Open government data: critical information management perspectives

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal:</td>
<td>Records Management Journal</td>
</tr>
<tr>
<td>Manuscript ID:</td>
<td>RMJ-08-2018-0023.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Research Article</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Open Government, Open Government Data, Information management,</td>
</tr>
<tr>
<td></td>
<td>Public Records Management, England</td>
</tr>
</tbody>
</table>
Open government data: critical information management perspectives

Abstract

Purpose

Open government data and access to public sector information is commonplace, yet little attention has focused on the essential roles and responsibilities in practice of the information and records management professionals, who enable public authorities to deliver open data to citizens. The article considers the perspectives of open government and information practitioners in England on the procedural and policy implications of open data, across local public authorities.

Design/methodology/approach

Using four case studies from different parts of the public sector in England (local government, higher education, National Health Service (NHS) and hospital trust), the research involved Masters level students in the data collection and analysis, alongside academics, thus enhancing the learning experience of students.

Findings

There was little consistency in the location of responsibility for open government data policy, the range of job roles involved, or the organizational structures, policy and guidance in place to deliver this function. While this may reflect the organizational differences and professional concerns, it makes it difficult to share best practice. Central government policy encourages public bodies to make their data available for re-use. However, local practice is very variable, and perhaps understandably responds more to local organizational strategic and resource priorities. A lack of common metadata standards for open data, different choices about which data to open, problems of data redundancy, inconsistency and data integrity and a wide variety of views on the corporate and public benefits of open data.

Research limitations/implications

The research is limited to England and to non-national public bodies and only draws data from a small number of case studies.

Originality/value

The research contributes to the debate about emerging issues around the complexities of open government data and its public benefits, contributing to the discussions around technology-enabled approaches to citizen engagement and governance. It offers new insights into the interaction between open data and public policy objectives, drawing on the experience of local public sectors in England.

Funding statement: This article presents the research findings from Trust in Digital Records (InterPARES Trust), supported by the Canadian Social Sciences and Humanities Research Council (SSHRC), grant number 895-2013-1004. [...] is one of
the European partners in the InterPARES Trust (IPT), a multi-national, interdisciplinary research project (http://www.interparestrust.org/).

Keywords: Open Government, Open Government Data, Information Management, Public Records Management, England
Open government data: critical information management perspectives

Introduction

Open government data and access to public sector information in the UK and across the world is in transformation and yet little attention has focused on the essential roles and responsibilities in practice of the information and records management professionals, who enable public authorities to deliver open government data to citizens. This article presents the research findings from [...] one of the European partners in the InterPARES Trust (IPT), a multi-national, interdisciplinary research project exploring digital records (http://www.interparestrust.org/). The article considers the perspectives of open government and information practitioners working in localities in England, set into an international academic research frame. It draws on practice experience across different types of local public authorities of the procedural and policy implications of open government data. It seeks to surface the significant but often-overlooked links between the effective management of information and the delivery of open government data.

Since 2013, [...] has run several linked projects that studied the role of the information and records management discipline in the context of new obligations on public sector bodies towards open government data (ie the proactive release of data by public sector organizations for re-use by third parties in the public benefit) and greater access to data for citizens. The research aims to develop a picture of implementation and compliance in the field outside central (national) government, using four case studies from different parts of the public sector in England. The lens through which it studied open government data was the professional discipline of information and records management, since these professionals have a critical but
under-recognised and hence often under-resourced role in the practical operation of open government data. It frames the findings from practice with a literature review drawing on academic research into the intersection between open government data and information management internationally. It concludes with implications for policy and practice.

In the policy context, public authorities are responding to the European General Data Protection Regulation (GDPR) which further constrains the requirements for managing and sharing personal data. Regulatory agencies including the UK Information Commissioner’s Office (ICO, 2017b) propose additional administrative data duties, such as the ‘duty to document’. In addition, citizens demand greater accountability and transparency in public processes. In this shifting climate, it is critical that open government data policy is better understood and framed holistically. Although related to freedom of information, access to public information and records management, as suggested by Janssen et al (2012) and Zuiderwijk et al (2014), open government data has distinct characteristics that merit investigation. For example, the relationships between open government data, public records, freedom of information, linked data and the role of information managers and data scientists are complex, overlapping and not fully understood, as set out by Luna-Reyes et al (2014), and Shepherd (2015, 2017). As government functions are increasingly delivered by commercial and third sector bodies in partnership with the public sector, issues of data provenance, guarantees of data standards and ownership need exploration. The secondary use of open government data is a significant resource to policy makers and for academic researchers, especially if datasets can be linked, as Safarov et al (2017), and Sexton et al (2017), explore.

Research methods
The foundation of the research was a literature review in two phases, on open
government data broadly and then focusing more closely on the related privacy
issues. The primary data about open government data in practice was collected in
four linked case studies, over a period of three years 2014-2016, in four different
local public sector settings: a local government authority, NHS England, a hospital
trust, and a university.

**Literature Review**

The literature review identified key literature about open government data and the
recordkeeper’s role. The second stage of the literature review focused more
specifically on the implications for information management of moves to extend open
government data and access to public sector information, including medical records
and patient data, in the context of data protection and privacy regulations in the UK
and EU. It analysed legislation, policy documents and research literature on the re-
use of public sector information, data protection, and use of medical records and
patient data. The focus was predominately England, but also looked outward to
legislation and policy in the devolved nations of the UK, as well as internationally.

**Case Studies**

The research method adopted was a qualitative, instrumental case study, extended
to four cases, each with some similarities and some differences, following Stake’s
(1994) typology. Stake identifies three main types of case studies (intrinsic,
instrumental, collective) although all have the characteristics of setting out the
historical and contextual background, information about the respondents and the
nature of the case. Through the collective case study, the researchers aimed to
provide insight into an overall research aim: to investigate the impact of open
government data requirements in local public authorities in England, through the lens of the practice of information and records management. Each case studied how open data were managed within one public authority, explored the reality of practice, and where the roles and responsibilities lie. These case studies enable comparative work, analysing the open government data policy setting, and how it interacts with information management roles. The research explored the role of the information and records manager in practice in an open government environment, given that it is a critical actor in the delivery of open government data to citizens (Thurston, 2012).

The geographical scope was England, bounded by the specific legislative and government policy requirements. First, a local government authority in the South East of England responsible for providing services to 1.5 million residents, directly or with commercial and third sector bodies, including transportation, schooling, social care, housing, environment, planning, libraries and culture. Secondly, NHS England which leads the NHS, nationally, regionally and locally, by setting strategy and priorities and implementing policies, distributing £100 billion annually, commissioning contracts for health services, and contributing to public debates on health care.

Given the unique size and role of this body it was not possible to anonymise this organization. Thirdly, an NHS hospital trust which delivers clinical services to a large urban population, as in-patients and in the community, from several hospital sites. This case is published by Chorley (2017). The fourth case study was a multi-faculty university employing about 1800 staff, undertaking research, higher education teaching of 20,000 students and knowledge exchange, funded by public and private funds. These four organizational settings had significant differences from each other in terms of size (number of employees, budget), organizational priorities and culture.
leadership, government policy context for core service functions, and client groups.

See Table 1 for brief details.

[insert Table 1]

In each case, the specific research objectives were (1) to identify the governing legislative and regulatory open government and open data frameworks; (2) to investigate existing organisational practices and job roles in delivering open government data and complying with public obligations; and (3) to develop better understanding of the critical information management issues, policies and guidance, relating to open government data. The research process was the same for each case. UCL research ethical approval and data protection registration was obtained for the IPT project overall, 2013-2018. Consent required the anonymisation of interview participants and the organizations, except for NHS England, retaining contextual data, including job titles and organizational type.

Data Collection

Research data was collected through a series of fifteen individual semi-structured interviews (40-60 minutes in duration) with information, records, data, governance and other professionals. The interviews were organised on a series of common themes drawn from the literature and designed to explore the research objectives.

See Table 2 for example questions. Data collection focused around guidance and regulations, implementation approaches, roles and responsibilities, recordkeeping and users of open government data. We did not specifically seek data about individual datasets, but rather the policy and operational context in which they might be made open. Interviews were usually audio recorded and then selectively transcribed.
Documentation supplied by interviewees, including internal policies and procedures, organization charts and manuals, was also analysed. Data coding and analysis was carried out using the pre-coordinate themes which framed the data collection (ie. role, responsibilities and job context; governing regulations, institutional policies and guidance for information governance, open data, information and records management, freedom of information and data protection; institutional practices for open government data, information and records management, staff awareness and training; future developments), extended by the findings of the data collection. An initial summary report of data was shared on the IPT project intranet (https://interparestrust.org/).

Literature Review

The literature review framed the research by drawing on academic research into open government data. The literature also helped to establish the themes which were explored in the data collection.

Novais et al (2013) provided a literature review of open government data studies from 2007 to 2012, in which they indicated the need to broaden the geographic coverage of open government data research and to improve the quality criteria for assessing open government data, including to ensure the trustworthiness of data (Ceolin et al, 2014; Léveillé and Timms, 2015). Novais et al suggested that the term, open government data, began to appear in 2007, following the publication of the original eight open data principles, which stated that open data should be complete, primary not aggregated, timely, accessible, machine-processable, non-discriminatory, non-proprietary and licence-free.
These principles framed much of the subsequent discussion but many technical, managerial, organisational and cultural barriers remain in delivering open data. Kitchin (2014) surveyed the data landscape, including big and open data, and in Chapter 3, considered the consequences and difficulties of moving from traditional, closed, access to data to more open data access. He provided a clear historical account of the open data movement, discussed further below.

Work on models and methods for evaluating open government data and benchmarking, includes Kalampokis et al’s (2011) stage model of open government data and the Open Data Institute (2015). Kalampokis et al built on models proposed to assess the progress of eGovernment, to incorporate open government data more fully, by outlining two dimensions (organizational & technological complexity and added value for data consumers) and four stages (Aggregation of Government Data, Integration of Government Data, Integration of Government Data with Non-Government Formal Data and, Integration of Government and Non-Government Data with Social Data). They hope that the model will should improve benchmarking and the construction of roadmaps for open government data. The information and records management community began to engage with open government data in early 2010s, encouraged by the development of records management guidance in the Open Government Guide (2015), an online resource developed by international civil society organisations to support governments in developing commitments for Open Government Partnership (OGP, http://www.opengovpartnership.org/), national action plans and by growing professional awareness of the importance of the management of the underlying data to the effective delivery of open government data to citizens. The UK and USA governments were among the eight founding
countries of the OGP in 2011 committed to fostering ‘a global culture of open
government that empowers and delivers for citizens, and advances the ideals of
open and participatory 21st century government’, based on access to information,
citizen engagement, fiscal transparency, and income and asset disclosure (Herrero,
2015). The influence of OGP action plans on practice and on research was
considerable. For example, the UK government (2016) national action plan for 2016-
2018 developed in collaboration with civil society organisations included
commitments to engage with data users about their needs and to gain their views on
priority areas for the development of the open data agenda, and to develop
<gov.uk>, the website for government publications, to make it more open and
accountable.

Privacy and data protection are often seen as problematic in an open data context,
inhibiting the public good in sharing data. Whilst sharing data may lie at the heart of
the open government data agenda, from a privacy perspective the right to limit the
sharing of identifiable personal information is cast as a fundamental human right. A
number of overlapping legal measures exist to protect privacy, including privacy
rights, which guarantee freedom from interference; data protection, which controls
the processing of personal data; and duties of confidentiality, which protect against
unauthorised or unreasonable breaches of confidence (Nuffield Council, 2015).

There is recognition that personal privacy is not always in the public interest if it
impedes other fundamental human rights and interests. For example, there is a
public good in the use of medical records and patient data to support advances in
medical, health and scientific research. However, there is also a public good in
respecting and protecting privacy, maintaining confidentiality, and limiting the use of
medical records and identifiable patient data. Balancing these public goods (the
public good in enabling research and the public good in protecting data) is increasingly challenging given the rapidly evolving mechanisms open to researchers and others to re-use and link data (Caldicott, 2013).

**Research Results and Discussion**

This section presents and discusses the main practice findings, illustrated by quotations from the interviews, referenced by case study and interview (eg 1:1) as in Table 1.

**Who is responsible for open government data in public authorities?**

Our first question sought to understand where responsibility for open government data sat organizationally and the professional groups and units that led on the issue. We were particularly looking for evidence that records and information managers were explicitly involved and their information skills acknowledged as relevant to open data. As might be expected, given that the four case study organizations had very different core mandates and significant differences in size and structure, there was little consistency in the location of responsibility for open government data policy, the range of job roles involved, or the organizational structures in place to deliver this function. That each case study took such radically different approaches to delivering the same open data function was perhaps unexpected. As this was a fairly new area of work, the local authority had established an Open Data Working Group in its Business Intelligence Team to develop policy, including governance and law, information governance, information access and records management. An operational team for Information Transparency had responsibility for data protection, freedom of information, data sharing, data security and environmental information,
but open data was not yet part of its functions. The authority sought to ‘make the
data we are creating as an authority, as part of our daily business, available for more
general use’ (interview 1:3) and equated open government data with
‘transparency…being able to give data to the public when they ask for it and being
open about it…being open about what we spend our money on’ (interview 1:1). This
authority seemed to have well established teams to deliver existing information
functions, and the cross-functional Working Group was making progress in defining
and identifying the policy aspects of open data as a preliminary step towards making
the new function part of an operational team.

In the NHS hospital trust, responsibility for open data was not yet established and no
job titles referenced that role directly. As in many health settings, the overarching
function was information governance, following the Caldicott Principles (2013) for
handling patient information across the NHS (2016). A traditional line management
structure for information governance sat within the directorate of corporate affairs.
Information governance and internal corporate records functions were fairly
traditionally conceived as having a primarily corporate focus, rather than external
obligations for openness and transparency, although with responsibility for freedom
of information requests. One interviewee anticipated that when the hospital trust
formally adopted an open data function, there would be a separation between policy
and operations:

I would expect to see it coming to Corporate Affairs and the stuff that’s around
releasing data, IG and FOI would probably come our way, but some of the
stuff probably more about policies, about how the Trust runs itself would
probably sit with the Trust Secretary. (3:1)
But that in practical terms existing staff would ‘just do it as we go along’ (3:1).

Another interviewee thought that open government and open data were not yet high priorities for the hospital trust (3:2). This position is somewhat at odds with the picture given by NHS England, guided by its obligations under the Health and Social Care Act 2012. At the time of the case study in 2015, NHS England had a policy unit to develop policies around information standards, open data and patient care data which sought to be a centre of expertise, acting as a ‘think-tank for NHS England’ and learning actively ‘from the best across the UK and internationally’ (2:3). NHS Digital collects, manages and publishes health and care data, under the direction of NHS England and the Department of Health (2:4). Our case study hospital trust noted that it reported large amounts of data to NHS England: much of the hospital’s open data is effectively published by a third party. However, the responsibility for open government data is not always completely clear in large and complex public services. Sometimes it is not clear how variants of datasets relate to each other, or who should publish which ones.

In the university case study, there was evidence of a very active approach to business intelligence, data assurance and responsive provision of personal data and corporate information, partly driven by the requirement to provide data to the Higher Education Statistics Authority (HESA) as a condition of grant funding. One interviewee from the strategy and planning unit admitted, however, ‘I’m not that familiar with the open data stuff’, which she saw as an external issue about access to data for academic research, rather than a corporate issue (4:1). The university records manager’s role had expanded rapidly from managing corporate records and information to include information assets, audit and compliance, and information governance. In order to deal with this changing environment,
We set up an Information Governance Group, … more involved in information security and information assurance. What we’ve had to do is to bring together representation from across the HEI including security, IT, academics, our Research Information Manager who handles research data management (4:2).

Our case studies suggested that proactive individuals could make progress in encouraging a public authority to develop policy and practice in open government data by bringing a group of interested officers together in a focused, short-term way, such as on a policy working group, with the longer-term aim of embedding the new function in an existing unit. The corporate environment could help or hinder such initiatives.

No one professional group routinely took on open government data functions: it depended more on the interests of individual staff taking advantage of corporate opportunities rather than yet forming part of formal role descriptions. Directors of Governance and Law, Information Governance specialists, Information Access officers, Data Sharing specialists, Corporate Communications officers, Business Intelligence officers, Heads of Data Policy, Freedom of Information and Data Protection Officers and Corporate Information and Records Managers. Post holders had backgrounds in law, ICT, information and cyber security, informatics, education, social sciences, media and communications, as well as records management. The research reflected a shift in some organizations to information governance as the overarching corporate function for open government data. Information and records managers often have the skills needed to manage and deliver open data, but in practice, their employing organisations did not always exploit these skills for public benefit.
What policy and guidance on open government data exists in public authorities?

Part of the explanation for the variation of approaches to the management of open government data discussed in the previous section might be found in the varied policy and guidance context from central government or national bodies varied from one part of the local public sector to another. All of our case studies were subject to FOIA 2000, and the Data Protection Act (DPA) 1998, replaced by the GDPR and DPA in 2018. Although there is no open government data legislation, there is central government policy directed at public bodies encouraging them to make their data available for re-use, including the Open Data White Paper and Open Public Services White Paper (UK government, 2011, 2012). The Open Government Partnership promotes the open government agenda through action plans: the UK issued its third action plan including open data commitments in 2016. The requirement to publish certain data to fulfil the government’s transparency agenda at <https://data.gov.uk/> facilitates central agencies in publishing data.

However, considerable local and sectoral variation in policy can be found in our four case studies. Some sectors are much more highly regulated, in particular in two of our case studies in the health sector. NHS England, including NHS Digital (2017), has a responsibility for improving the quality of health and care data and publishes a number of national guidance and policy documents. NHS Records Management Code of Practice for Health and Social Care (2006) forms part of NHS Digital’s guidance and policies on information governance, including the Information Governance Toolkit (2016). Several of our interviewees commented on the lack of connection between records management and information governance in the NHS, even though these policies ought to support each other. Records management was described as ‘not visible or vocal enough to be involved in open government data’
Information governance, by contrast, was well developed and although the IG Toolkit was ‘not directly mandated by legislation but it is…done without question’ (2:1). ‘The idea of patient confidentiality is one that has been at the heart of doctor-patient relationship since time immemorial’ (2:4) and data confidentiality is integral to information governance in the NHS. In the NHS ‘the more open you get, the more nervous people get’ (2:3) and ‘already there are large numbers of people who do not want their identifiable data used for anything other than their direct care’ (2:4), not least in the wake of failed data sharing initiatives such as care.data (Presser et al., 2015; McLeod and Childs, 2018), and concerns over data sharing with commercial companies such as Google Deepmind (ICO, 2017a). Much of the data held in the NHS is patient data and identifiable personal data cannot ever be open: such data will only be published if ‘aggregated and anonymised’ (2:4).

The national open data and transparency agendas of NHS England are not always reflected at a local level. Barriers to proactive publication of data, as one interviewee remarked, ‘the culture is more around publishing as little as possible’, ‘partly the culture, partly a lack of awareness and also partly …the technical capability’ (3:2). For policy areas not governed by legislation, such as open government data, there was a considerable delay between policymaking at national level, and the filtering down of policy and associated practice to local level.

By contrast, the local authority case study approach to open government data was longer established, although in practice only a few specific datasets were routinely made open. The Local Government Transparency Code (2015), together with guidance from the Local Government Association, provided a framework which was overseen locally by elected Councillors. One interviewee (1:2) commented on the tension between, on the one hand, central government policy to publish as much
data as possible and rules requiring publication of certain local financial data with, on
the other hand, allowing the authority to decide which data was of interest to local
public audiences and prioritise their resources to select and process that data for
publication.

Problems with delivering open government data in practice were also identified,
especially an assumption that data and information are held in well-established
electronic systems. Local authorities were expected to meet at least level three of
the ‘5 Star’ scheme when publishing data (a simple indicator of the ‘openness’ of a
dataset, such as its format and metadata, see <http://5stardata.info/>). In our case
study, the local authority records management systems were still predominately
based around paper-based records, while emails, HR and financial data were held in
separate digital systems. Many different formats were used for the creation and
business use of datasets, yet most of the data required re-formatting, additional
metadata and re-presentation to make it openly accessible. In order for the authority
to meet the requirements for opening up data, as well as assure data reliability and
quality of the data drawn from records systems, major changes in the way data and
records were managed and accessed might be needed.

While open government data remains unlegislated, it is largely at the discretion of
individual organizations whether they proactively publish datasets, what they publish,
and indeed for what purpose. Our university case study interviewees were mindful of
its statutory obligations to provide data to HESA and to comply with data protection
and freedom of information, but to go beyond that to ‘publish large swathes of data’
which might have reputational risks and would cost scarce resources was not a
current priority. Since opening up data consumes resources, the corporate, individual
and public benefits needed to be balanced.
If more data is to be made open, then choices have to be made by the creating organization about which data to open. At present, these choices are not themselves transparent, nor are the reasons for them explained in public. Open data is generally not contextualised and users just see individual datasets: a proper understanding of open data might also require knowledge of which data was not opened.

Open government data as a policy-driven activity is subject to localised pressures. Resources, technical capabilities and strategic and local policy priorities tended to drive the choices about which data to open. Secondary data use is therefore highly dependent on local pathways and citizens and researchers are faced with highly variable open datasets from place to place, making re-use problematic and unpredictable. Our case studies suggested that consistent policies are lacking between different sectors, different organizations and indeed between different teams and units within one organization, leading to inconsistency and a lack of transparency. While a single common mandated approach would not necessarily serve the needs of diverse organizations and their clients, such a wide range and the lack of even a common set of principles leads to confusion about open data priorities. Under UK FOIA (2000), The National Archives (2009) published a Code of Practice on Records Management -providing core common requirements to support freedom of information, which individual organizations could adopt in a way suited to their own structure and mandate: no similar guidance exists for best practice in open government data provision.

Publishing data is not enough: what else is required?

As Dawes (2010) stated, ‘publishing data is not enough’, it needs to exist within a technical infrastructure and information context that renders it meaningful and usable
to the public. Our interviewees reported that since data was not created with public access in mind, a considerable amount of work was needed to ensure that data was re-usable, including perhaps requiring staff to change their working practices to use open data formats for ‘business as usual’ (1:2). They were also concerned about problems of data redundancy, inconsistency and data integrity. One records manager (1:3) explained her responsibility for guaranteeing protective markings on records containing confidential or personal data, and for ensuring the protection of personal data and privacy of individuals before data is made open. The Open Data Policy in this case study gave heads of service the responsibility to ensure ‘that data published is as accurate and usable as possible’, but the practicalities of implementing data checking processes systematically across the authority to ensure this were considerable, according to our interviewees (1:2, 1:4). Another interviewee (3:1) said that there was a need to rebut ‘the misconception that publishing data would be easy and straightforward’ and commented ‘I’m sure the politicians do just think it’s the click of a button, but it’s not’.

Complete and accurate metadata sets are needed to establish the connection of the dataset to its creating context and to assist with interpretation, but available metadata may be limited to that captured automatically, perhaps with some additional contextual information provided by data creators, according to one interviewee (3:2). Record creators and information governance staff were concerned with the accuracy of metadata, but the presentation and ease of use of the data and explanation of the metadata may be more important to the end user (1:2). A lack of common metadata standards for open government data and lack of standardisation of terminology hindered the publication and use of open government data, according to one interviewee (1:3).
Denis and Goëta (2014) in their ethnography of open government data projects in French administrations noted three significant operations (exploration, extraction and rawification) needed in order to prepare open data for release. In our study, the quality of the data and the amount of work needed to make it open also affected access: cleaner data, for example without any personal data or in simpler formats which are easier to reprocess, were more likely to be made open, simply because it was easier to deal with. ‘Trying to work out what can be published’, as well as the technical and resource problems, was complex. Even once a decision was made to release data, ‘when we say it’s going to be open’, is that ‘just going to be for those researchers who are asking for it’ or something ‘that we’re going to stick on our website that anyone can come onto and look at’ (3:1)?

What are the information management challenges of the open government data environment?

The final issue which we sought to examine in our case studies was to draw out some of the information management challenges of open government data and the proactive release of government data, to consider how information managers might help their parent bodies to meet them. Our case study organizations trying to improve the proactive release of government data faced many information and records management challenges, including the management of risk, privacy and data security. Protecting personal data, especially sensitive personal data, from inadvertent disclosure is important and risky if not done well. Cleaning up data before proactive release is time consuming but essential to safeguarding data quality and privacy.
If data is aggregated, linked, anonymised and released openly, data inaccuracies may be impossible to identify and correct and will affect the understanding of the data. Inaccuracies in the data can occur at many stages in the record creating process, through inadvertent or deliberate mistakes in recording, through inaccurate transcription and faulty re-processing. This risk can be mitigated by reliable internal procedures, including staff training, and well-established information management policies and procedures that help to ensure proper data management from creation to public access. The authors would argue that accurate data is essential for internal corporate uses of data, as much as for open data release.

Technical capabilities of staff and of the institutional infrastructures were identified as issues by the interviewees. For instance, the large quantities of data and large size of datasets to be delivered through websites posed technical challenges, leading to websites freezing or collapsing and disruption to other digital services. Many local public sector organizations do not have the resources to establish sufficiently robust and scalable web services to deliver open government data. So far, shared open data portals outside national government have not emerged in England. Universities could develop shared open data sources for administrative and research data on the shared national models seen in the past to develop computing infrastructure (JANET network) and some research data (eg UK Data Archive and Administrative Data Research Network) but at present, most open data development is localised.

Metadata standards for open government data are currently lacking in principle and in practice. Metadata provision is not resource-neutral as most data is coded and described for its original business purpose, rather than for re-use. At present the processing of data prior to opening is labour-intensive and, until technology-assisted routine processes are developed, is likely to remain so. The push to open large
quantities of data has to be balanced with the need to ensure data and metadata quality and with effective use of scarce resources.

Many of the skills needed to deal with open government data are similar to those needed to manage access to information, and many of the principles for managing open data mirror existing information and records management principles and standardized processes, such as provenance, functional classification, and creation of standard metadata and description. Information managers are, therefore, well placed to play a leading role in open government data. However, in order to fulfil these roles and responsibilities effectively and support the development of re-useable and reliable open government data, they need to develop greater awareness of open data, better technical skills in digital data and systems, improved expertise in information security, in routines for anonymisation for personal data, in data analytics, in digital curation, and in semantic web technologies. They need to deploy their skills in the assessment and management of risk, environmental and privacy impact assessment. Ensuring that information and records management skills are fit-for-purpose and used for the public good in an open government data setting is a critical challenge.

**Conclusion**

This study sought to investigate the impact of the open government data environment in local public authorities in England, through the lens of the professional practice of information and records management. In particular, it aimed to identify key legislative and regulatory open government data frameworks and principles, through a study of the literature, and to investigate existing organisational practices and job roles in delivering open government data through four case
studies. The research results showed that in practice in the public authorities in the study, there was little consistency in the location of responsibility for open government data policy, nor agreement about which professional group should lead the activity, nor common approaches for the practical delivery of open data. This finding is perhaps not surprising given the different governance models adopted by the case study organizations and their different mandates. As a fairly new area of work, many different professions were involved: governance, law, information governance and records management, ICT, cyber-security and data management. This highlighted some of the managerial and organisational barriers which public authorities currently face. Although a single common approach would not be appropriate for such a variety of public organizations, the lack of consistency inhibits effective sharing of best practice. If public authorities shared good practice and benchmarked across different domains and between different professional groups, then the success stories could better inform authorities whose open government data policy and practice was not yet fully developed.

A number of other barriers to open government data were identified. The technical infrastructure needed to deliver open government data, website capabilities, storage of data, metadata, and the long-term provision of access still need development. This was a resource issue, both the costs of implementing improved technical systems, but also in terms of ensuring the necessary staff competencies, re-training and re-skilling staff on an ongoing basis. Other technical issues related to the internal systems used to create data and records, which are not yet attuned to open data as a business-as-usual issue. As a result, data needs extensive re-processing before opening.
Cultural barriers still exist in organisations. Authorities did not agree about the benefits of open government data, some saw it as an exercise in transparency and accountability, others a means of ensuring internal data compliance and governance. Few seemed yet to have considered the external needs of open data users, and this is an area in which further research would be beneficial to guide policy. If public authorities do not see the value to their business of making government data open, then they will not allocate the necessary resources to ensuring that it happens.

This research suggests that the skills of information and records management professionals could be better utilised to help public authorities to meet open government data challenges. The management of risk, balancing privacy and the public good in open data, against the reputational and individual risks of releasing data inappropriately, is a critical skill. Ensuring data accuracy by enabling datasets to be traceable to their original reliable data source and ensuring the data governance systems which surround them is also critical. Information and records managers are well placed to play a leading role in open government data, but they need to explain their unique contribution more clearly in a crowded and confusing field.

This work makes three key recommendations:

• Open government data policy and structures should be supported by best practice case studies and guidance. As public bodies implement open government data functions they would benefit from the shared experience and models from best practice.

• Designing information systems with the potential to support open government data could be improved if public bodies co-developed shared templates for systems requirements, including open government data metadata standards and standards.
for the creation of current data. At present, developments seem to be highly
localised.

- The job roles and organisational structures for delivering open government
data initiatives should be clearly articulated and take account of the full range of
information expertise including, law, ICT, cyber-security, information governance,
and information and records management.

The development of a code of best practice for the management of open government
data across the public sector would enable these recommendations to be taken
forward.

Funding: This work was supported by the Canadian Social Sciences and Humanities
Research Council (SSHRC), Trust in Digital Records in an Increasingly Networked
Society (InterPARES Trust), grant number 895-2013-1004.

References

system. Information: To share or not to share? The Information Governance
Review”, available at:  https://www.gov.uk/government/publications/the-information-

Ceolin, D, Moreau, L, O'Hara, K, Fokkink, W, van Hage, W R, Maccatrozzo, V,
the Reliability of Open Government Data”, in Laurent, A, O Strauss, B Bouchon-
Meunier, and R R. Yager (eds.) Information Processing and Management of


29 no. 4, pp. 258–268.

Stage Model”, in Janssen, M, H J. Scholl, M A. Wimmer, and Y Tan (eds.) Electronic
22878-0_20

Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures &
their consequences. London: SAGE Publications. doi: 10.4135/9781473909472

a Framework for Trust in Open Government and Open Government Information.” in
Les Objectifs Visés Par Les Systèmes de Gestion Documentaires: La Mise En Place
D'un Cadre de Confiance et de La Transparence de L'information Dans Un


in the Cloud: Challenges and Solutions, Presentations at the 21st Archival Sciences
Colloquium of the Marburg Archives School, pp. 43-64.


UK government. *Open Data White Paper: unleashing the potential*, Cm 8353, 2012, 
available at: 


### Table 1: overview of case studies

<table>
<thead>
<tr>
<th>Four InterPARES Trust case studies</th>
<th>Date of data collection</th>
<th>Research Assistant</th>
<th>Number of interviews/ job roles of interviewees</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Local government</td>
<td>June-July 2014</td>
<td>Jessica Page</td>
<td>4 interviewees: 1. Information Access Officer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Business Intelligence Officer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Records Manager</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Information Governance Officer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Archives</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Senior Data Sharing Specialist</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Head of Data Policy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Senior Information Governance Advisor</td>
</tr>
<tr>
<td>3. Hospital trust</td>
<td>June-July 2016</td>
<td>Katherine Chorley</td>
<td>3 interviewees: 1. Information Governance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Manager</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Corporate Records Manager</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Assistant Records Manager</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Records Manager</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Freedom of Information and Data Protection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Officer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Research Information Officer</td>
</tr>
</tbody>
</table>
Table 2: summary of semi-structured interview questions

<table>
<thead>
<tr>
<th>Interviewee’s role</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Please explain your role and the responsibilities of your department?</td>
<td></td>
</tr>
<tr>
<td>Guidance and Regulations</td>
<td></td>
</tr>
<tr>
<td>1. What OGD obligations and regulations are imposed by central government?</td>
<td></td>
</tr>
<tr>
<td>2. What OGD guidance is available?</td>
<td></td>
</tr>
<tr>
<td>3. What internal and external policies are used for publishing OGD?</td>
<td></td>
</tr>
<tr>
<td>4. Is 5-Star Scheme used?</td>
<td></td>
</tr>
<tr>
<td>5. Has OGD impacted on FOIA requests?</td>
<td></td>
</tr>
<tr>
<td>6. What metadata schema is used for OGD?</td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td></td>
</tr>
<tr>
<td>1. What are your responsibilities for OGD?</td>
<td></td>
</tr>
<tr>
<td>2. What challenges are there OGD?</td>
<td></td>
</tr>
<tr>
<td>3. How is authenticity, reliability and accuracy of OGD guaranteed?</td>
<td></td>
</tr>
<tr>
<td>Roles and Responsibilities</td>
<td></td>
</tr>
<tr>
<td>1. Who is responsible for OGD?</td>
<td></td>
</tr>
<tr>
<td>2. Which departments publish OGD? What kind of data?</td>
<td></td>
</tr>
<tr>
<td>Recordkeeping of open government data</td>
<td></td>
</tr>
<tr>
<td>1. How do OGD and RM systems interact?</td>
<td></td>
</tr>
<tr>
<td>2. What functional classification is used for OGD?</td>
<td></td>
</tr>
<tr>
<td>3. Is OGD linked to other data? How?</td>
<td></td>
</tr>
<tr>
<td>4. What OGD publication formats?</td>
<td></td>
</tr>
<tr>
<td>5. How to ensure accessibility of OGD?</td>
<td></td>
</tr>
<tr>
<td>6. What OGD Policy training?</td>
<td></td>
</tr>
<tr>
<td>Users of open government data</td>
<td></td>
</tr>
<tr>
<td>1. Which groups use OGD?</td>
<td></td>
</tr>
<tr>
<td>2. How does OGD encourage public participation?</td>
<td></td>
</tr>
<tr>
<td>End questions</td>
<td></td>
</tr>
<tr>
<td>1. What are future plans for OGD?</td>
<td></td>
</tr>
</tbody>
</table>