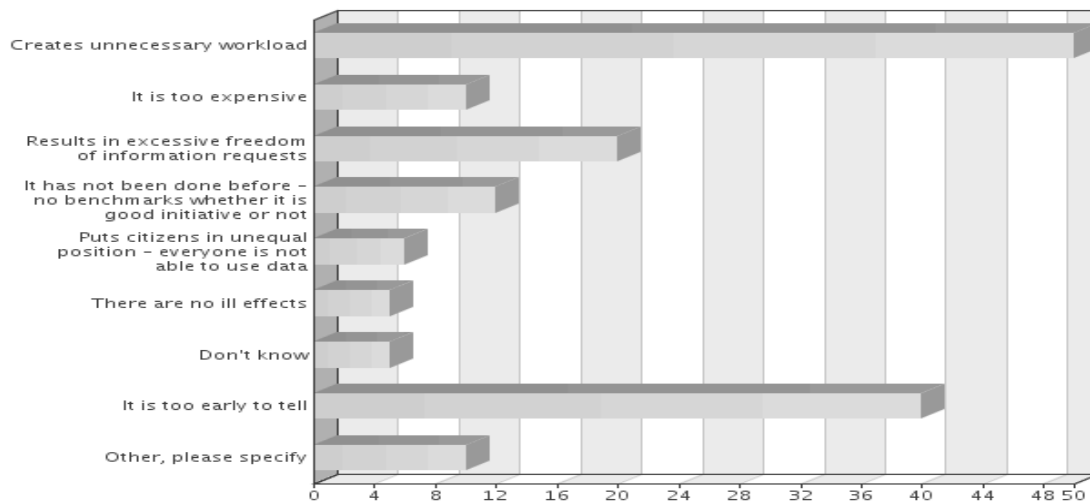
¹⁹⁹ Halonen, 2011

“There’s a serious workload for the outcomes. And it doesn’t create more FOI requests, because no-one is interested in data.”

“There is an additional cost attached to this initiative, and have we asked the public whether this money is well spent and whether they want this information?”

In your opinion, which ill-effects would you think publishing of expenditure data would cause in your council?

Please pick a maximum of three choices.



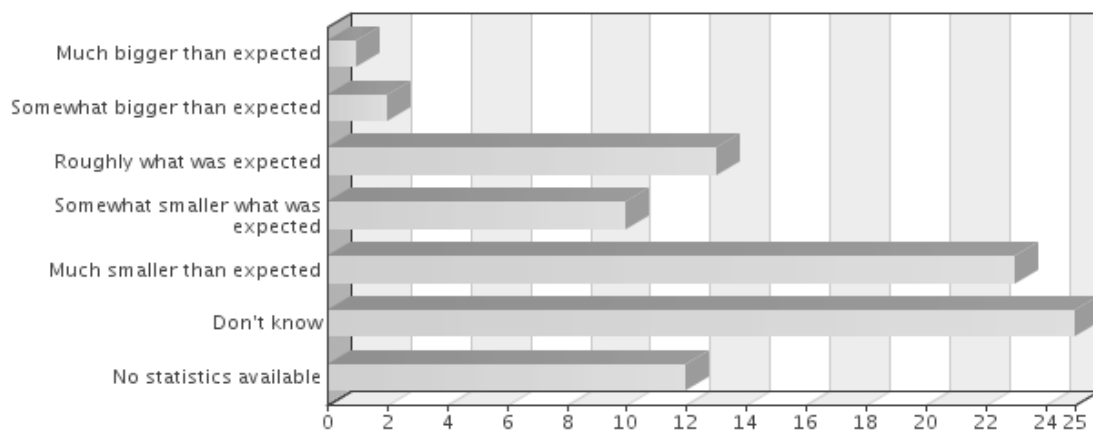
In terms of the wider implications of the agenda, perhaps the most striking survey result was the one regarding perception of public interest towards data. Only three (3) individual respondents held the view that the general interest towards data has been greater than expected, while of those who gave any opinion the vast majority thought that interest has been either somewhat smaller than expected or even much smaller than expected. Quite a few didn't have sufficient statistics available, and a large number of respondents chose the answer “don't know”.

The importance of public interest in data is especially high if the value of transparency is measured only in instrumental terms. Evidently, there is a need for more investment in putting the right data online, but it seems that the public interest for that is not there at the moment. One suggested that an answer to the lack of interest was to provide more useful data: “things about services, anti-social behaviour, engaging things people care about, not just financial stuff”.

“The same numbers appear to have been looking at the pages after Jan 2011 as before.”

“Hard to say as no previous experience, and with low publicity until Jan 2011 it is hard to gauge. Ask again in 6 months, as the public get used to it and there are more opportunities to explore the data and compare it through online applications.”

In your opinion, how would you describe the number of views and downloads of expenditure datasets in your council?



Although most councils collect statistics on data downloads and page hits, it seems to be rare to engage in dialogue with data users. Councils tend to focus on putting data online and complying with regulations. Seemingly, councils are not necessarily in a position to further engage with developers and other data users. None of the councils examined through interviews knew who, in fact, uses their data, which partly undermines the value of the scheme for public participation.²⁰⁰

Similar concerns were expressed in other questions, as well, for example when asked about objectives or ill-effects. Some respondents questioned the whole value of the agenda, based on the lack of interest. In the words of one individual respondent: “to the best of my knowledge, very few people are even interested in that level of detail, so I feel it has been largely a waste of time and money”. Interviews strengthened this perception further. Interviewees, regardless of their overall attitude towards the scheme, pointed out

²⁰⁰ Halonen, 2011

the general lack of interest and raised questions about whether this undermines the overall value of the whole agenda. A senior finance officer, who was generally supportive towards the agenda, summarised, “if people are not interested, let’s not waste our time and do something more useful”.²⁰¹

When asked about obstacles during the implementation process, opinions were more split. A few outstanding concerns emerged, however: inadequate advice from the central government (28%), lack of staff (26%), IT resources/technical issues (24%) and financial resources (21%). Furthermore, 26% did not observe any notable obstacles.

“Like most government initiatives, it was made up on the hoof. Initially, there was no clear guidance. When guidance did emerge, it was obviously flawed and has changed several times since. This meant implementation cost more than necessary.”

“The IT-technical issues are around how to extract personal data quickly and how to distinguish sole traders from non-commercial individuals. At the moment, we’re manually trawling through thousands of lines, which is very time-consuming.”

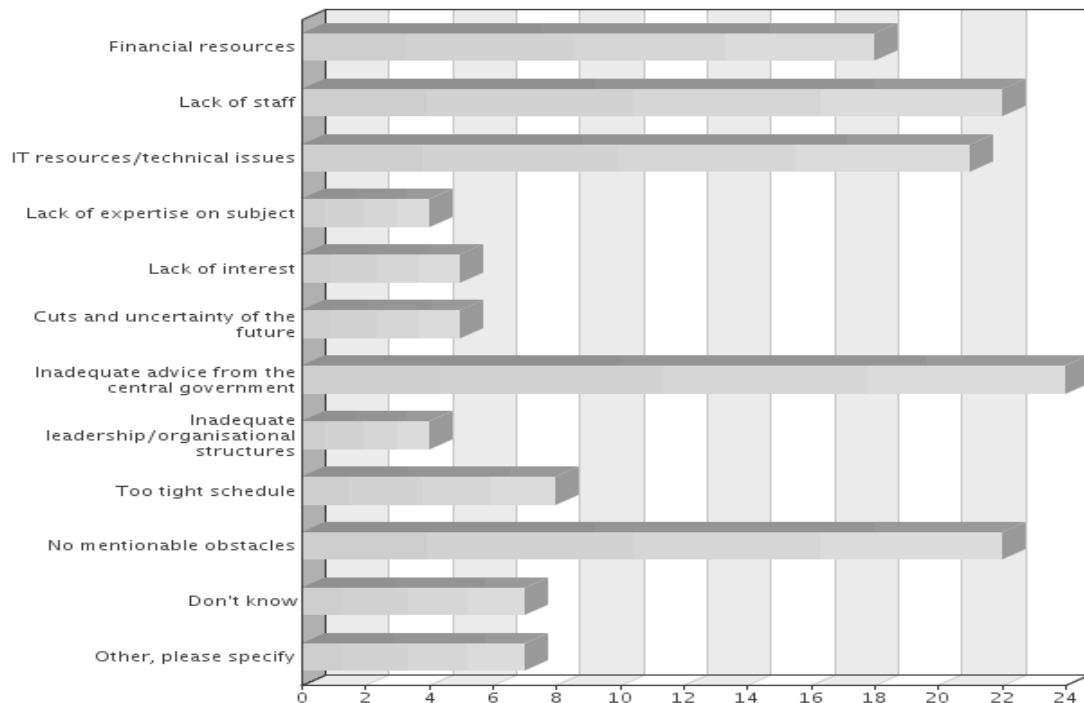
“Guidance wasn’t clear in all aspects – the ICO still haven’t advised if it is lawful to disclose payments to sole traders – so there are aspects where the government guidance is incomplete.”

“The council already published various pieces of information within its own transparency agenda and is keen to be open in its dealings. There have been inadvertent issues over confidentiality and data protection.”

²⁰¹ Ibid.

In your council, were there any obstacles encountered during the process of publishing expenditure data?

Please choose all applicable answers.



The question on whether councils have taken certain actions after they implemented transparency-agenda requirements was a little flawed, since it turned out that the answer options and the publishing data are not necessarily related in any way. Based on respondents' views, the expenditure-data scheme has not had any significant impact on wider organisational development so far, with only 10% saying that, in their organisation, record-management officers or FOI officers have been retrained, since the council started to publish data. Open-ended answers further clarify the lack of causality between transparency agenda and organisational changes.

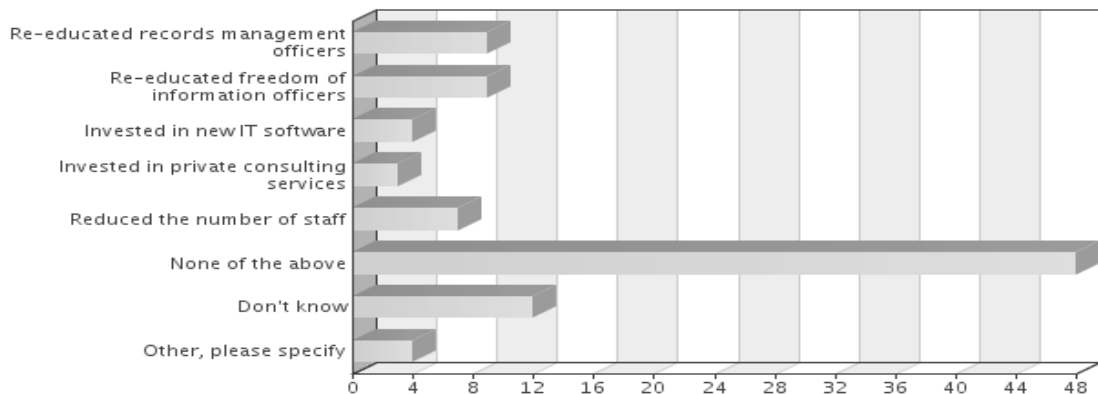
“I’m not aware of any changes – it’s just another task that has to be undertaken by already pressured staff.”

“Not related to publication of spending data though.”

“Additional resources in staff time & IT, but not an actual increase in staff numbers.”

Since your council started to publish expenditure data online, which of the following actions has your council taken?

Please choose all applicable answers.



The fears of data producers partly confirm the concerns about a certain gap in people's capabilities of using data. Although the expenditure-data initiative arguably increases transparency per se, many of its presumed positive impacts seem to be left unaccomplished so far.

A confident conclusion, based on survey results, is that overall respondents are relatively sceptical towards the benefits of the expenditure-data agenda itself, but supportive towards the idea of open data. This seems to be due to respondents' perception that while expenditure data without proper context might not be very useful to citizens, usability of data can be increased significantly if data are in the form of open linked data and thus significantly more re-usable. One particular answer roughly summarised the overall view of respondents: "the agenda on this issue is still somewhat confused, in that the stated objective is transparency whereas the wider benefit will be derived from open and linked data."

"The challenge is that all councils have pursued this in different ways. In a sense, a council can comply with the letter of the programme but not its spirit. Local government, by its nature, is inherently opaque. Therefore, the attempt to make it

transparent will be met with strategies and tropes that may render that transparency opaque. For example, search some of the various sites for sensitive words that would indicate spending on areas where the public, and particularly, the press would be interested and you are likely to find them covered by euphemisms or simply captured in a higher level, more abstract, category. The dialectic of transparency will continue.”

“Given that local government has traditionally been far more open than central government, it would be nice to see them open up a little bit more instead of deflecting it all our way.”

“Whilst we as a council agree that it is very important to do all we can to increase transparency and accountability – it costs money to do this and at a time when cuts to services are being made I am not sure it is seen as a high priority by our residents. So whilst we support this in principle – and are doing all we can to comply with the government's open-data movement – there is a limit to the amount of resources, which can be devoted to this.”

“If this level of openness does have a genuine benefit, perhaps it should be extended to contractors and charities that provide public services and services that were previously in public ownership such as energy.”²⁰²

Although the overall level perception of open data was positive, certain qualitative criticism emerged. Arguments here were seemingly similar to general arguments against expenditure data at large: data, no matter how open, is relatively useless if the public interest towards it is as low as it is. According to critics, there is no proper return of investment in terms of human and financial resources used to implement open-data strategies. Also, the whole idea of making profit out of publicly financed data raised questions of equality and fairness: is it really the purpose of public-sector organisations to provide free data for commercial exploitation?

It seems that views and expectations of central government and open-data advocates are better matched than those of local government officers. The primary focus of central government policies and statements has been around efficiency, accountability and the commercial value of data, whereas local government officials emphasise usability of and user interest in the data, the costs of the scheme and the workload it causes, thus having a much more practical mindset.

²⁰² Halonen, 2011

Implementation of the transparency agenda has revealed that public authorities are not necessarily ready for the open-data revolution – yet. Officials don't see the potential benefits of open data but are more concerned about the personal workload that opening up of datasets could cause. It has to be noted, though, that the whole agenda is still at very early days, which is arguably one of the main reasons why officials haven't yet seen the benefits.

As became evident in the case study for expenditure data, some of the data producers have initially feared that proactive publishing of open data would radically increase the number of freedom-of-information requests, thus causing a heavier burden for officials. The reasoning behind this argument derives from the perception that published data lacks relevant context and hence lacks information value. This is expected to lead to more freedom-of-information requests, as people want to understand the context behind the data.

Some data producers and analysts have also estimated that the number of requests will decrease, as citizens can find the information they want straight from the website. However, in order to achieve this benefit, the data-release system must be comprehensive and simple enough for citizens to use. Some examples suggest that citizens might get confused by the number of datasets and will not necessarily find the information they would be interested in. One successful option has been to present, by default, only the data that citizens have asked for previously via a FOI. Those who are interested in more data can then find it if they so wish, but the idea is to make the interface as simple as possible and not to put off those who are only looking for the most basic democratic data.²⁰³

“We had to find a way of hiding a lot of data.”²⁰⁴

However, initial findings on the impacts of the publishing of expenditure data indicate that the actual impact on freedom-of-information requests has so far been minimal. The publishing of expenditure data has not increased the number of

²⁰³ Interview, a local government civil servant

²⁰⁴ Interview, a local government civil servant

requests, as some feared it would do.²⁰⁵ In addition, relatively few people are genuinely interested in going through vast amounts of information on government spending. Those who, in fact, are interested in it are likely to request that information regardless of whether the data is published proactively or not.

The release of public data and the engagement of citizens in governance are claimed to widen the public sphere where relevant discussion that leads to policymaking is taking place.²⁰⁶ The empirical evidence on whether this has been the case has been difficult to measure so far. However, from interviews of public officials, it has been possible to identify whether open-data principles have been successfully embedded into processes of data management and other public-sector work, thus improving efficiency in the long run.

In terms of public-sector agencies benefiting from open data internally, questions concerning internal leadership, officials' data literacy and relevant technological skills are highly relevant. Interviews of data producers suggest that it is by no means clear that public officials would understand the data and all the significant implications of the data. Many analysts argue, however, that by opening up datasets, employees with a certain level of data literacy would learn more about how the organisation works, i.e. financially, thus improving its efficiency. One interviewee in recent UCL research argued, "around a third to 50 per cent of entries on the budget used to mean nothing to anyone but accountants, but now authorities are forced to look into it, which is especially useful for members".²⁰⁷

It increasingly looks like the efficiency savings and improved quality in public management will not be achieved by 'armchair auditors' and other external pressure, but by the increased level of knowledge and understanding within organisations themselves. This takes a certain amount of leadership and vision.

"I think our view as a professional organisation is that benefits are probably going to be more realizable by those people who are actually engaged in public services, so that could be people from the voluntary sector, frontline

²⁰⁵ Worthy et al., 2011

²⁰⁶ Coleman, Emer: *From New Public Management to Open Governance – a New Future Or The More Things Change The More They Stay The Same?* University of Warwick, 2011, <http://davepress.net/FromNPMtoOpenGovernance.pdf>,

²⁰⁷ Worthy et al., 2011

workers in public services, empowered by having data openly available to them, and giving them the right sort of encouragement and training and so on for them to analyse and make use of data in new ways.”²⁰⁸

Several hurdles remain, however. As one senior official interviewed for this report put it:

*“We are not near the stage where open data would be embedded into decision-making processes or public-sector management or data management. My perception is that there are very few people in the government who recognise the value of data at all. There is an awful lot of projects that have stalled because of lack of technological knowledge”.*²⁰⁹

Also, it has proved difficult to get public organisations joining the projects that aim at opening up datasets and simultaneously improving the internal data-management processes. The level of collaboration between bodies has not been ideal.

“It has been difficult to persuade partner organisations to nominate open-data representatives, since those organisations are increasingly subject to financing cuts, which effectively decreases their interest to work for something they don’t see as part of their day-to-day work.”²¹⁰

An optimistic example comes from the London borough of Redbridge, where the level of performance has improved significantly after the officers learned how to access data and analyse data properly.²¹¹

“Redbridge is a good example where that has started to happen. They are making data available within the organization, and that is empowering people to think about problems and using new ways and starting to develop new relationships with people they are serving, informed by that data and the potential solutions to whatever the problems are.”²¹²

²⁰⁸ Interview, an academic/researcher

²⁰⁹ Interview, a local government civil servant

²¹⁰ Interview, a local government civil servant

²¹¹ Interview, a local government civil servant

²¹² Interview, a local government civil servant

The data-release system implemented by the council has emphasised easy access to data by both citizens and council officials. Initial experiences suggest that the feedback has been very positive and that council officials have adopted the system as a way of trading internal statistics between various departments, which has already created efficiency savings.²¹³ Officers also argue that the level of knowledge needed to implement a comprehensive data-release service is not excessively high.

“We are trying to debunk the idea that it is very technical and takes lot of expertise and resources. By getting more people to understand how simple it is to do it and what the benefits are, we will get more momentum in the service areas to try it.”²¹⁴

There are also arguments that state agencies, by nature, are not good at using data, and therefore there is a need for third parties who do understand data and can do all these wonderful things that make life easier.²¹⁵ This might be the case at the aggregate level, but arguably it only further illustrates the demand for more education of public officials. Successfully embedded data-release systems prove that it is far from impossible for public-sector organisations to manage data efficiently and in an innovative manner.

The question of data literacy meets support in the findings of the expenditure-data survey, which reveals that the level of understanding of open-data criteria is, in fact, not terribly high.²¹⁶ It seems that data producers believe that their data meets the open-data criteria, when in reality that is not necessarily the case.²¹⁷

Moreover, it seems that public officials don't see technological issues as part of their work.²¹⁸ In terms of effective data management, that can be a significant problem, since it is argued that in the future the boundaries between IT, records management and data management will become increasingly vague. Moreover, a certain degree of data management will arguably be something that each official is expected to be able

²¹³ Interview, a local government civil servant

²¹⁴ Interview, a local government civil servant

²¹⁵ Interview, a local government civil servant

²¹⁶ Halonen, 2011

²¹⁷ Ibid.

²¹⁸ Interview, a local government civil servant

to do. Producing good quality data is increasingly seen as being as important as other frontline services public authorities provide.²¹⁹

Problems arise in a scenario, where the parties don't interact with each other sufficiently or don't understand each other's issues. These problems have already been witnessed in the relationship between IT and record-management departments, although their work should be interrelated in many ways, especially when it comes to applying open data in a beneficial manner.²²⁰

“Most local authorities don't have anybody in the driving seat for information management, across the organization, to ensure they do get a join-up between the different disciplines.”

Steve Saxby fears that the change that the open-data initiative is expected to bring will not necessarily go far enough. The danger is that public officials perceive the initiative as too “geeky” and don't really understand it. What is required, according to Saxby, is “a refinement of the message so that the case of better access and use of PSI within government and beyond is tailored to departmental circumstances.”²²¹

Arguably there is a certain danger of the policy initiative getting stuck in a rut, if central government policy-makers and those public officials whose task it is to implement the policy don't speak the same language, or if officials look upon it purely as a cost, or something that just has to be done. In terms of the transparency agenda, it indeed seems like the expectations of central and local government haven't really matched so far. The British government has invested a lot of resources in attracting the best possible know-how into their central government, but as Saxby points out, it is far from clear that this know-how has been distilled into everyday policy implementation.²²² Hence, it is crucial to recognise the need for a fundamental change in the attitude, approach and policies.²²³

²¹⁹ Interview, an academic/researcher

²²⁰ See, for instance: Cummings, Kate & Findlay, Cassie: *Digital recordkeeping: are we at a tipping point?* Records Management Journal, Vol. 20, No. 3, 2010, pp. 265-278. and Moss, Michael: *'Without the Data, the Tools are Useless; Without the Software, the Data is Unmanageable'*, Journal of the Society of Archivists, Vol. 31, No. 1, April 2010, pp. 1-14.

²²¹ Saxby, Steve: *Three years in the life of UK national information policy – the politics and process of policy development*, International Journal of Private Law, Vol. 4, 2011, pp.1-31.

²²² Saxby, 2011

²²³ Ibid.

Another potential aspect of the internal benefits is indeed the concept of new information management. Steve Bailey has estimated that efficient records management creates significant efficiency savings for the whole organisation.²²⁴ Furthermore, there is no technical reason why the data management could not be implemented according to open-data principles: if all public information were released automatically into the public domain in open-data format instead of locking the data up in PDF's that are deposited in electronic record-management systems, there would be a significant decrease in the burden for freedom-of-information and record-management officers.

However, the status quo of many public-sector organisations is that the data-management software doesn't necessarily support the critical demands of open data. Hence, the cost of embedding open data into everyday data management could be significant. Moreover, the change from paper-based office environment to a digital one is far from over in many organisations. The issue of reliable digital long-term preservation is also very topical: Paul Jaeger and John Carlo Bertot point out that several American investigations have found that US government agencies "often have insufficient policies and procedures to ensure that digital-born information is retained in coherent fashion".²²⁵ All in all, the disciplines of records management, IT and data publishing at the moment don't interact sufficiently, and the roles and responsibilities within information management are far from clear. In particular, those organisations that have outsourced their IT, data and information management often lack the expertise needed.²²⁶

There is a strong chain of arguments in favour of digital data-management systems, which could involve all data regardless of its format and formal status as a record, and this has typically been important in deciding how to handle that piece of information. Difficulties might arise, however, in terms of procurement guidelines and whether the purchase of such software would fit into organisations' procurement policies.

In terms of data literacy, interviews suggest that in cases where individual officers are fully aware of the open-data principles the results also seem to be the best:

²²⁴ Bailey, Steve: *Measuring the impact of records management: Data and discussion from the UK higher education sector*, Records Management Journal, Volume 21, Issue 1, 2011, pp. 46-68.

²²⁵ Jaeger & Bertot, 2010

²²⁶ Interview, an academic/researcher

“It takes a lot of energy and some vision to see the wider importance of open data. If you are a busy official, then opening up data is not going to be near your list of priorities these days, unless you’ve really become interested in it”.²²⁷

As another interviewee explained, it is impossible to change officials’ mindsets by legislation – there has to be a personal side to the issue, and officials need to be able to see the internal benefits themselves.²²⁸

5.1.3 Transparency

Measuring transparency is a tricky task. It is not simply a question of the actual amount of information available, but also a question of equality in the accessibility and usability of that information.

The value of transparency increases when data is given a proper context and thus people can truly understand and use it. Therefore, as Jaeger and Bertot rightly argue, transparency must encompass all aspects of information access: physical access, intellectual access and social access.²²⁹ Physical access refers to the ability to reach content, social access to share content and intellectual access to fully understand content.²³⁰ The real value of open data for transparency lies in the fulfilment of these primary criteria.

When English local authorities were surveyed in early 2011, most of the respondents saw increasing transparency as the most significant positive impact of the open-data agenda.²³¹ Moreover, the majority of those who didn’t see any significant positive impact in the agenda were still supportive of the idea of increasing transparency as such. Interestingly, however, the information value of data was quite low since the proper context was not provided. This set-up proves to be an intriguing starting point for assessing the real transparency value of open data.

²²⁷ Interview, a local government civil servant

²²⁸ Interview, a central government civil servant

²²⁹ Jaeger & Bertot, 2010

²³⁰ Ibid.

²³¹ See [expenditure-data](#) case study

Development of any sophisticated and comprehensive method for measuring the quantitative value of open data for transparency was unfortunately well outside the scope of this study. However, the normative question of the justification of transparency is equally important and, in fact, more approachable in terms of the purposes of this study. It is highly debatable whether transparency can be measured quantitatively at all.

According to several democracy theorists, the fundamental justification of the democratic system lies in the principle of each citizen being able to see that their interests are treated equally in society.²³² To take this idea further, if a society is to be capable of enlightened discussion and decision-making, then its members ought to have a profound historical and societal knowledge as well. In order to achieve this, the information that forms human thought and ideas has to be as free and easily accessible as possible.

Furthermore, Robert Dahl argues that an advanced and justified democratic society would be possible only if demos had the same level of knowledge as the political elite have.²³³ Dahl goes on to say that access to information *per se* is not enough, that citizens must also have access to alternative sources of information, and that is information not under the control of government or dominated by any other point of view.²³⁴ Dahl's argument supports the call for opening up private corporate data alongside public government data.

Effectively, justification must encompass both reactive and proactive aspects of transparency, as well as the two-way open-data -powered transparency. The main epistemological difference between the "reading society" that reactive freedom-of-information legislation can provide and the "writing society" that an open-data regime could possibly provide lies in the source of initiative: in a reactive freedom-of-information regime, government is obliged to release public information that citizens ask for, whereas in proactive transparency, government decides to make a certain kind of information available to citizens, effectively undermining the need for freedom-of-information requests.

²³² See, for instance: Christiano, Thomas: *The Authority of Democracy*, The Journal of Political Philosophy, Volume 12, No. 3, 2004, pp. 266-290; Dahl, Robert A.: *Democracy and its critics*, Yale University Press, New Haven, 1989, pp. 84-87.

²³³ Dahl, Robert A.: *Democracy and its critics*, Yale University Press, New Haven, 1989, pp. 338-339.

²³⁴ Dahl, Robert A.: *On Democracy*, Yale University Press, New Haven, 1998, p. 97.

In a transparent society, the ideal is that every member of demos has an equal level of physical, intellectual and social access to information and can equally re-use public information and thus take part in public discussion. Following the Habermasian tradition of the public sphere²³⁵ and the criterion of equal advancement of interests, it is questionable whether a mere proactive disclosure really satisfies citizens' right to gain public information of their choice and thus freely and fully take part in the public discussion. When the initiative for granting access to public information lies in the hands of a small elite – as in the case of proactive disclosure – it is possible that citizens are not free to access public information of their choice and thus may not be able to satisfy their fundamental need to confirm their equal treatment. In other words, the knowledge gap between demos and the political elite prevails.

This potential risk of widening the gap in intellectual access – and hence reducing equality – is an issue that the open-data community can not afford to neglect. The physical and social aspects of information access can be approached relatively easily by means of enhancing the infrastructure of information management, but intellectual access drills into the core issues of equality and education and the social policies that aim to support them.

At the moment, it is hard to argue confidently for or against the increased level of transparency due to open data. Obviously, there is more data available for citizens, but whether that data holds relevant information value for all the public and thus really enhances all crucial aspects of transparency is another, considerably wider issue.

5.2 External benefits

5.2.1 Entrepreneurship

Arguably the most widely cited and influential paper on the economic impact of open data is the one by Rufus Pollock, in which he argues that releasing data at a marginal cost would result in revenues of ca. £6bn annually.²³⁶

²³⁵ See, for instance: Habermas, Jürgen: *The Structural Transformation of the Public Sphere. An Inquiry into a Category of Bourgeois Society*, MIT Press, 2001

²³⁶ Pollock, 2008

Most interviewees for this study agree emphatically with Pollock’s main argument of marginal-cost release. However, of all local-government authorities surveyed earlier this year, only 13% feel that releasing commercial value would be one of the open-data objectives most likely to be achieved by releasing expenditure data. The data producers’ perception is, of course, only one opinion on the subject, and there is obviously no comprehensive evidence on the economic benefits or on the possible lack of them.

In addition to Pollock’s paper, there are quite a few studies that argue in a similar fashion. The principal source of economic benefits seems to be the release of data at a marginal cost, which makes it possible for small enterprises to exploit data commercially.²³⁷ If there was a charging regime installed, economists fear that it would smother the SME’s and give a disproportional edge to the big corporations.²³⁸

“We need to promote entrepreneurship and compete in a global digital space. We need to keep a good digital infrastructure, and if we charge for data we would cut the entrepreneurial side of it, SME’s in particular. The Googles and Microsofts can operate anyway.”²³⁹

One emerging open-data ecosystem theory emphasises the role of big corporations, which provide the platform for smaller companies to operate on. In this model, open data enables the interaction but this activity still needs the technological platform that couldn’t necessarily be provided without big corporations like Google or Microsoft.²⁴⁰

Reports from around the post-industrialised world suggest that governments worldwide are looking for possibilities of supporting the private sector using open data. It is thus possible that a new form of national competition is arising.

“Advice for Finland: keep it free! The world is increasingly data, and you don’t have time to lose the competitive edge. Go free, go quickly. The return in charging is not high with all the costs included in the charging process.”

²³⁷ See, for instance: Koski, 2011.

²³⁸ *ibid.*

²³⁹ Interview, a local government civil servant

²⁴⁰ Interview, a central government civil servant

Developers are smart, and they will find a way of using data anyway, you can't stop them.” ²⁴¹

Governments can support SME's by releasing as much data as possible for free, and by focusing on datasets that are most beneficial for enterprises. Identifying these datasets is thus important, and increased collaboration with entrepreneurs and data providers is needed, ideally in a non-procurement setting: it is considered highly questionable whether bureaucrats are the best-informed people to decide what datasets are most beneficial for entrepreneurial purposes.

According to some, there is a risk of creating a vicious circle of endless enquiries on what datasets should be released first. Potential developers ask data providers what data do they have and providers ask developers what do they want to have, thus paralysing the whole initiative. In London, the London Data Store has aimed at solving this problem by consulting the developers before releasing anything and then focusing firstly on those datasets whose release was given the most support by the developer community. From then on, it was relatively easy to release the datasets as early as possible.²⁴²

In order to gain significant economic benefits, data that is opened must hold a certain market potential. Expenditure data, for example, is considered to have minimal market potential by local-government data producers. There are, however, practical examples available of businesses that operate by either analysing data or by creating applications for data, which only illustrates the fact that it is really difficult to consider what kind of datasets might be potentially most useful for enterprises.

There are a few examples of business models that companies use. Arguably the most typical one is to add value to raw data and consequently to monetise this in the following ways:

- a) By advert-funded web distribution
- b) By revenue-sharing with app developers who innovate with the company's API feeds
- c) By serving the operators and businesses in the market with analytics

²⁴¹ Interview, a local government civil servant

²⁴² Interview, a local government civil servant

As is the case with many other potential benefits, the economic benefits have been mostly based on estimates, so far. It is still very early days, and a positive outcome will only be achieved if certain criteria are met. Some critical views, however, have already emerged:

“Personally I am very doubtful that there is a great deal of opportunity to generate profitable new information products and services on the back of public-sector OpenData. I just can’t see colossal opportunities in bus departures, live train info, government expenditure data, planning, environment, health and crime data, which are some of the topics that have attracted the most clamour from advocates. Of course there will be some small businesses that will survive or flourish, but this is not the great information nirvana that we were promised by Rufus Pollock, Sir Tim Berners-Lee and Nigel Shadbolt. Maybe I am too impatient, but in the fast-moving world of digital with all that suppressed entrepreneurship, surely we should be seeing a bit more by now?”²⁴⁴

If it eventually turns out that the value of the positive economic impact has been overestimated, the need for measuring the normative value of transparency and democratic accountability is even more crucial. At this time, however, there are plenty of SME’s that operate with open data, and hence there is an area of research for more business-minded researchers than the author of this report.

5.2.2 Democratic accountability

The external benefit of democratic accountability is arguably related to the benefits and the initial objectives of freedom of information: in a democratic society, there is a fundamental need for citizens to be aware of how such decisions that affect their everyday life are being made and how the basic social infrastructure works.

²⁴³ <http://placr.co.uk/blog/2011/07/the-growth-case-for-open-data/>

²⁴⁴ Feldman, Steven: #OpenData – Who Profits?, <http://govinthelab.com/steven-feldman-opendata-who-profits/>

In order to achieve the desired benefits, the initial data-release phase should be undertaken with democratic impacts in sight. Data producers ought to identify who they want to access and use their data. If they want all citizens to be able to access and analyse datasets effectively, then they need to provide a simple front end for the data-release service. The open-data community easily overestimates the level of interest towards data and the abilities of citizens to understand data. Also, data must have information value that is relevant to citizens' democratic rights.

There are potential obstacles on the path to the data-powered ideal of democratic accountability.

Firstly, existing research suggests that the social effects of other forms of transparency, namely reactive freedom of information, have been relatively modest. The number of people who enforce their rights to scrutinise public authorities is really low. Generally, they are either organised civil groups – journalists, enterprises, NGOs – or they are individuals who are already interested in public policy and involved in public discussion by other means.²⁴⁵

Freedom of information therefore has not increased the number of people joining the public discussion and scrutinising politicians' and public officials' actions; it has merely provided better opportunities for those who are already empowered and active. FOI has not increased the general level of understanding of politics either.²⁴⁶ Whether that has been the case with open data is not within the scope of this study, but given the theories of John Zaller, it is highly probable that the increased amount of information available has merely strengthened the pre-existing attitudes of citizens and has not necessarily increased knowledge or understanding.²⁴⁷

Secondly, and partly relating to the first point, there is the question of data users' motives and whether they, in fact, support democratic accountability. Many observers share the concern that government data is mainly used for the purposes of "gotcha" journalism, where data is deliberately interpreted in a way that shines bad light on the government, even if there is no apparent reason for that.²⁴⁸ In particular,

²⁴⁵ Hazell et al., 2010

²⁴⁶ Ibid.

²⁴⁷ Zaller, John: *The Nature and Origins of Mass Opinion*, Cambridge University Press, 1992.

²⁴⁸ See, for instance: Fung, Archon & Weil, David: *Open Government and Open Society*, in Lathrop, Daniel & Ruma, Laurel (ed.): *Open Government: Collaboration, Transparency, and Participation in Practice*, O'Reilly

those public officials who are responsible for publishing data share this perception. Many of them argue that there are no real accountability benefits in open data if scrutiny is targeted at single spending issues without noticing the wider context around the spending.

However, in terms of journalists' – and other data users' – motives, one must remember the fundamental right to express one's views regardless of whether they support or undermine the actions of public officials or incumbent government. In that sense 'gotcha' journalism can not really be considered a threat to democratic accountability.

In terms of accountability, these two issues are, fortunately, not as worrying as they might sound. For the sake of accountability, it is important that the relevant data is out there and that people are aware of it and can access it effectively. Services like Theyworkforyou and OpenlyLocal have arguably lowered the threshold for scrutinising data and simultaneously increased the number of people capable of reaching crucial democratic data. As with the other potential benefits and application areas, the primary issue is the ability of people to access and use the data. If data is presented in a vague jungle of spreadsheets and csv files, the research suggests it is extremely unlikely that a large number of citizens would be interested in going through them.

Again, the question of measurement is interesting. Accountability by concept is not anything to be measured by the amount of datasets released or downloaded. In my perception, accountability derives from each citizen's ability to see that the government operates legally and treats her interests equally, hence it derives from the same normative ground as the legitimisation of democracy itself. When the crucial information that affects people's lives is increasingly in digital form, for the sake of accountability it is important that people have equal access to that data – that is relevant technology at hand and relevant skills. Public-sector data producers can enhance this access by providing data inventories and data portals comprehensive and simple enough for everyone to use.

Another issue regarding accountability is whether private sector data should be made subject to same transparency regulations as public sector data. It is increasingly unclear what can be labelled as a public task and what information should thus be available for public scrutiny. It could be argued, that private sector organisations hold such power and influence on wider society that their actions should be made subject to FOI. In case of open data, a service called OpenCorporates²⁴⁹ aims at making basic corporate information more transparent and it is likely that this issue will attract even more interest in the future.

5.2.3 Case: Open data and cultural heritage

Galleries, libraries, archives and museums (GLAM) form the nucleus of cultural-heritage organisations, which have significant amounts of relevant data that could be put into use for the common good.

These cultural-heritage organisations have not necessarily been the most visible actors in the open-data discussion so far, but the potential benefits of opening cultural-heritage data have recently gained more and more attention. The LODLAM movement (linked open data in libraries, archives and museums) has increasingly started to gain support: at the European level alone, there is a large amount of cultural-heritage data that could be released into the public domain, and its potential benefits for education purposes alone are massive.

In terms of opening up cultural-heritage data, the crucial issue is metadata and how to open it sustainably and efficiently. With sufficient metadata, it is possible to describe the whole collections and provide possibilities to create dynamic connections between different repositories and thus enhance access to digital material significantly. An example is Europeana²⁵⁰, a EU-funded initiative to create a single portal where citizens could freely explore the digital resources of European museums, libraries and archives. Europeana is powered by open metadata, which directs the user to the original source of the item. The goal is that access to all European cultural-heritage material would be possible through one data portal.

Other similar examples include Theatricalia²⁵¹, a British database of theatre performances, places and people, and JISC OpenART²⁵², an open database of British

²⁴⁹ <http://opencorporates.com/>

²⁵⁰ <http://www.europeana.eu/portal/>

²⁵¹ <http://theatricalia.com/>

artworks from 1660 to 1735. Both services are created under open-data guidelines and aim to enhance opportunities to access historically and culturally important information.

The open-data movement in cultural heritage seems to be intrinsically different in its goals and values to that in some other social areas. Cultural data is not intended to create massive financial benefits for any of the stakeholders but to help people with learning and educating themselves. It is a great example of the diversity of open data and how different goals can potentially be achieved with it.

5.2.4 Participation

What is political and social participation, and how should we measure it? If we accept the traditional definition of participation that, for instance, Robert Putnam used in his influential piece *Bowling Alone*²⁵³, then it is unlikely that open data has had any significant impact on the level of participation. There are, however, many calls for redefining participation from a certain duty-based citizenship to more of an engaged form of citizenship.²⁵⁴

The crucial question for initial consideration is whether the increased amount of data available and improved access to public information has changed citizens' behavioural patterns. Can it be plausible that a mere increase in opportunities to re-use public information could change those patterns significantly and, if so, in what way? We should bear in mind that overall citizens are extremely lazy in requesting public information via tools that are already available. In terms of data and particularly in its re-use, there is the extra barrier of possession of the relevant skills.

Russell Dalton argues in favour of a generational shift in the types of citizenship norms that people stress. People who reached the age of 18 by the end of World War II score higher on citizenship duty than younger generations, who emphasise values such as self-expression and freedom of choice. Where open data could hypothetically promote participation is in terms of the engaging type of citizenship instead of duty

²⁵² <http://www.jisc.ac.uk/whatwedo/programmes/inf11/infrastructureforresourcediscovery/openart.aspx>

²⁵³ Putnam, Robert: *Bowling Alone: The Collapse and Revival of American Community*, Simon & Schuster, New York, 2000.

²⁵⁴ See, for instance: Dalton, Russell: *Citizenship Norms and Political Participation in America: The Good News Is... the Bad News Is Wrong*, CDACS Occasional Paper 2006-01, October 2006, <http://www8.georgetown.edu/centers/cdacs/cid/DaltonOccasionalPaper.pdf>

citizenship. We should therefore examine the demographics of data users and their motivation to use data in order to approach this question.

The positive effect on empowerment and participation seems to be two-fold: while the number of data users remains low, in many cases during the initial open-data policy drafting phase the general public was involved in the decision-making process via various discussion events and policy workshops. For example, in the Greater London Authority, the initial phase of setting up the London Data Store was to engage with over 60 potential developers and consult with them about the form and purpose of the data portal. According to the Director of Digital Projects of the GLA, Emer Coleman, the consulting process was very successful.

In Manchester, the local open-data advocacy group Open Data Manchester have successfully set up fora for public authorities and developers to discuss open data and its usability in an “enlightened manner”.²⁵⁵ Similar examples can be found in several communities and local authorities. An example comes from the Open Kent project, where various developer days have been held. According to key people behind the project, Open Kent has operated with a fairly limited budget and limited resources so far, but nevertheless managed to set up a portal with a relevant API, which provides members of the public with an opportunity to search for data and re-use it in their chosen way. In terms of that objective, the project is arguably a success, but the project leader feels that the engagement of all stakeholders has been relatively low, which highlights the wider-scale key problems of the open-data movement.²⁵⁶

The key to success in all these examples of successful user engagement seems to be the high level of mutual trust between authorities and developers. Hence, they have been spared from the overly political gotcha-culture, which has daunted some of the open-data discussion (more on Chapter 6).

In that sense, open data has arguably already transformed the traditional culture of bureaucracy, moving it a bit towards the ideal of collaborative democracy introduced by Beth Noveck, for instance.²⁵⁷ Open data has been an excuse for public authorities to open up their policy-drafting methods, even if the data itself has not necessarily

²⁵⁵ Interview, an open data advocate

²⁵⁶ Interview, a local government civil servant

²⁵⁷ Noveck, Beth Simone: *The Single Point of Failure*, in Lathrop & Ruma (ed.): *Open Government: Collaboration, Transparency, and Participation in Practice*, O’Reilly Media, Sebastopol, 2010, pp. 49-70.

been opened. A similar scenario can be seen in the case of public consultation on open data, which has been going on during the autumn 2011.

“In the spirit of openness it’s not enough to have an exclusive group of experts, we need to talk to members of the activist community and also try to understand other bodies. How can we develop this in an open way to make sure that we’ll get it right?”²⁵⁸

Open data has effectively removed the monopoly status of public authorities in terms of rendering information. That has not, however, resulted in a revolution of citizen-created public services that some politicians and activists alike have seemingly hoped for.

“There was this notion of getting people to use data, and you would be able to kick-start this new type of engagement. What we have seen with the release of government data is that this has not happened. There was this very naïve notion that if you release data, then people will go and do stuff. People have to earn a living.”²⁵⁹

In terms of public participation, it seems that the most beneficial aspects of open data have been relatively small-scale innovations: things that make everyday life just a tiny bit easier and more convenient.²⁶⁰ Also, in terms of increasing the number of politically-engaged citizens, it is hard to find evidence of whether open data has played any significant part in this. Thanks to open-data powered services such as Theyworkforyou and OpenlyLocal, it is easier for citizens to get relevant information and possibly therefore join the public discussion, but whether there is any empirical evidence for this happening remains to be seen.

Wider political change that has arisen due to data is extremely hard to prove, and so far the research indicates that expectations of the data-powered political implications have been overly optimistic.²⁶¹ It very well may be that open data is only a part of a

²⁵⁸ Interview, a central government civil servant

²⁵⁹ Interview, an open data advocate

²⁶⁰ Interview, an academic/researcher

²⁶¹ See, for instance: McClean, Tom: *Not with a Bang but a Whimper: The Politics of Accountability and Open Data in the UK*, paper presented at the American Political Science Association Annual Meeting, Seattle, Washington, 1-4 September 2011, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1899790

bigger change in culture and attitudes, but not a significant catalyst for change in itself.

“Lots of open-data stuff hasn’t made much of an impact on politics. The reasons for that are that there are actually quite significant barriers in its use.”²⁶²

The question of data users and their motives is thus essential. Davies has identified six main motivational clusters of open government data users in the UK:

- Government focused
- Technology & innovation focused
- Reward focused
- Digitizing government focused
- Problem solving
- Social or public-sector entrepreneurialism²⁶³

Despite the observation that open government data has allowed “new actors using data to enter public debate”, Davies argues that those actors are already empowered individuals who are technically skilled and interested in issues.²⁶⁴ The case resembles the identification of FOI users, who are also found to be those who are already engaged and politically active.²⁶⁵ The threshold for using data is relatively high, and therefore further education is badly needed in order to gain a more data-literate public.

Ben Worthy et al. have identified the main users of local-government expenditure data by interviewing public officials who publish the data. Findings further illustrate the need for engagement between data producers and potential users, since data publishers simply don’t know who uses their data at the moment.²⁶⁶ Among the

²⁶² Interview, an academic/researcher

²⁶³ Davies, Tim: *Open data, democracy and public sector reform. A look at open government data use from data.gov.uk*, University of Oxford, August 2010, p.23. <http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf>

²⁶⁴ Davies, 2010

²⁶⁵ Hazell et al., 2010

²⁶⁶ Worthy, Ben et al. *Town Hall Transparency? The Impact of Freedom of Information on Local Government in England*, The Constitution Unit, December 2011. <http://www.ucl.ac.uk/constitution-unit/research/foi/foi-and-local-government/town-hall-transparency.pdf>

identified users, the biggest groups seem to be journalists and members of the general public.²⁶⁷ The use of data by businesses seems to be significantly lower.²⁶⁸

Open data therefore does not necessarily increase the actual number of empowered citizens or the number of civil-society institutions engaging with data. However, the quality of engagement can be increased due to the better opportunities for those who would engage anyway. Furthermore, it is quite unclear whether the model of civic participation would work in a situation where financial rewards are unlikely.²⁶⁹ Arguably, there has to be a personal connection with the data, in other words, data users download and use data for their own personal benefits.

The causal connection between public data provision and participative democracy is therefore far from clear. There may well be benefits, but so far the evidence has been lacking and, according to some, the level of research on the social implications of open data has so far been insufficient.²⁷⁰ Open data can be seen to promote a certain type of engaging citizenship among younger people who already possess both the skills to use data and an interest to do that. A large-scale boost in participation, whether based on duty or engagement, is harder to single out.

It may well be that the sheer number of data users is irrelevant, but the quality of user engagement in policymaking and the efficiency in creating smaller-scale implications and spreading them to the wider society might be the way to go in the future. In order to achieve this, the data that is initially released should be data that potential developers find useful and interesting. Moreover, the examples suggest that there has to be a deep understanding and trust between public authorities who publish data and developers who use data. It might even be that the best examples are created by those who are already engaged in public services and the political sphere.

In the future, social scientists should apply sophisticated empirical methods in measuring the exact impact of open data on public participation. One potential way would be to measure behavioural patterns of different groups of people and see whether data users' patterns differ significantly from others'. The change in data users' own political behaviour should also be measured in order to exclude the

²⁶⁷ *ibid.*

²⁶⁸ *Ibid.*

²⁶⁹ McClean, 2011

²⁷⁰ Interview, an academic/researcher

possibility that data users are politically more engaged to start with, as might be the case, given the findings of Davies.

It seems evident that in order to generate any real benefits for participation, the threshold for using data effectively must be as low as possible. It is questionable whether that threshold can be lowered enough for all citizens to be able to use data in a relatively short time period. At the moment, it is thus advisable to concentrate on those potential data users who are already engaged in the public sphere in some form, be it within the public sector or voluntary organisations or some other organised civil-society movement. The potential users need to have a certain level of resources and knowledge available. It is highly unlikely that anyone without the skills or financial resources to spend time on scrutinising public data would actually start doing it. More education is needed, but before the overall data skills are sufficient, the initial focus should arguably be on those who are already engaged.

5.2.5 Case: Guardian's Data Journalism

One large group of data users who are already engaged in the public sphere are journalists. Open data is claimed to provide some completely new ways for journalists to interact with citizens and to provide news. This movement, which has been partially ignited by the open-data movement, is called data journalism.

Data journalism derives from the concept where journalists create news that is based on analysis of open datasets. Most of the articles are published online and are open to instant feedback by readers. Articles are either in a traditional text form, or perhaps more often, in a form of a data visualisation or a web application. Helsingin Sanomat, the leading Finnish newspaper, has described data-journalism application as a 560x400 ppi application, which is encouraged to be coded by using a Javascript-PHP-MySQL-Apache combination and which should be embedded into an Iframe application.²⁷¹

In the UK, The Guardian has been the pioneering newspaper in the field of data journalism, after they started their Datablog in 2009. According to Simon Rogers, the editor of Guardian Datastore and Datablog, the main things in data journalism are its flexibility in searching for new ways of storytelling and secondly its ability to make data more accessible and understandable to the general public.²⁷² The exponential increase in

²⁷¹ <http://blogit.hs.fi/hsnext/aihe/hs-open>

²⁷² Guardian Datablog, 28 July 2011: *Data Journalism at the Guardian: what is it and how do we do it?*

the amount of data available has arguably made it more important to be able to sieve the relevant data from the masses and to present it in an understandable way. Moreover, Rogers argues that open data has effectively made it possible for everyone to analyse relevant data and create information. In a sense, open data has created a more democratic or egalitarian type of journalism. In order to gain maximum benefits for journalists' purposes, data that is published should be as disaggregated as possible.²⁷³

Arguably, the most visible example of The Guardian's data-journalism aspirations has been the WikiLeaks case, although it wasn't really started by open data. In 2010, the online whistleblowing service WikiLeaks published thousands of classified documents in huge datasets, which several journalists filtered into public use by creating search platforms and visualisations that were based on the original datasets. It has been argued that without journalists' efforts the impact of release would have been much smaller, and potentially insignificant.²⁷⁴ Similar approach is used also in other data-journalism cases, for example in the joint project between The Guardian and London School Economics, where London riots of August 2011 were analysed.

However, several questions remain that are yet to be comprehensively answered. On one hand, it is believed that release of government datasets undermines the role of established media, since every citizen is now expected to be capable of obtaining public information and spreading and re-using it via various online tools. On the other hand, open data offers an array of new opportunities for journalists, and it is assumed that the role of journalists will transform into one of a data analyst, who filters relevant information from the data to the greater public. A new set of skills is needed in order to find newsworthy information from the vast datasets effectively.

Furthermore, the rise of new digital media is feared to cause the knowledge gap to grow even more significant²⁷⁵, which underlines the importance of understanding journalists' role in the new information provision regime. Also, it is increasingly important to analyse journalists' ways of using open data in order to eventually assess its wider impact on society, especially since the journalists' use of freedom-of-information requests has even been seen to result in decreasing public trust of government.²⁷⁶

<http://www.guardian.co.uk/news/datablog/2011/jul/28/data-journalism>

²⁷³ Interview, a news editor

²⁷⁴ Roberts, Alasdair: *WikiLeaks: The Illusion on Transparency*, International Review of Administrative Science, 2012.

²⁷⁵ Prior, Markus: *News vs. Entertainment: How Increasing Media Choice Widens Gap in Political Knowledge and Turnout*, American Journal of Political Science, Vol. 49, No. 3, July 2005, pp.577-592.

²⁷⁶ Hazell et al., 2010

Journalists are one of the largest single groups of freedom-of-information requesters and thus hold great power in terms of filtering public information for the greater public. It is estimated that only about 0.02% of UK citizens have made a request themselves, but most of the population is still aware of the concept of FOI and FOI legislation. The most typical way of citizens getting hold of information released through FOI requests is through mass media.²⁷⁷

According to some research, journalists' use of FOI tends to decrease public trust in government, with only 3% of the sample indicating an increase and 58% a decrease.²⁷⁸ The MPs' expenses scandal is a frontline example of a high-profile case powered by journalists' FOI requests. At the local council level, journalists prefer to request information on council spending, which is seen to underline either the journalists' pursuit of holding officials to account or possibly a wider conflict between parts of the mass media and government officials. Similar findings have already been made in the case of journalists' use of expenditure data.²⁷⁹

Strictly in terms of journalists' use of FOI, Holsen et al. have interviewed journalists and argue that the FOI Act has merely given another chance of obtaining information, but in general it has made little difference to their reporting.²⁸⁰ Studies referred to in the introductory section have also concluded in a similar fashion, which clearly leaves space for a study with more direct emphasis on journalists' use of open data and changes that might have occurred since implementation of open-data schemes.

Alastair Roberts has argued in favour of newspapers' importance in distilling newsworthy information to the general public from increasing amounts of raw data. He uses WikiLeaks as an example: he argues that without several newspapers collaborating with WikiLeaks, the leaks would have remained insignificant in terms of citizens' awareness of the issues. According to Roberts, the general public is not as eager to obtain information from raw datasets as some may have believed or hoped, not even when information concerns highly controversial issues.²⁸¹

It seems quite widely accepted that highly-informed and well-educated citizens look for news-related information online while people with lower socioeconomic status prefer to

²⁷⁷ Hazell et al., 2010

²⁷⁸ Hazell et al., 2010

²⁷⁹ See, for instance: Halonen, 2011

²⁸⁰ Holsen, MacDonald & Glover: *Journalists' use of the UK Freedom of Information Act*. Open Government Journal, Volume 3, Issue 1, 2007.

²⁸¹ Roberts, 2012

look for entertainment thus skipping news-related content.²⁸² Accordingly, it is presumably unlikely that people who haven't submitted FOI requests would actively look for government datasets. Journalists' role in filtering and providing information should thus be greater than in the analogical world, where citizens in general used to look for more news-related information than in the digital world. Internet use is also much more fragmented than use of the old-fashioned media has been.²⁸³

Data journalism as a concept is very much taking shape, and many guidelines are yet to be created. The trend of readers using more digital sources to find information than traditional newspapers seems strong, and in this trend data journalism can prove to be a significant factor. Internet as a platform gives much more opportunities for flexible and interactive information provision than paper-based newspapers.

This chapter has examined the benefits of open data and simultaneously discussed whether these benefits are realistically achievable. Despite the generally positive approach to open data, some problems have also emerged during the take-off process. In the next chapter, I will further identify the main causes of concern and argue how these problems could be solved.

²⁸² Prior, Markus: *News vs. Entertainment: How Increasing Media Choice Widens Gap in Political Knowledge and Turnout*, *American Journal of Political Science*, Vol. 49, No. 3, July 2005, pp.577-592.

²⁸³ Ibid.

6. Emerging problems and how to solve them

The pursuit of best practices can sometimes be best achieved by first identifying some examples of the worst practices, or at least some problems and setbacks that have emerged. It is not the purpose of this report to be overly negative towards the development and applicability of open data, but it is essential to realise that not all decisions have been ideal and to find answers on how things could be done better.

6.1 Confusion on policies

One of the biggest problems perceived by both data publishers who were surveyed for expenditure-data research and those interviewed solely for this paper was the inadequate advice they received from central government.²⁸⁴ As one respondent had put it:

“Like most government initiatives it [the transparency agenda] was made up on the hoof. Initially there was no clear guidance. When guidance did emerge, it was obviously flawed and has changed several times since.”²⁸⁵

Arguably, there are many issues that affect the overall efficiency and feasibility of open-data policies. In the context of the UK, several historical and political factors must be taken into consideration, as well as the vague relationship between local and central government and the institutions and organisations working within them. The level of mutual trust between local and central government authorities is also debatable.

Most of the interviewees in this study agree that there have been several problems in terms of getting the organisational policy responsibilities clear. The vagueness of policies is considered highly problematic in terms of gaining true benefits:

“From my point of view the main thing is to influence policies on open data. There’s no legal obligation to publish open data, and until then it’s not going

²⁸⁴ Halonen, Antti: *Data without context? A study on the UK local government transparency agenda and its impacts on information management*, Unpublished MRes dissertation, University College London, 2011.

²⁸⁵ Ibid.

to happen uniformly or successfully. Until that happens, we won't see the real benefits. That's where we should be focusing.”²⁸⁶

The emergence of the Public Sector Transparency Board and Public Data Corporation first, and later the Public Data Group and Data Strategy Board, was hoped by some to solve this, at least partly, although the initial purpose of the PDC itself is subject to fierce criticism. It is feared that the establishment of such a centralised agency will merely benefit the existing bureaucratic power structures.²⁸⁷ It is questionable whether this framework would result in the “creative destruction” that is claimed to be necessary for successful policy reform.²⁸⁸

Drawing from the experiences of the UK, the perceptions of UK-based open-data actors and existing theoretical research on public policy, I argue in favour of strong leadership within organisations, clear guidelines, a strong institutional basis and a comprehensive licensing framework for opening up datasets. It is essential that potential developers have clear knowledge on how and where they are allowed to use and re-use data and that data producers are well aware of data management, data protection and publicity issues, which have so far proved to be relatively unclear for many of them.

Data producers should have clear guidance on what data they should publish first, or preferably be encouraged to open up all datasets that don't have secrecy or privacy complications. Focus on expenditure data has seemingly led to a situation where data producers feel that they have already met the requirements, even if they still have lots of good data that would be more beneficial to users than expenditure data is.²⁸⁹ Moreover, the level of awareness of the definition of openness should be increased. Currently, there are many data publishers who genuinely believe that their data is open when, in fact, it is not.²⁹⁰

One solution to the integrated data policy issue would be to centralise the formulation of national information policies under one umbrella organisation. Arguably, this has been the case with the Office of Public Sector Information, which

²⁸⁶ Interview, a central government civil servant

²⁸⁷ Interview, a local government civil servant

²⁸⁸ See [Chapter 4](#).

²⁸⁹ Interview, a local government civil servant

²⁹⁰ Halonen, 2011

was later merged into the National Archives. Janssen, who argues in favour of integrated information policies across the EU, makes a similar argument.²⁹¹ However, she warns of the possibility that a single information policy could potentially lead to a scenario where re-use regulation – if based on charging – could take priority and thus the public would have to pay for data on more occasions than would have been the case under the traditional freedom-of-information legislation.²⁹² Hence, it is important to focus heavily on the issue of charging and the very definition of openness, if it is desired to keep data truly open.

6.2 Privacy issues – a non-issue?

Marco Fioretti argues that privacy issues, in terms of open data, have nearly always been a non-issue.²⁹³ His argument is based on the perception that datasets that are opened up are intrinsically not personal data. A similar argument has also been made in the case of other sensitive information, e.g. data concerning national security.

Fioretti's argument has a point, but misses one important aspect of open data, the interconnectedness that is achieved by linked data. Even if one particular dataset is not personal information, it is possible that after linking a myriad of datasets together certain individuals can be identified. Thus the concern about privacy is not necessarily unwarranted, and, in fact, it is an increasingly difficult question to address comprehensively.

A new dimension to the privacy discussion emerged after the Autumn Statement, when the Chancellor announced the sharing of certain medical records in unidentifiable form for private firms' and analysts' re-use. However, it is questionable how unidentifiable these datasets will be in the long run. Moreover, there are arguments that this data sharing, in fact, has nothing to do with the definition of "open data" as used e.g. in this report, but discussing it as part of open-data policies can blur the meaning of openness. Potential privacy risks that are involved in the sharing of personal data can thus lead to a loss of public trust in open data.²⁹⁴

²⁹¹ Janssen, 2011

²⁹² Janssen, 2011

²⁹³ Fioretti, Marco: *Open Data: Emerging trends, issues and best practices*, http://www.lem.sssup.it/WPLem/odos/odos_2.html

²⁹⁴ See, for instance: <http://www.timdavies.org.uk/2011/12/02/3090/>

As Kieron O’Hara concludes, an intrinsic aspect of transparency is indeed trust, and – according to him – this trust cannot be achieved if privacy protection is not embedded in the transparency programme itself.²⁹⁵ Effectively, this means being transparent on transparency, for example by opening up the methods and tools of anonymising.

6.3 Politicised environment and gotcha culture

In terms of media publicity, the most widely reported open-data issue so far has arguably been the refusal of Nottingham City Council to publish its expenditure data as it was requested to by the Department of Communities and Local Government.²⁹⁶

Instead of discussing the potential benefits of transparency and accountability, the main arguments in the Nottingham case were highly politicised. The Conservative-led DCLG accused the Labour-led Nottingham City Council of wasting public money and hiding behind a veil of secrecy, while the council leaders claimed that the whole transparency agenda is just window-dressing and a Conservative gimmick for further cuts.²⁹⁷

This kind of debate becomes problematic if the participants take too partisan a tone. As, for instance, John Zaller²⁹⁸ and Angus Campbell et al²⁹⁹ have theorised, highly partisan individuals tend to see social issues through a “perceptual screen”, which effectively means that their perception of a piece of public policy is more dependent on their political affiliation than the substance of the policy itself.

A case like open data, which is not traditionally a left- or right-wing policy, can suffer from a highly partisan environment. Open-data advocates also don’t perceive themselves as political figures. It is possible, however, to identify some trends in the

²⁹⁵ O’Hara, Kieron: *Transparent Government, Not Transparent Citizens: A Report on Privacy and Transparency for the Cabinet Office*, September 2011,

<http://www.cabinetoffice.gov.uk/sites/default/files/resources/transparency-and-privacy-review-annex-b.pdf>

²⁹⁶ Halonen, 2011

²⁹⁷ See, for instance: Nottingham City Post, 24 May 2011: *Council’s anger as credit-card spending row ‘backfires’*; Nottingham Evening Post, 15 April 2011: *Councillor and minister in TV clash on spending*; BBC, 4 February 2011: *Ministers step up criticism of Labour’s Nottingham fortress*, http://www.bbc.co.uk/blogs/johnhess/2011/02/the_war_of_words_between.html; BBC, 29 March 2011: *Pickles threatens Nottingham council with court action*, <http://www.bbc.co.uk/news/uk-england-nottinghamshire-12900725>; Parliamentary debate on 9 February 2011, HC Deb, 9 February 2011, c383

²⁹⁸ Zaller, 1992

²⁹⁹ Campbell et al: *The American Voter*, University of Chicago Press, 1960.

development of open data that have connections to either the communitarian left, which emphasises the sharing of common resources, or the libertarian right-wing movement, which has traditionally supported small government and the cutting of bureaucracy. Moreover, the open-data movement does not operate in a political vacuum - it is not apolitical, and different political streams behind the movement must be examined. The release of expenditure data, in particular, is claimed to be driven by mostly political motives.

“The open-data movement thus far – and I’m basing this on conversations with members of the movement – has primarily been concentrating on overcoming technical barriers, and it has not really concentrated on overcoming social and political barriers. That will have to change. To give an example, two or three months ago there was an article that was published in the Canadian Public Administration Journal, which said look, the rhetoric of the open-data movement is – when you boil it down to its very essentials – very similar to the rhetoric of advocating New Public Management. [...] There is confusion among the open-data movement because they didn’t see themselves in this light at all. All they assumed was that open data would be this very democratic and liberating thing, so there is a degree of political naivety involved.”³⁰⁰

In addition to the political dimension, the level of trust in politicians and public authorities could potentially dictate the discussion. As Chapter 4 pointed out, it is highly questionable whether any of the preceding e-Government initiatives have resulted in an increased level of public trust, and it is thus unlikely that open data would provide significantly different results in that area in the short term. It is naïve to assume that merely by releasing public data the level of public trust would be automatically improved.

It is therefore vital to understand not only the substance of the released information, but also the logic of the principal data users, including the media. The first phase of the transparency agenda concentrated on expenditure data, and one of its most visible impacts so far has been the number of stories on public spending and whether it has been justified. Many data producers, however, criticise the sensationalism of

³⁰⁰ Interview, an academic/researcher

public affairs and feel that the media try to deliberately create dramatic headlines about spending just to increase their sales.³⁰¹

The media is not solely to blame for the gotcha culture. It is hard to believe that the political cat-and-mouse game that has been going on between the DCLG and local authorities would result in increased trust in public authorities in any way. Talk of “stopping reckless waste” adds nothing to the overall value of open data, but it is an essential part of the political culture.

The history and political culture of the UK indeed suggest that there is a deep mistrust between local and central government.

“There is a kind of ideological tension in how central government feels that it should be able to tell local government what to do and equally local government tend to be reluctant because of their traditional autonomy and status as an elected body.”³⁰²

Evidently – given the historical and political frameworks in the UK – it has not been an ideal decision to start the transparency agenda with an initial focus on expenditure data. There has been the clear risk of the whole open-data issue and its potential benefits getting buried underneath the debate on expenditure data. Many data producers in local government lack the motivation to continue with open data and feel that the job is done when expenditure data is published, regardless of whether the data format meets the definition of openness or not.

“No, honestly, when they first announced the release of spending data, I didn’t think that this would a very good idea. It seems to be purely political rhetoric.”³⁰³

To summarise, in order to get sustainable use of data for public good, a certain level of trust between public authorities and data users must be established. Also, public authorities should not see citizens as mere annoyances who try to obstruct their daily work, and equally citizens must give up the perception of civil servants and

³⁰¹ Halonen, 2011

³⁰² Interview, a local government civil servant

³⁰³ Interview, an open data advocate

politicians as crooks who deliberately aim to waste public money for their own personal benefits, when overall that clearly is not true.

“What you have to do in this relationship is to create trust. You’re trying to create an environment where the public sector and private developers can meet in a non-procurement setting. Generally what happens is that public officials think that they are going to be criticised or feel that someone will try to sell them something. Neither of those situations is particularly desirable.”³⁰⁴

6.4 Data hugging

“The main challenge is to change the culture in government.”³⁰⁵

According to interviews of those who are close to public-sector data publishing, one of the main barriers – if not the main barrier – for effective data production has been reluctance on the part of public officials. Interestingly, politicians have been relatively supportive of the open-data agenda, but the civil service has not.³⁰⁶

“Biggest obstacle? Public officials. Period.”³⁰⁷

The Weberian tradition of bureaucracy indeed suggests that bureaucracy is a somewhat secretive organism by nature: it protects its own power and capabilities by keeping crucial information close to itself. Moreover, the British civil service has arguably enjoyed a reputation of upholding a certain degree of secrecy during the last century.

“The British state regulated the economy and its relations with civil society through secrecy for decades.”³⁰⁸

³⁰⁴ Interview, an open data advocate

³⁰⁵ Interview, a local government civil servant

³⁰⁶ Interview, a local government civil servant

³⁰⁷ Interview, a local government civil servant

³⁰⁸ Interview, an academic/researcher

However reluctant some public officials are to release datasets, there is no need to generalise about the whole civil service as overly secretive or being against the open-data movement. In some cases, civil servants have, in fact, led the initiative. Yet the status quo is unlikely to be changed if there is no strong legislative obligation to open up data – the release of datasets means additional work, and in the case of local-government expenditure data the public officials’ perceptions were significantly influenced by their own interests. In addition, as already noted, public officials are cautious towards any potential risk of being accused of some wrongdoing. There has to be a push to open up data but also support in terms of helping officials to open up the right data with data-protection regulations taken care of.

Interestingly, Alon Peled has argued that in many cases the public officials, in fact, have nothing against transparency *per se* - that is being transparent to the general public. However, according to him, there is a certain inter-departmental battle taking place within public-sector organisations.³⁰⁹ Open data could significantly help public officials in transferring information between organisations, but this inter-departmental battle seems to hinder the development.

Moreover, the open-data revolution should be able to cover the whole stretch of public-sector work. There are officials who are genuinely interested in open data and argue emphatically for it. There just needs to be both a personal side to it and also a legislative push, as well as the support already mentioned above.

“I think we recognize this has the potential to fundamentally change how and what we’re doing at the government.”³¹⁰

Former director of digital engagement, a member of the Transparency Board, Andrew Stott calls the reluctance of releasing data “data hugging”.³¹¹ He argues that authorities come up with excuses for not releasing data, though not necessarily deliberately. These reasons include:

- Not enough resources
- Release threatens privacy of individuals

³⁰⁹ Peled, Alon: *When Transparency and Collaboration Collide: The USA Open Data Program*, Journal of the American society for information science and technology, Vol. 62, Issue 11, November 2011, pp. 2085-2094.

³¹⁰ Interview, a central government civil servant

³¹¹ Andrew Stott at the Open Knowledge Foundation Conference in Berlin, June 2011.

-
- Data quality is not sufficient
 - Bad advice from central government
 - There is no interest in data
 - Data ownership – “it is our data!”

Data producers in local authorities did indeed use similar arguments when they were asked to describe their perceptions on the transparency agenda!³¹²

In some cases, the question of resources might be justified. Public authorities must be able to focus on their frontline services, and quite a few have difficulty in seeing the opening up of their datasets as a frontline service, especially when they are experiencing heavy financial cuts. If it turns out that in the long run there is no proper financial return on investment for opening up datasets, it may well be that the usefulness of the whole agenda will be reviewed. However, one must bear in mind that it is considerably easier to come up with costs than it is to measure exact benefits, and therefore the resource argument must be taken with a pinch of salt. The question of costs is also related to the level of IT infrastructure already in place. With a significant investment in agile government IT solutions, it is possible to lower the costs in the long run.³¹³ Here again, the difficult reality concerning the level of IT knowledge within organisations is apparent.

The examples of Redbridge and the London Data Store suggest that the resources needed to set up coherent data portals are not too high. According to the London Data Store, the financial sacrifice needed to set up a web interface was not at all excessive. More than financial resources, it is a question of human resources: will and sufficient technological competence to do it.³¹⁴ Moreover, the normative side of the question should not be overlooked.

Interviews suggest that the cultural change can best be achieved at an individual level. So far, the open-data movement has gone forward mostly thanks to a few enthusiastic individuals and leaders, who have pressured their respective agencies from within. Emer Coleman suggests that the culture shift in bureaucracy indeed

³¹² See the [expenditure-data](#) case study

³¹³ Stephen, Justine et al.: *System Error – Fixing the flaws in government IT*, Institute for Government, 2011, <http://www.instituteforgovernment.org.uk/publications/23/>

³¹⁴ Interview, a local government civil servant

needs a change in attitudes towards risk.³¹⁵ Transparency can expose wrongdoings and mistakes, but public authorities should see this more as an opportunity than a threat.

“Open data is a brave new world and hasn’t been explored. Officials would have to give up their control, and therefore we have to convince very risk-aware people.”³¹⁶

In order to change this culture, there is a need for strong grass-root examples to firstly convince ministers and key politicians, who can then start talking about open data and eventually convince the civil service. Persuading officials is thus considered a good idea, but the problem is that legislation can not achieve this change in mindset.³¹⁷ Jonathan Gray from the Open Knowledge Foundation has described the desired method as “not necessarily hard legislation, more like soft policy”.³¹⁸

However, this scenario of change requires a change in attitude from civil society, as well. The “gotcha” culture, where mistakes by public authorities are overly scrutinised, does not lead to a more trusting and open society, as I have argued above. The issue of mutual trust is crucial.

The questions of data ownership and control are interesting. Some public authorities perceive data as a product of their labour and therefore resist the idea of releasing it to the greater community for the purposes of free re-use.³¹⁹ There is reasoning behind this reluctance, and typically it derives from the idea of selling datasets in order to fund other public services. Similar reasoning is evident in the possible charging regimes proposed in the Public Data Corporation consultation. However, one interviewee from the civil service criticises this approach and argues that, in the case

³¹⁵ Coleman, Emer: *From New Public Management to Open Governance – a New Future Or The More Things Change The More They Stay The Same?* University of Warwick, 2011,

<http://davepress.net/FromNPMtoOpenGovernance.pdf>

³¹⁶ Interview, a local government civil servant

³¹⁷ Interview, a central government civil servant

³¹⁸ Gray, Jonathan: *Interview for University of Southampton open data study*,

<http://jwyg.okfn.org/2011/12/06/interview-for-university-of-southampton-open-data-study/#comments>

³¹⁹ Interview, a local government civil servant

of most public agencies, data selling is only opportunistic and not by any means necessary for the authority.³²⁰

Secondly, civil servants seem to be reluctant about the idea of losing control of the data. Typical arguments have related to concern about what will happen to data if it is released openly. They are worried about the possibility of fraud and all other kinds of misuse.³²¹ According to interviewees, however, this is only a part of the culture change that is necessary.

6.5 Data quality and context

“There are some technical aspects, the formats of data and data classification. They are pretty fundamental things that I believe need to be addressed in order to enable publishing of open data in the most beneficial way.”³²²

A general knowledge of information studies tells us that data without a proper context is relatively meaningless in terms of its information value.³²³ With a proper context, however, data can gain certain information value, and ultimately can lead to increased knowledge and wisdom. Visualisation of this basic information hierarchy can be seen in Figure 2.

³²⁰ Interview, a local government civil servant

³²¹ Halonen, 2011

³²² Interview, a local government civil servant

³²³ See, for instance: Ackoff, R.L.: *From Data to Wisdom*, Journal of Applied Systems Analysis, Volume 16, 1989, p. 3-9.; Zeleny, M.: *Management Support Systems: Towards Integrated Knowledge Management*, Human Systems Management, Volume 7, 1987, pp. 59-70.

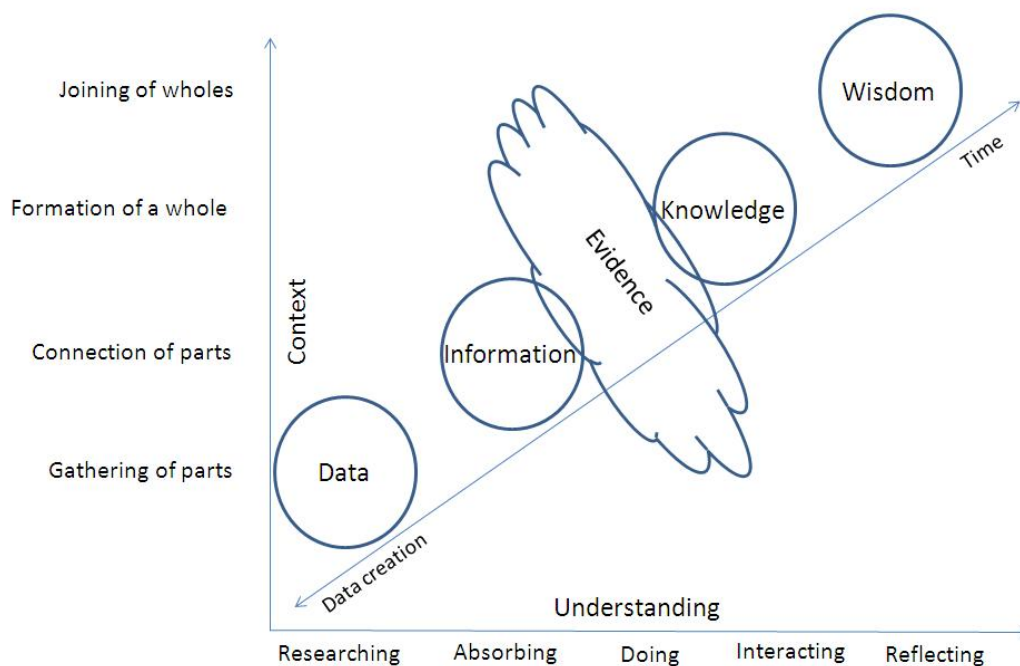


Figure 2: Visualisation of information hierarchy³²⁴

It has turned out that overall many public-sector data producers are worried about data quality and whether there is any proper use of data. In the case of expenditure data in particular, the context and quality do not seem to be there, and therefore the return on investment is claimed to be insufficient.³²⁵ One data producer who was interviewed for the expenditure-data survey summarised the case as follows:

“Public trust requires that data is supplied in context. The current process does not do this. Data of questionable quality will be twisted and edited to say whatever the individual wants it to say and related FOI requests are inevitable.”³²⁶

The Public Sector Transparency Board and many open-data advocates have taken a stance that data producers should not worry too much about data quality but first emphasise the publishing itself. The improvement of quality would then come from active users, who would give feedback to data producers and possibly even provide

³²⁴ Basic model adapted from Hey, Jonathan: *The Data, Information, Knowledge, Wisdom Chain: The Metaphorical Link*, published at Intergovernmental Oceanographic Commission- OceanTeacher: a training system for ocean data and information management, December 2004.

³²⁵ Halonen, Antti: *Data without context? A study on the UK local government transparency agenda and its impacts on information management*, Unpublished MRes dissertation, University College London, 2011.

³²⁶ Ibid.

relevant metadata. If the data is out there early enough then its quality will be improved, the argument goes.

“Data is always incorrect. If you worry about the quality, then you will never put it out.”³²⁷

“If we don’t publish it, its [quality] is never going to be improved, that is the reality.”³²⁸

However, this will need a certain number of active data users who are able to improve the quality of data effectively. If the demand is not there, then crowdsourcing will not provide desirable results. So far it seems that the level of engagement has been relatively low, but many data producers have actively sought to create new ways of engagement and feedback.³²⁹

Since the principal users of open data are – according to government policies – supposed to be citizens themselves, and taking into account the growing demand for a more user-perspective approach in information management, in the academic sense it sounds useful to further engage end-users in providing context and thus produce better quality data. According to several theories of crowdsourcing, it is the end-users who should have the best capabilities to contextualise information.³³⁰

“We are moving towards a world where data is seen less as owned by organisations and more by the community and individuals. There’s an argument that they’ve got a vested interest in improving the quality of data.”³³¹

Steve Bailey and Jay Vidyarthi argue that information management is specifically lacking in the depth and sophistication of its knowledge about the needs and

³²⁷ Interview, a local government civil servant

³²⁸ Interview, a local government civil servant

³²⁹ Interview, a local government civil servant

³³⁰ Boleman, Leo & Coleman, Emer: *New Development: What Lies Beyond Service Delivery? Leadership Behaviours for Place Shaping in Local Government*, Public Money and Management, Volume 28, Issue 5, 2008, pp. 313-318.

³³¹ Interview, a central government civil servant

objectives of information users.³³² In this regard, it is important to hear data users' views on open data and later build on those responses.

Since the whole open-data movement generally breathes through online collaboration and participation, it is a logical first step and in accordance with the arguments of Bailey and Vidyarthi to assess the question of data contextualisation by the wisdom of the crowds, that is attempt to create relevant metadata via collaborative crowdsourcing methods. Leo Boland and Emer Coleman even argue that public authorities should advocate raw datasets instead of contextualised data in government transparency.³³³

“Developers said don’t get into discussion about formats, just as long as it is not PDF’s.”³³⁴

However, the relevance of crowdsourcing as a tool for contextualisation relates to citizens' use of public data and whether they are interested in providing context to data in the first place. This is related to the concept of public participation. As is already mentioned, Tim Davies has identified the main motivations for data users, with some of the biggest clusters being those who engage with data because they are focused on creating new online platforms and improving government efficiency, thus being at least theoretically interested in collaborating in online metadata mapping projects.³³⁵ The problem with the status quo, however, is the fact that most data providers don't interact with data users and thus miss out their experiences of data.³³⁶ This connection is important to further examine the different user-driven scenarios of context provision.

Arguably, further engagement of data users in context provision could be helpful in order to further satisfy the demands of access, interpretation and use when utilising open data. The possible wider context of data is thus related to a combination of various types of metadata from different sources of linked open data, created by a

³³² Bailey, Steve & Vidyarthi, Jay: *Human-computer interaction: the missing piece of the records management puzzle*, Records Management Journal, Vol. 20, No. 3, 2010, pp. 279-290.

³³³ Boleman, Leo & Coleman, Emer: *New Development: What Lies Beyond Service Delivery? Leadership Behaviours for Place Shaping in Local Government*, Public Money and Management, Volume 28, Issue 5, 2008, pp. 313-318.

³³⁴ Interview, a local government civil servant

³³⁵ Davies, 2010

³³⁶ See, for instance: Worthy et al., 2011

myriad of data users. This does not intrinsically undermine the role of the organisation providing the data, since the initial context of the data derives from the actions of the organisation itself.

In order to comprehensively analyse the level of public understanding of data and therefore its information value, further research on data users is needed. However, the fundamental questions of users' equal capabilities to exploit data remain.

6.6 Data divide and lack of interest

“Perhaps the most stressing or serious problem, in a sense that deserves the most attention, is the idea that by releasing the data you are empowering the already powerful.”³³⁷

When local-government data producers were asked about their perceptions of the public interest towards expenditure data, the results were surprisingly clear, indicating a rather low number of data downloads.³³⁸ When there are certain costs involved in the opening process, it might prove to be a significant argument against open data if the interest remains as minimal as it has been so far. However, different evidence is also available, as the police website showing crime statistics visualised on a map³³⁹ crashed due to heavy visitor traffic immediately after its launch.

The crime-map and expenditure-data examples further illustrate the fact that data needs to have relevant context and value added in order to be understood by the general public. The role of intermediaries is significant. The data itself might be useless to people without certain technical skills and a vision of how to utilise the data. However, a whole new question is how these skills are divided in society. Citizens without these skills are dependent on skilled people who can process data and potentially exploit it commercially.

Pippa Norris's influential book on the digital divide raised concerns on the inequality in access to relevant digital tools and the Internet.³⁴⁰ A similar theme has occurred in

³³⁷ Interview, an academic/researcher

³³⁸ Halonen, 2011

³³⁹ <http://www.police.uk/>

³⁴⁰ Norris, Pippa: *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*,

the open-data discussion, since the threshold for using data is seen to be relatively high: besides access to online resources, potential users need to possess some sophisticated technical skills. It is argued that because quite a small minority of people have these skills, open data can even widen the knowledge gap between those who are already empowered and those who face a threat of exclusion from society.³⁴¹

“There are technical barriers and there are also very high knowledge barriers in using information. Data is just a bunch of numbers, and you have to know what they mean. These are not skills that everyone in the community has.”³⁴²

Despite certain concerns, it seems that the question of the data divide is categorically different from the original concept of the digital divide. As it seems that the most beneficial solutions to start with are those where data is used by those community groups and public service providers who are already empowered and engaged, it is questionable whether it is a feasible expectation that data should be used by all of the public in the first place.

In the initial phase, it may well be sufficient that those with certain public service aspirations in mind use the data, and with time and education the number of users would increase. If we accept this argument, then it would be wise to stop pretending that open data is something truly empowering that would transform the otherwise passive society into a dynamic hub of active citizens. MySociety and the Open Knowledge Foundation are famous examples of community groups that aim at lowering citizens' threshold for using data.

“We shouldn't expect that the average member of the public is going to be using this. If we are going to release data, we should realise that it is almost certainly going to be journalists or some organized civil-society groups who are going to be using this. That is significant, because if you're looking at the rhetoric that is coming from the current government, it's all about reducing the importance of intermediaries and making the bureaucracy responsible

Cambridge University Press, Cambridge, 2001

³⁴¹ Gurstein Michael, B.: *Are the Open Data Warriors Fighting for Robin Hood or the Sheriff? Some Reflections on OkCon2011 and the Emerging Data Divide*, <http://gurstein.wordpress.com/2011/07/03/are-the-open-data-warriors-fighting-for-robin-hood-or-the-sheriff-some-reflections-on-okcon-2011-and-the-emerging-data-divide/>

³⁴² Interview, an academic/researcher

directly to the public at large, so there is a sense in which the rhetoric is going in a very different direction than the likely reality.”³⁴³

These challenges are not tough enough to degrade the value of open data. It is important to provide education on data skills and to improve equality in accessing the Internet, but as the research shows that the level of users is still relatively low, the risk of a massive increase in the knowledge gap is arguably rather low at this time. In addition, many of the organised civil-society communities that use data have real public-benefit aspirations in mind and aim to increase knowledge of data and also to solve more immediate social problems with the help of data.³⁴⁴

The engagement of potential data users is thus crucial, but the ability to spot those who download datasets and re-use them might be difficult if sufficient software is not in place. As is the case with government IT in general, the level of software compatibility does not seem to be so high that it would be economically sustainable to install a method of measuring the level of user engagement and data use.

Data education and better engagement of data users could potentially be a cure for the data divide, as well. However, this seems to be an issue that has lacked attention so far. Discussion of open data has been mostly focused on technical issues and the potential economic impacts. The social implications of data now require much more research.

³⁴³ Interview, an academic/researcher

³⁴⁴ See, for instance: <http://www.mysociety.org>

7. Conclusion

The focus of this paper has been to gain a better understanding of the open-data movement, its goals and development and some of the impacts that have been achieved with open data. In terms of the applicability of data, the key issues that emerged during the research process can be listed as follows:

- The government decision to initially emphasise expenditure data was not ideal
- Instead of increasing the sheer number of data users, it is more important to improve the quality of engagement by further empowering those who are already engaged
- There has to be both a legislative push and personal support for public-sector data providers
- To argue for the democratic impacts of data, more research is needed

Open data clearly can be seen as an integral part of the wider emergence of the digital infrastructure, both within government and also in society at large. Governments across the post-industrialised world have developed digital platforms for providing public services and in order to gain efficiency benefits in terms of the level of governance. New models of citizen participation are also being discussed and facilitated.

The United Kingdom has taken a firm approach to open data. Data transparency has been emphasised by both the previous and current Prime Ministers, and the Cabinet Office has, in fact, gone far enough to declare open data as its single most important piece of public policy.³⁴⁵ The UK has effectively taken a highly institutionalised route to open-data policies with strong political support.

This institutional approach has had its pros and cons. On one hand, due to strong political support, open data has gained visibility and public authorities have been effectively forced to open up their datasets, despite certain resistance from some parts of the civil service. On the other hand, the open-data discussion has taken a fairly politicised tone at times, which – if continued – can lead to blurring of the

³⁴⁵ Tim Kelsey at FOI Live at UCL 23.6.2011. Video accessible at <http://www.ucl.ac.uk/constitution-unit/events/foilive-2011>

intrinsic objectives of open data. Open data does not operate in a political vacuum – the proactive release of datasets usually aims at achieving certain targets, which are set up by political processes. Issues like public spending and the level of crime have traditionally been really close to Conservative policies in the UK, and those were the exact issues that the initial release of datasets was hoped to shine light on.

The intentions are evidently good, but not all decisions have been ideal. The decision to emphasise expenditure data has not yet proven to be the right one. Countries that have not yet established their open-data policies should therefore think twice about what kind of data should be opened up first, if it is not possible to release all the public data at once. Data transparency is not just about public spending and salaries of senior civil servants, even though the tabloid press might be interested in exactly those issues. Public spending is a delicate matter, with highly ideological aspects, and there is a certain risk of twisting the debate. One reason for the initial emphasis on expenditure data has been a political one: the historical narrative of transparency in the UK suggests that the current government has always been keen to pinpoint the financial mishaps of the previous government, and the Conservative discourse on transparency has emphasised the issue of public-sector inefficiency and the reduce of waste. Another potential reason is the fact that expenditure data is something that every public organisation must have at hand, if they have operated in accordance with the law.

Many data producers in local government feel that the demand to publish all spending above £500 was not a fair one, and subsequently the motivation to open up further datasets has been low. Moreover, it is argued that expenditure data lacks the necessary context and therefore has no proper information value. This can lead to confusion and even misinformation on how public funds are being used and thus undermine the overall value of the whole open-data scheme. In London, the developer community was consulted before any concrete measures were taken, and the result of the consultation revealed that the most popular datasets were those concerning traffic and crime, instead of expenditure data. The Cabinet Office has also emphasised the importance of public engagement, and at this time it seems that the best solutions need comprehensive community co-operation and engagement with citizens and businesses in order to materialise.

The open-data community is getting bigger and bigger, and simultaneously open data, as an issue, is gaining interest from different parts of society. Governments face the risk of being excluded from the crucial development phase if they do not actively take part in the discussion and facilitate development by progressive data re-use policies.

Data hackers are extremely well connected and so eager to develop new applications and innovations based on data that they are likely to do it anyway, regardless of government support. Governments should therefore support their work and also create policies that support the applicability of data. There are plenty of examples where developers have decided to approach the grey area by scraping datasets from websites, despite the fact that it might be technically illegal to do so. Contemporary society is increasingly dependent on data, and citizens with data processing skills are essential in order to fully capitalise on all the potential that lies within public data.

It is crucial that we understand this profound change in the infrastructure. The 21st century world is increasingly digital, whether we like it or not. The Vice President of the European Commission has only very recently likened data to gold. However, more important than merely helping entrepreneurs and their use of data is to further educate people on the importance of data and information to our society. We can not afford to widen the gap between digital haves and have-nots. It is increasingly important to assess the social implications of open data.

The number of individual data users is still so small that, in terms of filtering the information to the wider public, media plays a crucial role, just as FOI does. So far, the British media has focused mostly on separate expenditure issues revealed by the transparency agenda. This has twisted the big picture on data transparency and is not the way to go in the future. If data transparency is to achieve its objectives of positive democratic impact, the public discussion must change from the current gotcha-culture, where single incidents are scrutinised at the expense of more comprehensive interpretation of societal issues. As pointed out earlier, the social implications of open data seem to have been neglected in research so far and need much more emphasis in the future.

Research suggests that the expectations of the positive social impacts of open data have so far been overly optimistic. The data users are a marginal group, and it is extremely difficult to prove that re-use of data has so far distilled significant social value into the wider society. However, in terms of the applicability of data, this paper has argued that, instead of the sheer number of data users or the type of data (health, crime, traffic, expenditure, etc.), a more important issue is to empower those who are already engaged in data use in some way and also integrate data users and developers efficiently into the policymaking process. It is highly questionable whether politicians and public officials themselves have the best knowledge of applying data for the public good, and therefore a design approach would potentially be desirable.

At this time, we should not hide any more behind the “it is too early to tell” argument, despite there being a certain level of truth in this. There are strong normative arguments for open data, and bold decisions have to be made if we wish to gain the expected benefits.

The ethnographic side of the study reveals that the open-data community is relatively unanimous on the objectives and benefits of open data. Open data seems to be most applicable in small-scale solutions that aim to make life easier, but openness in general reflects the whole society. A bigger impact could therefore be achieved in changing the mindsets of people, both public officials and individual citizens. Open data is not necessarily a change in itself but may facilitate the change, as has been the case with several open-data projects where the policymaking process has been opened up.

Lessons

This paper argues that without engaging data users in open-data policies effectively, little positive development is likely to be achieved in terms of applicability of data. Open data only has value in its use, and at the moment it seems that the most efficient way of facilitating that use is to:

- a) Provide more education on the data society
- b) Further engage those that are already empowered
- c) Encourage third-party sites and initiatives

d) Do not focus initially on any particular type of data but try to release it all

In the case of more education, the main purpose is not to make citizens more likely to build apps and websites and profit from them financially, but to understand the basics of the new digital world and be able to manage their own lives within it. Learning the basic programming skills and how to interact with computers is one of the key skills that the future generation should possess.

Given the current low interest in raw data, the role of third parties in disseminating information for the greater public is significant. Similar challenges are faced by every social initiative that relies on public participation. People have to earn a living, and it is a very small group of people who, in fact, are ready to give up their time to produce services by themselves. However, there are many organised civil-society communities that have the resources to create these services. Also, local authorities with enthusiastic officials have proved that open data can be applied in very efficient and beneficial ways. The key issue just seems to be some existing framework, around which initiatives can be built.

The applicability of data is effectively linked to the initial objectives of open data. The value of open data is built on an uncertain variable and on how people use it – it is difficult to form a single “one size fits all” model, where value of applicability is measured using single variables. Also, the very definition of open data is still relatively vague and needs more clarification, if we wish to avoid things being sold as open data when they clearly are not.

At this time, it would be useful to disaggregate the objectives of open data and measure the impact on each application area separately. Different datasets are beneficial for different sectors. A good starting point for this sector-specific approach would be to design a measurement model for the democratic impact of open data. Economic impacts can be measured relatively easily with the current methods, but the possible changes in our society due to digitisation of the core infrastructures and the abilities of citizens to manage their lives within it pose challenges for the legitimate and democratic transparency regime. Truly democratic transparency requires more than just the release of open data. It needs citizens who can see that their interests are treated equally in society. If it is hoped that open data will provide

the catalyst for this, then the barriers to access, use and interpretation of data need to be as low as possible and citizens must be able to request for data they truly want to have themselves.

8. References

Interviewees

I wish to thank each of the following people for kindly giving their time to be interviewed for this project. Without their contribution this project would not have been possible. All views presented in this report are their personal opinions and not of any organisations they might represent. When quoted, only a vague description of their job responsibility is used, thus protecting the anonymity of opinions.

In alphabetical order:

Charlotte Alldritt

Cabinet Office Transparency Team

Beth Brook

Information Policy Advisor, the National Archives

Diane Cabell

Acting director, iCommons Ltd

Emer Coleman

Director of Digital Projects, Greater London Authority

Tim Davies

Independent consultant/researcher & PhD student, University of Southampton

Clive Davis

Software Development Group Leader, London Borough of Redbridge

Lee Edwards

Chief Information & Communications Technology Services Officer, London Borough of Redbridge

Martin Ferguson
Policy Director, Socitm

Anne Fitzgerald
Professor, Law and Justice Research Centre, Queensland University of Technology

Roger Hampson
Chief Executive, London Borough of Redbridge

Janet Hughes
Head of Scrutiny, Greater London Authority

Matthew Kerr
Technical Transformation Officer, Kent County Council

Tom McClean
LSE Fellow in Political Sociology, London School of Economics

Matthew Pearce
Standards Advisor, the National Archives

Simon Rogers
Editor, Guardian Data Blog

Fabrizio Scrollini
PhD student, London School of Economics

Julian Tait
Founding Member of Open Data Manchester

Ben Worthy
Research Associate, The Constitution Unit, University College London

Events attended

FOI Live: Open data, FOI and the future, University College London, 23.6.2011

Open Knowledge Conference, Berlin, 30.6-1.11.2011

Open Data Meetup, London, 3.10.2011

Open Data and Cultural Heritage, London, 6.10.2011

Open Government Data Camp, Warsaw, 20.-21.10.2011

Open Knowledge Finland Meetup, Helsinki, 30.11.2011

Anonymity, Privacy and Open Data, London, 7.12.2011

UK government documents

APPSI: *Exploiting governments information assets for public good*,
<http://www.nationalarchives.gov.uk/documents/Exploiting-government-information-assets-for-public-good.pdf>

Cabinet Office: *Making Open Data Real: A Public Consultation*,
<http://www.cabinetoffice.gov.uk/resource-library/making-open-data-real-public-consultation>

Cabinet Office: *A Consultation on Data Policy for a Public Data Corporation*,
<http://www.cabinetoffice.gov.uk/sites/default/files/resources/A-Consultation-Data-Policy-for-a-Public-Data-Corporation.pdf>

Cabinet Office: *Further Detail on Open Data Measures in the Autumn Statement 2011*,
http://www.cabinetoffice.gov.uk/sites/default/files/resources/Further_detail_on_Open_Data_measures_in_the_Autumn_Statement_2011.pdf

Cabinet Office: *Open Public Services White Paper*, July 2011,
<http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf>

Cabinet Office: *Open Government White Paper, 1993*, <http://www.official-documents.gov.uk/document/cm22/2290/2290.pdf>

Home Office, Committee on the Civil Service, Her Majesty's Stationery Office, London 1968, <http://www.civilservant.org.uk/fultonreport.shtml>

Letter from APPSI to Minister Michael Wills MP, 21 July 2008,
<http://www.nationalarchives.gov.uk/documents/appsi-minister-letter-ec-submission-07-2008.pdf>

Letter from APPSI to Minister Lord McNally MP, 6 July 2010, <http://www.nationalarchives.gov.uk/documents/appsi-minister-letter-ec-submission-07-2010.pdf>

Local Public Data Panel Minutes, <http://data.gov.uk/blog/local-public-data-panel>

New Public Sector Transparency Board and Public Data Transparency Principles, <http://data.gov.uk/blog/new-public-sector-transparency-board-and-public-data-transparency-principles>

PM David Cameron's letter to Government departments, 31 May 2010, <http://www.number10.gov.uk/news/letter-to-government-departments-on-opening-up-data/>

PM David Cameron's letter to Cabinet Ministers on Transparency and Open Data, 7 July 2011, <http://www.number10.gov.uk/news/letter-to-cabinet-ministers-on-transparency-and-open-data/>

The Coalition: our programme for government, http://www.cabinetoffice.gov.uk/sites/default/files/resources/coalition_programme_for_government.pdf

The National Archives: *Open Government Licence for public sector information*, <http://www.nationalarchives.gov.uk/doc/open-government-licence/>

Transparency Board Minutes, <http://data.gov.uk/blogs/leadership>

UK Government Licence Framework, <http://www.nationalarchives.gov.uk/information-management/government-licensing/the-framework.htm>

UK Parliament Commons, Written Answers, <http://www.theyworkforyou.com>

EU Documents

European Directive 2003/98/EC of 17 November 2003 on the re-use of public sector information. Official Journal of the European Union 345/90, http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive/psi_directive_en.pdf

European Public Sector Information Platform, Topic Report 22: *Simplifying PSI re-use in the United Kingdom: the UK Government Licensing Framework and the Open Government Licence*

Speeches

The Chancellor's Autumn Statement 2011, http://cdn.hm-treasury.gov.uk/autumn_statement.pdf

Secretary of Communities and Local Government, Eric Pickles: A new era of transparency will bring about a revolution in town-hall openness and accountability, 4 June 2010, <http://www.communities.gov.uk/news/corporate/1606882>

Speech by the Chancellor of Exchequer, George Osborne, at Google Zeitgeist 16 May 2011, http://www.hm-treasury.gov.uk/press_48_11.htm

Speech by Tony Blair at the Campaign for Freedom of Information's annual Awards ceremony, 25 March 1996, <http://www.cfoi.org.uk/blairawards.html>

Kroes, Neelie: *Opening Remarks, Press Conference on Open Data Strategy*, Brussels, 12 December 2011, <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/872&type=HTML>

Prime Minister David Cameron, HOC Liaison Committee, evidence from the Prime Minister, 6 March 2012, <http://www.parliamentlive.tv/Main/Player.aspx?meetingId=10438>

Bibliography

2020 Public Services Trust: *Online or In-line: The future of information and communication technology in public services*, March 2010, http://clients.squareeye.net/uploads/2020/documents/online_or_inline.pdf

Ackoff, R.L.: *From Data to Wisdom*, Journal of Applied Systems Analysis, Volume 16, 1989, pp. 3-9.

Andersen, K. N. ; Henriksen H.; Medaglia, R.; Danziger, J.N.; Sannarnes, M.K.; Møyfrid K.; and Enemærke, M., 2010. 'Fads and Facts of E-Government: A Review of Impacts of e-Government (2003-2009)'. International Journal of Public Administration, Vol 33 (11) 564-579.

Armstrong, Cory L.: *Providing a clearer view: An examination of transparency on local government websites*, Government Information Quarterly, Volume 28, Issue 1, January 2011, pp. 11-16.

Bailey, Steve & Vidyarthi, Jay: *Human-computer interaction: the missing piece of the records management puzzle*, Records Management Journal, Vol. 20, No. 3, 2010, pp. 279-290.

Bannister, Frank & Connolly, Regina: *The Trouble with Transparency: A Critical Review of Openness in e-Government*, Policy & Internet, Volume 3, Issue 1, Article 8, 2011, <http://www.psocommons.org/policyandinternet/vol3/iss1/art8/>

Becker, Howard S. & Geer, Blanche: *Participant observation and interviewing*, in Seale, Clive (ed.): *Social Research Methods – a reader*, Routledge, New York, 2004.

Berners-Lee, Tim & Shadbolt, Nigel: *There's gold to be mined from all our data*, <http://eprints.ecs.soton.ac.uk/23090/1/Times%20OpEd%20TBL%2DNRS%20Final.pdf>

Bertot, John Carlo; Jaeger, Paul T.; and Grimes, Justin M.: *Using ICTs to create a culture of transparency: E-Government and social media as openness and anti-corruption tools for services*, *Government Information Quarterly* 27 (2010), p.264.

Boleman, Leo & Coleman, Emer: *New Development: What Lies Beyond Service Delivery? Leadership Behaviours for Place Shaping in Local Government*, *Public Money and Management*, Volume 28, Issue 5, 2008, pp. 313-318.

Campbell et al: *The American Voter*, University of Chicago Press, 1960.

Chapman, Richard; Hunt, Michael: *Open Government in a Theoretical and Practical Context*, Ashgate Publishing, Aldershot, 2006.

Christiano, Thomas: *The Authority of Democracy*, *The Journal of Political Philosophy*, Volume 12, No. 3, 2004, pp. 266-290.

Clarke, Paul: *There's data, and there's data*, <http://paulclarke.com/honestlyreal/2010/06/theres-data-and-theres-data/>

Coleman, Emer: *From New Public Management to Open Governance – a New Future Or The More Things Change The More They Stay The Same?* University of Warwick, 2011, <http://davepress.net/FromNPMtoOpenGovernance.pdf>

Colquhoun, Anna: *The Cost of Freedom of Information*, The Constitution Unit, December 2010, <http://www.ucl.ac.uk/constitution-unit/research/foi/countries/cost-of-foi.pdf>

Cummings, Kate & Findlay, Cassie: *Digital recordkeeping: are we at a tipping point?* *Records Management Journal*, Vol. 20, No. 3, 2010, pp. 265-278.

Dahl, Robert A.: *Democracy and its critics*, Yale University Press, New Haven, 1989.

Dahl, Robert A.: *On Democracy*, Yale University Press, New Haven, 1998.

Dalton, Russell: *Citizenship Norms and Political Participation in America: The Good News Is... the Bad News Is Wrong*, CDACS Occasional Paper 2006-01, October 2006, <http://www8.georgetown.edu/centers/cdacs/cid/DaltonOccasionalPaper.pdf>

Davies, Alysia & Lithwick, Dara: *Government 2.0 and Access to Information: Recent Developments in Proactive Disclosure and Open Data in the United States and Other Countries*, Library of Parliament of Canada, 15 April 2010, <http://www2.parl.gc.ca/Content/LOP/ResearchPublications/2010-15-e.pdf>

Davies, Tim: *Evaluating the Autumn Statement Open Data Measures*, <http://www.timdavies.org.uk/2011/12/02/3090/>

Davies, Tim: *Open data, democracy and public sector reform. A look at open government data use from data.gov.uk*, University of Oxford, August 2010, <http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf>

Deloitte: *Unlocking Growth – How open data creates new opportunities for the UK*, Deloitte 2011, http://www.deloitte.com/view/en_GB/uk/market-insights/deloitte-analytics/2d376cdde7e14310VgnVCM1000001a56f00aRCRD.htm

Denzin, Norman K. & Lincoln, Yvonna S (edit.): *Strategies of Qualitative Inquiry*, Sage Publications, Thousand Oaks, 1998.

Eaves, David: *Access to Information is Fatally Broken... You Just Don't Know it Yet*, <http://eaves.ca/2011/03/30/access-to-information-is-fatally-broken-you-just-dont-know-it-yet/>

Eaves, David: *Case study: How Open data saved Canada \$3.2 Billion*, <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>

Erkkilä, Tero: *Reinventing Nordic Openness: Transparency and State Information in Finland*, University of Helsinki, 2010.

Fioretti, Marco: *Open Data: Emerging trends, issues and best practices*, http://www.lem.sssup.it/WPLem/odos/odos_2.html

Gant, Jon; Turner-Lee, Nicol: *Government Transparency: Six Strategies for More Open and Participatory Government*, The Aspen Institute, February 2011, p.17, http://www.knightcomm.org/wp-content/uploads/2011/02/Government_Transparency_Six_Strategies.pdf

Gray, Jonathan: *Interview for University of Southampton open data study*, <http://jwyg.okfn.org/2011/12/06/interview-for-university-of-southampton-open-data-study/#comments>

Guardian Data Blog, <http://www.guardian.co.uk/data>

Feldman, Steven: *#OpenData – Who Profits?*, <http://govinthelab.com/steven-feldman-opendata-who-profits/>

Fung, Archon & Weil, David: *Open Government and Open Society*, in Lathrop, Daniel & Ruma, Laurel (ed.): *Open Government: Collaboration, Transparency, and Participation in Practice*, O'Reilly Media, Sebastopol, 2010.

Gurstein, Michael B.: *Are the Open Data Warriors Fighting for Robin Hood or the Sheriff? Some Reflections on OkCon2011 and the Emerging Data Divide*, <http://gurstein.wordpress.com/2011/07/03/are-the-open-data-warriors-fighting-for-robin-hood-or-the-sheriff-some-reflections-on-okcon-2011-and-the-emerging-data-divide/>

Habermas, Jürgen: *The Structural Transformation of the Public Sphere. An Inquiry into a Category of Bourgeois Society*, MIT Press, 2001.

Halonen, Antti: *Data without context? A study on the UK local government transparency agenda and its impacts on information management*, Unpublished MRes dissertation, University College London, 2011.

Hazell, Robert; Worthy, Ben; Glover, Mark: *The Impact of the Freedom of the Information Act on Central Government in the UK – Does FOI work?* Palgrave MacMillan, 2010.

Hey, Jonathan: *The Data, Information, Knowledge, Wisdom Chain: The Metaphorical Link*, published at Intergovernmental Oceanographic Commission-OceanTeacher: a training system for ocean data and information management, December 2004.

Holsen, Sarah et al.: *Journalists' use of the UK Freedom of Information Act*. Open Government Journal, Volume 3, Issue 1, 2007.

Huijboom, Noor & Van den Broek, Tijs: *Open data: an international comparison of strategies*, European Journal of ePractice, No. 12, March/April 2011, http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012_1.pdf

Inglehart, Ronald & Abramson, Paul R.: *Measuring Postmaterialism*, American Political Science Review, Vol. 93, No. 3, September 1999.

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

Kalela, Jorma: *Making History. The Historian and Uses of the Past*. Palgrave Macmillan, 2012

Korn, Naomi & Oppenheim, Charles: *Licensing Open Data: A Practical Guide*, accessible at http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf

Koski, Heli: *Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth?* The Research Institute of the Finnish Economy, September 2011, <http://www.etla.fi/eng/julkaisuhaku.php?type=details&id=1853>

La Raja, Raymond: *Does Transparency of Political Activity Have a Chilling Effect on Participation?* Meetings of the Midwest Political Science Association, Chicago, March 31 – April 3, 2011.

Lathrop, Daniel & Ruma, Laurel (ed.): *Open Government: Collaboration, Transparency, and Participation in Practice*, O'Reilly Media, Sebastopol, 2010.

Leadbitter, Charles: *The Civic Long Tail*, September 2011, <http://www.demos.co.uk/publications/theciviclontail>

Longo, Justin: *#OpenData: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse?*, Public Policy and Governance Review, Volume 2, No. 2, Spring 2011, pp. 38-52.

Mayo, Ed & Steinberg, Tom: *The Power of Information*, http://webarchive.nationalarchives.gov.uk/20100413152047/http://www.cabinetoffice.gov.uk/media/cabinetoffice/strategy/assets/power_information.pdf

McClean, Tom: *Not with a Bang but a Whimper: The Politics of Accountability and Open Data in the UK*, paper presented at the American Political Science Association Annual Meeting, Seattle, Washington, 1-4 September 2011, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1899790

McKinsey Global Institute: *Big data: The next frontier for innovation, competition and productivity*, McKinsey Global Institute, 2011, http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation

Moss, Michael: 'Without the Data, the Tools are Useless; Without the Software, the Data is Unmanageable', Journal of the Society of Archivists, Vol. 31, No. 1, April 2010, pp. 1-14.

Norris, Pippa: *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*, Cambridge University Press, Cambridge, 2001.

Noveck, Beth Simone: *The Single Point of Failure*, in Lathrop & Ruma (ed.): *Open Government: Collaboration, Transparency, and Participation in Practice*, O'Reilly Media, Sebastopol, 2010, pp. 49-70.

O'Hara, Kieron: *Transparent Government, Not Transparent Citizens: A Report on Privacy and Transparency for the Cabinet Office*, September 2011, <http://www.cabinetoffice.gov.uk/sites/default/files/resources/transparency-and-privacy-review-annex-b.pdf>

Patashnik, Eric: *Reforms at Risk: What Happens After Major Policy Changes Are Enacted*. Princeton University Press, 2008.

Pina, Vicente; Torres, Lourdes; Royo, Sonia : *Is e-government leading to more accountable and transparent local governments? An overall view*, Financial Accountability & Management, Volume 26, Issue 1, February 2010, pp. 3-20.

Poikola, Antti; Kola, Petri; Hintikka, Kari: *Public Data: an introduction to opening information resources*, <http://www.scribd.com/doc/57392397/Public-Data>

Pollock, Rufus: *Open Data: a means to an end, not an end in itself*, Open Knowledge Foundation Blog, 15 September 2011, <http://blog.okfn.org/2011/09/15/open-data-a-means-to-an-end-not-an-end-in-itself/>

Pollock, Rufus: *The Economics of Public Sector Information*, University of Cambridge, 2008, <http://econpapers.repec.org/paper/camcamdae/0920.htm>

Prior, Markus: *News vs. Entertainment: How Increasing Media Choice Widens Gap in Political Knowledge and Turnout*, American Journal of Political Science, Vol. 49, No. 3, July 2005.

Raymonds, Eric: *Cathedral and the bazaar*, <http://catb.org/~esr/writings/homesteading/cathedral-bazaar/cathedral-bazaar.ps>

Roberts, Alasdair: *Blacked out: Government secrecy in the information age*, Cambridge University Press, New York, 2006.

Rubin, Herbert J. & Rubin, Irene S.: *Qualitative Interviewing: The Art of Hearing Data*, Sage Publications Ltd, Thousand Oaks, 2005.

Saxby, Stephen: *The development of UK Government policy towards the commercialization of official information*, International Journal of Law and Information Technology, Issue 3, 1996, pp. 199-233.

Saxby, Steve: *Three years in the life of UK national information policy – the politics and process of policy development*, International Journal of Private Law, Vol. 4, 2011, pp. 1-31.

Schellong, Alexander & Stepanets, Ekaterina: *Unchartered waters – the State of Open Data in Europe*, CSC Public Sector Study Series 1/2011,

http://assets1.csc.com/de/downloads/CSC_policy_paper_series_01_2011_unchartered_waters_state_of_open_data_europe_English_2.pdf

Stephen, Justine et al.: *System Error – Fixing the flaws in government IT*, Institute for Government, 2011, <http://www.instituteforgovernment.org.uk/publications/23/>

Taggart, Chris: *What's that coming over the hill? Is it the Public Data Corporation?* <http://countculture.wordpress.com/2011/01/14/whats-that-coming-over-the-hill-is-it-the-public-data-corporation/>

UN e-Government Survey 2010, <http://unpan1.un.org/intradoc/groups/public/documents/UN-DPADM/UNPAN038853.pdf>

Vickery, Graham: *Review of recent studies on PSI re-use and related market developments*. Information Economics, Paris, 2011. http://www.epsiplatform.eu/news/news/review_of_recent_psi_re_use_studies_published

Vleugels, Roger: *Overview of all 86 FOI countries (2008)*, <http://www.statewatch.org/news/2008/sep/foi-overview-86-countries-sep-2008.pdf>

West, Darrell M.: *E-Government and the Transformation of Service Delivery and Citizen Attitudes*, Public Administration Review, Volume 64, Issue 1, February 2004, pp. 15-27.

Worthy, Ben: *John Major's Information Revolution? The Code of Access ten years on*, Open Government: a Journal of Freedom of Information, Volume 3, Issue 1, 2007, <http://www.opengovjournal.org/article/view/994/792>

Worthy, Ben et al.: *Town Hall Transparency? The Impact of Freedom of Information on Local Government in England*, The Constitution Unit, December 2011. <http://www.ucl.ac.uk/constitution-unit/research/foi/foi-and-local-government/town-hall-transparency.pdf>

Zaller, John: *The Nature and Origins of Mass Opinion*, Cambridge University Press, 1992.

Zeleny, M.: *Management Support Systems: Towards Integrated Knowledge Management*, Human Systems Management, Volume 7, 1987, pp. 59-70