Future Internet **2014**, *6*, 397-413; doi:10.3390/fi6020397

OPEN ACCESS

future internet

ISSN 1999-5903

www.mdpi.com/journal/futureinternet

Article

Privacy and Open Government

Teresa Scassa

Faculty of Law, University of Ottawa, 57 Louis Pasteur, Ottawa, ON K1H 6W9, Canada;
E-Mail: tscassa@uottawa.ca; Tel.: +1-613-562-5800 (ext. 3872); Fax: +1-613-562-5124.

Received: 25 February 2014; in revised form: 3 May 2014 / Accepted: 22 May 2014 / Published: 18 June 2014

Abstract: The public-oriented goals of the open government movement promise increased transparency and accountability of governments, enhanced citizen engagement and participation, improved service delivery, economic development and the stimulation of innovation. In part, these goals are to be achieved by making more and more government information public in reusable formats and under open licences. This paper identifies three broad privacy challenges raised by open government. The first is how to balance privacy with transparency and accountability in the context of "public" personal information. The second challenge flows from the disruption of traditional approaches to privacy based on a collapse of the distinctions between public and private sector actors. The third challenge is that of the potential for open government data—even if anonymized—to contribute to the big data environment in which citizens and their activities are increasingly monitored and profiled.

Keywords: open government; privacy; data protection; big data; open data; personal information

1. Introduction

The global Open Government movement places the transparency and accountability of governments at its core. Other objectives include increased citizen participation and engagement with government, economic development, and the stimulation of innovation [1]. While these goals are all of indisputable importance, their pursuit may have unanticipated consequences for citizen privacy, especially in the oft-promoted open-by-default approach to the management of government data. The privacy implications are rooted in some core realities: governments collect a vast range of data, including the personal information of individuals; the large volume of data already available in our "big data" environment means that even apparently innocuous or anonymized government data sets can contribute to the identification of individuals when matched with other available data; and, in some cases, the very technologies increasingly used by governments to interact with citizens generate new layers of personal information. Thus although the Open Government movement is not meant to displace privacy as an important social value, its implementation does raise privacy challenges.

This paper begins by identifying and outlining three core components of open government. Section 3 situates the open government movement within the broader global information context. Section 4 identifies and discusses certain privacy challenges raised by open government. These include the difficulty in balancing the goals of transparency and accountability with privacy in the context of "public" personal information, the collapse of public and private distinctions, and the role of government data in big data. While the issues discussed in this paper have a broad application, specific examples are drawn primarily from the Canadian and U.S. contexts.

2. Open Government

Yu and Robinson [2] situate the roots of the Open Government movement in the post-World War II era, although pressure for greater government transparency and accountability may well have a much longer history [3]. Nevertheless, the contemporary Open Government movement combines values of transparency and accountability with modern information technologies and the economic value of data [2,4]. This version of Open Government has taken the world by storm within a relatively short period of time [2], and has enlisted the commitment of a significant number of the world's countries [5]. Many of the world's major democracies—and some emerging democracies as well—have subscribed to the Open Government Partnership (OGP), an international multi-stakeholder movement dedicated to advancing open government. States that are members of the OGP commit to establishing and meeting goals within the framework established by that organization. Other commitments to open government are also seen in the recent signing of the Open Data Charter by G8 leaders [6], and in the work being done by the Organisation for Economic Co-operation and Development (OECD) on open government [7].

The goals of the contemporary Open Government movement go beyond transparency and accountability, although these are still central objectives [3,5] In addition to transparency and accountability, open government seeks to increase citizen participation in governance through a variety of new forms of engagement, some of which may include social media. In addition, open government seeks to promote innovation and economic growth through the release by governments of large amounts of data in reusable digital formats and with few or no restrictions on reuse [4]. At the municipal level this can include real-time GPS transit data feeds, restaurant inspection reports, and permit data. At other levels of government it may include geospatial data, statistical data, public health information—the possibilities are virtually limitless.

These goals of open government can be loosely organized into three core components: open access, open data and open engagement. Although there is some overlap between these three components, they each have distinct features that are important to a consideration of privacy issues. Open access builds upon the "right to information" movement, which emphasizes the fundamental rights of citizens to government transparency and accountability [3]. Access to information regimes have been traditionally based upon individual requests for access to government information, and privacy considerations have long been built into the processes regarding the disclosure of information through these mechanisms. Such regimes are premised upon evaluation by data custodians of the potential privacy implications of the release of the sought-after data, as well as decision-making processes that may permit partial release of information with redaction of personal information. While open access can be enhanced by streamlining access to information procedures, improving response times and reducing costs for access [8], one of the ways in which the open government movement seeks to enhance access to government information is by encouraging its proactive release [1]. This is sometimes called "open by default". For example, in its 2012 Open Government Directive, the Obama administration in the US set targets and guidelines for the proactive release of data in the hands of government departments and agencies [8]. Similar targets have recently been adopted in the UK [9] and in Canada [10]. The proactive release of some categories of data can further the goals of access to information, particularly where it is the kind of data that is regularly sought by civil society groups through access to information requests. The open access dimension of open government responds to the needs of civil society groups that seek to hold government accountable [3]. It is also linked to citizen engagement in society, as an informed public is better placed to give input on policy questions. In spite of the overlaps, proactive disclosure will not result in a complete conflation of open access with open data. Access to information regimes will still be necessary for those seeking the release of data that is not selected for proactive disclosure, and they will also be important for enabling individuals to access their own personal information in the hands of government. From a privacy perspective, what is significant is the shift from a case-by-case assessment of the privacy impacts of the release of certain types of data towards a single upfront assessment of data sets that are designated to be open by default.

Open data overlaps to some extent with the proactive release of data under "open access", but it encompasses much more than the proactive release of data sought by civil society groups with a view to greater government transparency and accountability [2]. Janssen notes that there are many similarities between open access (or the right to information) and open data, but she emphasizes as well the fact that open data embraces goals that go beyond transparency and accountability, and include those related to innovation and efficiency [3] (p. 4). Open data is thus the component of open government that best serves the goals of supporting innovation and economic development.

Open data involves the publication of non-personal, non-confidential government data in reusable electronic formats and under an open licence [11]. Openness has been defined in this context, as arising where "[a] piece of data or content is open if anyone is free to use, reuse, and redistribute it—subject only, at most, to the requirement to attribute and/or share-alike" [12]. While some open data sets may be important to civil society groups seeking to hold the government to account, the open data movement has a strong innovation orientation [4], and thus it frequently includes important sets of base geographical or geospatial data, as well as other data sets with significant potential for commercial application. Ideally, the data sets are released in formats that have been developed according to standards that will facilitate their reuse in conjunction with data from other sources. The open data movement also encourages the making of data sets available through a single portal and in a way that makes them easy to locate [11]. The goal of open data is not just to make government information *public*; it is to make it as useful as possible. To some extent, it is also to push for reuse of government data by the private sector [4,13]. Governments may use competitions (app contests, hackathons) and other incentives to encourage individuals to find innovative uses for the data sets that they release [5,11,14]. From a privacy perspective, it is important to note that the inclusion of data sets in an open data portal generally means that the sets have been judged to be free of protected personal information. Whether this is actually the case is, as will be discussed below, sometimes contestable.

The third component of open government is open engagement. Open engagement has several dimensions, some of which overlap with open access and open data. For example, the Open Government Declaration [1] includes holding governments to account and monitoring their activities as one element of citizen engagement; this form of engagement is dependent upon the ability to access information and government data. At the same time, open engagement can involve greater citizen participation in government decision-making and planning, using novel methods [15]. Some governments have experimented with the use of social media such as Facebook, Twitter and Google Plus, for interactive citizen participation in consultations and in providing feedback on government initiatives [16]. On a more practical and day-to-day level, open engagement may also involve the use of online tools to facilitate the reporting by citizens of problems in their neighborhoods, or to engage citizens creatively in planning or other processes [16,17]. As will be discussed in greater detail in Section 4.2, these new forms of citizen engagement, particularly where mediated by private sector companies, can have privacy implications.

3. Open Government in Its Broader Context

The same influences which have radically changed the manner in which information is generated and communicated in other spheres have also been driving forces in moving forward the Open Government agenda [18]. Powerful technologies for gathering and processing information are now in the hands of ordinary individuals, leading to new forms of creative and collaborative engagement across a broad range of fields. There has been a stunning growth in user-generated content of all kinds [19,20]. The crowd-sourcing of information, citizen science, and the crowd-funding of innovative, entrepreneurial and creative activities are also blossoming. Contemporary culture has embraced digital openness and sharing through open licensing, open source software, and open data standards. These developments have led to a so-called “democratization” of knowledge creation and dissemination [19,21,22]. The generation and processing of information can now be carried out by any individual in possession of the powerful and relatively inexpensive software and hardware that is so readily available today. Easy access to the Internet also means that the dissemination of information need no longer be channelled through established and authoritative intermediaries [19]. These changes have implications for privacy, as they increase the number of potential publishers of information, while at the same time decreasing the levels of control over what information is disseminated. Data protection regimes have tended to exclude “private” activities in relation to personal information from their purview, leaving citizen-driven and “non-commercial” dissemination outside the reach of data protection laws.

This vibrant and disintermediated information context also means that open government is not exclusively under the direction or control of governments. Indeed, the difficulty of controlling information is one of the great challenges (and perhaps blessings) of the information age. As a result, open government can be said to have both its official and its “unofficial”—or even illicit—versions. In addition to the movement by governments towards the greater and more proactive disclosure of data prompted by the shift towards open government, the same digital technologies that facilitate (and even demand) open government, also permit citizens to "open" governments that otherwise seek to be closed. This can happen in a number of different ways. In some cases, where governments are seen to be too slow to react to particular problems or crises, citizens have used available technology to collect, compile and disseminate information in a manner that is essentially in competition with the government's role as an information source. In Japan, for example, the tragic experience of the 2011 tsunami and nuclear disaster proved to be a driving force for open data after citizens and the private sector used crowdsourcing to create important disaster-related data tools for the public [23]. In Canada, the failure of the federal government to provide information about the decades-long problem of missing and murdered aboriginal women prompted the group Anonymous to establish its own interactive online map detailing cases of murder and abuse [24]. Wikileaks has been a thorn in the side of many governments as it has provided a forum for whistleblowers to publish government information about otherwise secret or non-transparent processes [2]. Similarly, intense U.S. government secrecy about its spying and surveillance practices was dealt a blow when Edward Snowden, in conjunction with the Guardian newspaper, began releasing highly sensitive information detailing the practices of not only U.S.-based agencies but also those of some of its allies [25].

Although the publication of confidential or secret government information may be illegal, these activities are also very much a product of the democratization and disintermediation of knowledge creation and dissemination. Just as traditional cultural industries have been challenged by the widespread adoption of the technologies of reproduction and dissemination, the traditional channels for holding governments accountable (or not) face stiff competition from citizens armed with cell phone cameras and data sticks. Even while governments struggle to appear more transparent and more accountable by embracing open government, individual and group actors are choosing their own routes to achieving those same goals. This contextualization of the open government movement is important because it underlines the fact that the driving forces are not all within government control. While this has important lessons for government, it is also significant from a privacy perspective [2]. While proponents of Open Government still maintain the importance of privacy and the need for governments to protect the personal information they gather from their citizens, a government's ability to control the information it possesses may be considerably less robust than public sector data protection laws require. As a result, there may be significant risks to privacy from "opened" government.

4. Privacy and Open Government

Privacy is not an easy concept to define [26], and attempts at definition have ranged from the classic "right to be left alone" [27] to more nuanced concepts of privacy as integral to dignity, autonomy and integrity [28], or as consisting of "many different yet related things" [26] (p. 9). Although there are many theories and critiques of privacy, it is generally the case the contemporary data protection norms are based upon a liberal view of privacy premised both upon the autonomous individual who can make informed choices about the sharing of her personal information, and a certain separation between public and private—between the individual and the state [29]. In this view, individual privacy is seen as essential to a person's ability to act autonomously and to be free from undue interference from the state.

Bennett and Raab argue that the enactment of public and private sector data protection laws in Western democracies over the last quarter of the twentieth century was a reaction to the "recognition of the power of new information and communication technologies in the hands of large public and private agencies" [29] (p. xviii). In liberal democracies these laws have tended to reflect a conception of privacy as an individual right, and one that is relatively subjective. The goal is to preserve autonomy by protecting privacy, but at the same time to leave room for autonomy by preserving choice in relation to one's personal information. In other words, the legal framework for protecting privacy reflects a view that different individuals in different contexts may make different choices about the degree of privacy protection they want. It is no surprise, then, that data protection laws tend to be based on "control" models [29,30] that permit individuals some latitude in choosing whether and to what extent their personal information will be collected, used or disclosed. In the public sector context, where individuals may have fewer choices about whether and when the government will collect personal information, the obligations on government to protect this data tend to be more onerous.

Bennett and Raab [29] (p. 6) suggest that the protection of privacy in relation to information in the hands of government "is arguably related to wider attitudes about participation in public affairs and about trust in the authority of government agencies". It is worth noting, therefore, that the protection of personal information in the hands of government may also serve some of the goals of open government, particularly those relating to the encouragement of citizen participation and engagement with government.

The privacy challenges brought about by the open government movement relate to an evolving technological context that increasingly blurs the boundaries between public and private, that sees a shift in relationships between governments and citizens, and that has created an increasingly democratized and disintermediated information environment. In this section, three particular privacy challenges for open government are discussed. The first relates to finding the appropriate balance between values of privacy and those of transparency and accountability in the context of public personal information, the second is the continuing collapse of public and private spheres and the implications for traditional models of data protection, and the third is the risk of disclosure of personally identifying information through open data.

4.1. Finding the Right Balance: Public Personal Information

The first privacy challenge is part of an overarching problem: how can the objectives of open government be balanced with privacy values? As noted earlier, data protection regimes are aimed, in part, at boosting citizen confidence or trust in government so as to enhance public participation. In this sense, at least, the protection of privacy would seem to match the goals of open government. Yet the transparency and accountability objectives of open government push towards greater disclosure of information in the hands of government, and it is here that there is the greatest tension between privacy and open government [31]. There is a large volume of personal information in the hands of government, and in some cases, principles of transparency and accountability require its disclosure. To what extent, though, do the demands for proactive disclosure of data in reusable formats create privacy risks for so-called "public" personal information. "Public personal information" is information about identifiable individuals that is in the hands of government, and that is mandated by law or regulation to be made public. It may include information about political campaign contributions, public servant salary information, building or renovation permits, land titles information, and so on.

Perhaps because they have long been required to operate under the public eye, and because courts have a very long history of publishing the reasons for their decisions—which often contain detailed personal information—the judicial branch of government has had considerable experience with the challenges of balancing privacy with transparency and accountability. It has been suggested that the principle of open justice dates back to the Magna Carta [32], and is reflected in the U.S. in the First Amendment right of the public to have access to court proceedings [33]. In state constitutions it exists in the form of an open courts rule [32]. Information related to court proceedings is presumptively public, notwithstanding the fact that much of it may involve highly sensitive and personal information. In Canada, the open courts principle plays a similar role, with the Supreme Court of Canada describing it as "necessary to maintaining the independence and impartiality of courts" and "integral to public confidence in the justice system and the public's understanding of the administration of justice" [34] (paragraph 25). Because trials themselves are by default public, and because courts publish reasons for decisions that often contain a detailed recitation of key facts, in the context of the courts, we often see a level of disclosure of highly personal information that is not replicated elsewhere in government.

It is instructive to consider some of the challenges that courts and administrative tribunals have faced with respect to the balance they strike between openness and privacy, particularly in the Internet age. Thus, in the U.S. and Canada, although court and tribunal reasons for decisions have always been considered public in nature, the ability to make these documents publicly available over the Internet, through court and tribunal websites, for example, has sparked controversy over whether the nature and quantity of personal information available in these documents should be made available on such a large, public and readily searchable scale [35,36,37]. This has prompted consideration of whether Internet publication is necessary or desirable to achieve the goals of transparency. Others have debated *what* information in tribunal proceedings must be disclosed to ensure transparency, with some tribunals choosing to redact certain information before the online publication of decisions [35,38].

The experience of courts and tribunals will be instructive in considering the balance between privacy on the one hand and transparency/accountability on the other when considering how to deal with "public" personal information in the context of open government. There is a large body of personal information in the hands of governments that is "public" according to various laws or regulations. In many cases, decisions around the public nature of the information were made in an era before the Internet. Because this information is considered to be "public", its availability under access to information regimes or through the consultation of public registers may seem self-evident. Yet these "public" access points contain their own limitations; an effort—sometimes even travel—is required to manually search public records. In an era of open government, it may seem self-evident that these records and registers be digitized and made available for a more open release; yet doing so may raise significant privacy concerns [39]. Digitized information can be rapidly copied, mined and matched, and can be used for a broad range of purposes that many would consider privacy invasive.

An example of a controversial use of public registry data is found in the publication by a newspaper of interactive, online maps showing the names and addresses of all registered gun owners in two New York counties following the tragic school shooting in Newtown Connecticut [40,41]. These maps, by combining public registry data obtained through an access to information request with a Google Maps interface, matched registry information with specific houses in a searchable and interactive format. The "public" personal information was thus published in a format that made it much more accessible both in terms of its presentation and its dissemination. It was also placed within a polemical context. When the maps were published online, citizens expressed outrage either at being identified as a household for which a gun permit had been issued, or at being identified as one for which one had not. There was a strong sense that this information, which had been acceptably public when contained in a register accessible only through a government office or an access to information request, was unacceptably public when it was represented on an interactive map and posted on the Internet. As an additional complicating factor, either the data was released by government with inadequate metadata, or its users did not take appropriate care in examining the metadata. Data that was accurate within the particular limitations of the gun registry was wildly inaccurate when mapped in this manner. The inaccuracies compounded the privacy problems associated with the release of the data [42].

There is precedent for the refusal to disclose public personal information in reusable digital formats. In *Matter of New York Times Co. v. City of New York Police Dept.* [43], the appellate division of the New York State Supreme Court considered a request filed by the New York Times for access to an electronic copy of a database of the names and addresses of all residents of New York City with handgun licences. The appellate division of the New York State Supreme Court denied disclosure of the database notwithstanding that the information it contained was a matter of public record. The court stated: "The fact that Penal Law §400.00 (5) makes the name and address of a handgun license holder "a public record" is not dispositive of whether respondent can assert the privacy and safety exemptions to FOIL [Freedom of Information Law] disclosure" [43]. The court went further, noting that this was so "especially when petitioners seek the names and addresses in electronic form" [43]. It also indicated that other case law supported the view that the disclosure of a person's home address "implicates a heightened privacy concern."

Another example of the conflict between public government data and privacy can be found in the Proposition 8 map [44]. Proposition 8 was a proposed amendment to the California constitution that sought to ban gay marriage. Because it was to be put to a referendum, and because election campaign contributions are a matter of public record, it was not difficult for opponents of the proposition to create an online interactive map that matched the name of each donor who supported Proposition 8, along with the amount of their donation to their street address. While the goal of public political campaign finance lists is transparency and accountability, this information can have significant privacy impacts depending on how it is disseminated publicly. The experience around Proposition 8 led to calls for changes to the laws around disclosure of campaign finance information in order to protect donor privacy [45]. The Proposition 8 map is a high-profile example of a major challenge in the context of public personal information. Although the problem of "public" personal information pre-dates the open government movement, it will only be exacerbated in the context of open data. If such personal information is considered "public", should a government make the data available through proactive disclosure? How can the competing goals of privacy protection and transparency be appropriately balanced? As the experience of courts and tribunals shows, it may sometimes be necessary to place limits on the digital disclosure of some of the information in a "public" record in order to achieve this balance.

In theory, private sector data protection laws—where they exist—would apply to the collection, use and disclosure of publicly available personal information provided by government. However, such laws will generally only apply where there is a commercial dimension to the re-use of the information. Private or non-commercial activities generally fall outside the reach of data protection legislation. In addition, uses that serve freedom of expression goals, such as journalism (broadly defined) may also be exempt [46]. Even where there is a commercial use, data protection laws will generally require that such use be consistent with the purposes for which the information is made available to the public [47]. Yet these purposes may not always be entirely evident. For example, in the context of court or tribunal decisions, the reasons for publication may range from the promotion of transparency and accountability in the administration of justice to the provision of useful precedents to those who may engage in similar proceedings. It may also not be entirely clear what re-uses actually serve a stated purpose. In the Proposition 8 case, if the purpose for the disclosure of the information is to ensure transparency and accountability in the context of campaign financing, it might be possible to argue that being able to identify trends or patterns in financing by citizens based on particular geo-demographic considerations also serves the goals of transparency and accountability by revealing important information about patterns of contributions [45]. Ultimately, the open licences that would accompany data released as open data have no mechanism in them to limit the use of data to particular purposes, including those for which the data was originally collected.

These examples illustrate that the balance between privacy on the one hand and transparency and accountability on the other can be difficult to achieve. Certainly, the experience of courts and tribunals demonstrates that in some cases privacy interests must give way to transparency and accountability. In the government context, while some personal information is considered to be "public", its reuse is limited to purposes consistent with the goals of its original collection. Not only are these purposes often ambiguous, they are likely impossible to police or enforce in the fast-moving data world. Yet a tighter control over this information would limit transparency and accountability. Thus one challenge for governments is to reconsider the nature and extent of "public" personal information that is disclosed in light of both privacy and transparency considerations. This may involve both a consideration of the potential harm (direct and indirect) to individuals that may flow from disclosure as well as the extent to which all of the personal information in the records at issue is necessary to achieving the goals of transparency. It may also require more explicit statements of the purposes for which the data was collected, and, by extension, the purposes for which its re-use is authorized.

4.2. A Collapse of Public and Private

A second privacy challenge lies in the fact that data protection has largely been structured along public and private lines [29]. Thus public and private sector actors tend to be governed by different regimes. Governments tend to be held to fairly strict standards with respect to personal information collected from individuals. By contrast, the regulation of privacy practices in the private sector may be relatively loose, with a focus on obtaining customer consent to increasingly complex and often largely unread privacy policies [48]. In Canada, private sector data protection laws set a flexible framework for the collection, use and disclosure of personal information; in the U.S., such norms tend either to be voluntary or sectoral, except in the case of the personal information of children [49].

This separation of public and private sectors may be increasingly difficult to manage. The extent of the problem is evident in recent controversies over government surveillance of citizens in both the U.S.

and in Canada. In both countries, governments have tapped into the vast information resources of private sector companies—sometimes with the consent and cooperation of those companies [50,51,52]. Following public outcry over leaked information about the spying activities of the U.S. National Security Administration (NSA), the Center for Democracy and Technology in the U.S., in concert with major Internet companies, announced its support for bills that would permit corporations to publish data about requests they receive from state authorities for customer information [53]. This illustrates how the line between public and private sectors can dissolve in the context of privacy; it also suggests that transparency in relation to government activities may also require transparency on the part of private sector corporations.

The blurring between public and private occurs in other contexts as well. The open engagement component of open government involves governments finding new ways to engage citizens. This may involve the use of social media tools as platforms for discussion or exchange of views. In some instances, private sector companies also provide the digital infrastructure for online feedback or complaints mechanisms, or government information is shared with the public using the services of a private sector company. In these examples, the private sector company acts as an intermediary between citizen and government—sometimes as a client of the government, but other times simply as the provider of a platform chosen by the government as a medium for interaction.

In some cases, governments may choose to interact with citizens through social networking platforms such as Facebook, Twitter or Google Plus [15,54,55]. These companies collect, use and disclose consumer data in ways that are insulated from the stricter norms that would otherwise protect citizen data shared with government. In the case of Facebook, for example, there have been numerous high profile privacy concerns [56,57]. The common feature in all these cases is that citizens are encouraged to communicate information to governments via a private sector intermediary. Some of this information may include personal information—either directly (e.g., a named individual providing an opinion about an issue within the community) or indirectly (e.g., the individual providing information in text or photo format (or in a combination of the two) that can lead to her identification.) It may also be the case that the service requires users to register or create an account in order to participate. In these instances, although citizens may consider themselves to be directly engaging with government, they are doing so via a private sector intermediary that may (or may not) also be collecting and recording their personal information. The standards for data protection for private sector companies are typically quite different from those applicable to government, and there may be a gap between expectations and reality [58]. As Bertot *et al.* [59] (p. 32), observe, "[b]y adopting the use of specific social media tools, government agencies appear to be tacitly endorsing the privacy, security and other policies employed by these social media providers". Further, this form of citizen engagement may also permit governments to gain access to an additional layer of personal information where the individuals making the contributions can be linked to social networking profiles [58]. The blurring of the separation between public and private sectors in the data protection context will create privacy challenges in the context of open government. More and more data will be generated by governments carrying out what appear to be traditional government functions (delivery of services, interaction with citizens). Yet the lines blur between the data that are collected, used or disclosed by the public sector and which is collected, used and disclosed by private sector actors. In a data protection environment that places different controls on government than it does on the private sector, both the inherent ambiguity and the drift towards the private sector should be matters of concern from a privacy perspective.

4.3. Big Data, Open Government and Privacy

A third privacy issue relates to the release of government data through open data programs within the broader big data context. "Big data" is a term used to refer to the massive and ever-expanding volume of digital data [60,61]. Although government data sets may, on their own be "small data" [62], when released as open data they enter the big data ecosphere. Such data sets, made publicly available in reusable digital formats, may appear to be free of personal information or may have been anonymized. Nevertheless, their use in combination with other available data, can pose real risks to privacy. Sophisticated algorithms can be used to match these different data sets and to re-identify specific individuals, contributing to widespread practices around profiling of individuals [63,64]. The oft-cited examples of the release of anonymized customer data by both AOL and by Netflix [64,65] illustrate how the many minds that are the engine of Web 2.0 can, working alone or in combination, quickly lead to the identification of specific individuals in anonymized data sets based upon pattern identification and data matching. Ironically, anonymization, while protecting privacy, may undercut the usefulness of the data. As Ohm notes, "Data can either be useful or perfectly anonymous but never both" [64] (p. 1704). Thus the objective of open data—the release of useful data sets in reusable formats—does not always sit easily with privacy principles.

Public sector data protection laws define personal information broadly, and most definitions will include information that does not directly identify an individual, but that may lead to such identification [66,67]. As a result, there is a real risk that open data initiatives may lead to breaches of data protection obligations. This has raised concerns that privacy laws may actually be in tension with the potential benefits of open data [31]. At the same time, the massive release of government data through open data programs may add further fuel to the insatiable engines of big data, and thus may further undermine the control that individuals are meant to have over their personal information.

While transparency and accountability are touted as prime motivators for open government, the language of entrepreneurship and innovation is also clearly present in the rhetoric around open government [4,18]. While many users of open data are interested in building apps and tools based on government data, there is nothing to prevent the use of open data in profiling, data mining, and other activities which have privacy implications for individuals. The re-identification and profiling risks related to open government data are only likely to intensify as governments at all levels diversify and amplify the information they collect. Popular concepts such as "smart cities" or "intelligent transportation" envisage an environment in which all manner of data about individual activities is collected and recorded either by government entities or by the private sector service suppliers with which they have contracted. Although the primary objectives may be to improve services or planning, the data will be highly sought after by the private sector and by researchers for a broad range of purposes. The release of such data, even if anonymized, may only serve to enhance privacy risks. Yet, as noted earlier, privacy restrictions on open data may undermine data quality, hampering re-use of the data, or blunting transparency and accountability.

This third consideration—the impact of big data on privacy in the context of open government—raises concerns about the extent to which government data will contribute to increased profiling of individuals or to the re-identification of specific individuals. These are clearly privacy risks; the challenge is to assess these risks in light of the benefits of potentially greater transparency and accountability, improved services, and increased innovation. In a sense, then, this returns us to the first issue considered—the need to recalibrate the balance between privacy interests and the competing interests present in the context of open government.

5. Conclusions

This paper has placed open government within the evolving and emerging Web 2.0 environment, and has identified a number of privacy challenges that arise in consequence. The discussion focussed on three particular issues: the balance between privacy protection and the goals of transparency and accountability in the context of public personal information; the collapse of the public/private framework for data protection; and the relationship of open data to big data. The diminishment of privacy generally may lead to harms such as the loss of dignity or autonomy, or may raise concerns about security or physical integrity.

In considering the balance between transparency and accountability on the one hand and privacy on the other, the challenge for governments is to achieve the desired transparency and accountability without exposing individuals to an unwarranted loss of privacy. In the case of public personal information, governments at all levels must consider whether the amount of personal information disclosed in the public records in the analog environment is appropriate and necessary in a digital and networked environment. In a sense, because public personal information has already been publicly accessible, it may require a reassessment of the degree of openness—a potential clawback, as it were, of some degree of transparency in the interests of privacy. It may require—additionally or alternatively—a greater focus on both data quality and meta data—as illustrated by the example of the gun permit map. More thoughtful approaches on how to effectively communicate meta data may mitigate some of the potential negative impacts of disclosure.

The blurring of public and private sectors in the collection, use and dissemination of personal information, is present as well in the context of open government. To the extent that engagement with government creates data trails in the private sector, and activities in the private sector create data trails that are shared with government, there is a risk not only to individual privacy, but also to the relationship of trust that is meant to exist between citizens and their government.

Finally, the release of government data sets through open government raises challenging questions about the extent to which individuals may be identifiable when this data is combined with other available data. Beyond the challenges for governments of complying with data protection norms, while also making data open, this raises questions about the balance between privacy interests and those of transparency and accountability. The innovation/entrepreneurship goals of open data—which may overshadow goals of transparency and accountability—are a complicating factor in this area.

Each of these areas raises difficult privacy issues, some of which can be linked to policy or privacy issues already encountered in earlier contexts. In the context of a very rapidly evolving environment for open government, where the pressure is on governments to proactively release data, and where the rhetoric is dominated by calls for transparency and accountability, it is important not to lose sight of privacy. In many cases, addressing privacy issues will require small focussed actions, rather than high level legislative change. Necessary action may include the development of guidance on when data sets considered for release may raise privacy issues, and guidance as well on when those privacy issues are overridden by the need for transparency and accountability. Improved approaches to meta data may also be necessary. Policies on use of social media should address privacy and may require additional notice to citizens invited to engage with government through these means. Overarching these smaller steps, however, is the much bigger question of whether, how, and to what extent privacy concerns can be reconciled with the goals of open government.

Acknowledgments

This paper is based upon a presentation made at the Symposium on International Reactions and Responses to the Snowden/NSA Dilemma hosted by Pauline C. Reich of the Asia-Pacific Cyberlaw, Cybercrime and Internet Security Research Institute at Waseda University School of Law in Tokyo, Japan on 14 December 2013. The author gratefully acknowledges the support of Waseda University School of Law, the Canada Research Chairs Programme, and the Social Sciences and Humanities Research Council of Canada. The research assistance of Christiane Saad is gratefully acknowledged. Many thanks to Christopher Berzins, Barbara Bucknell, Patricia Kosseim, Tracey Lauriault, Pauline C. Reich, and Pamela Robinson. Thank you also to anonymous reviewers for their thoughtful and helpful comments.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Open Government Declaration. Open Government Partnership, 2011. Available online: http://www.opengovpartnership.org/about/open-government-declaration (accessed on 29 April 2014).
2. Yu, H.; Robinson, D.G. The New Ambiguity of "Open Government". *UCLA Law Rev. Disc.* **2012**, *59*, 178–208.
3. Janssen, K. Open Government Data: Right to Information 2.0 or its Rollback Version? Available online: http://ssrn.com/abstract=2152566 (accessed on 26 May 2014).
4. Bates, J. This is what modern deregulation looks like: Co-optation and contestation in the shaping of the UK's Open Government Data Initiative. Available online: http://ci-journal.net/index.php/ciej/article/view/845 (accessed on 29 April 2014).
5. Huijboom, N.; Van den Broek, T. Open data: An international comparison of strategies. *Eur. J. ePract.* **2011**, *12*, 1–13. Available online: http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012_1.pdf (accessed on 2 May 2014).

6. Vincent, J. G8 Open Data Charter will "Increase Transparency" and "Fuel Innovation". *The Independent*, 19 June 2013. Available online: http://www.independent.co.uk/life-style/gadgets-and-tech/g8-open-data-charter-will-increase-transparency-and-fuel-innovation-8665696.html (accessed on 2 May 2014).
7. Innovative and Open Government: An Overview of Recent Initiatives. Available online: http://www.oecd.org/governance/ministerial/46342001.pdf (accessed on 2 May 2014).
8. McDermott, P. Building open government. *Gov. Inf. Q.* **2010**, *27*, 401–413.
9. McClean, T. Not with a Bang but a Whimper. The Politics of Accountability and Open Data in the UK. American Political Science Association: Washington, DC, USA, 2011. Available online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1899790 (accessed on 2 May 2014).
10. Canada's Action Plan on Open Government. Government of Canada: Ottawa, ON, Canada, 1 August 2013. Available online: http://data.gc.ca/eng/canadas-action-plan-open-government (accessed on 2 May 2014).
11. Ubaldi, B. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. *OECD Work. Paper. Public Gov.* **2013**, doi: 10.1787/5k46bj4f03s7-en. Available online: http://dx.doi.org/10.1787/5k46bj4f03s7-en (accessed on 2 May 2014).
12. Open Definition. Available online: http://opendefinition.org/ (accessed on 29 April 2014).
13. Robinson, D.; Yu, H.; Zeller, W.P.; Felten, E.W. Government data and the invisible hand. *Yale J. Law Tech.* **2009**, *11*, 160–175.
14. Eyler-Werve, K.; Carlson, V. *Civic Apps Competition Handbook*; O'Reilly: Sebastopol, CA, USA, 2012.
15. Bertot, J.C.; Jaeger, P.T.; Munson, S.; Glaisyer, T. Engaging the public in open government: The policy and government application of social media technology for government transparency. *IEEE Comput.* **2010**, *43*, 53–59.
16. Mergel, I. Goverment 2.0 Revisited: Social Media Strategies in the Public Sector. *PA Times* **2010**, *33*, 7–8. Available online: http://faculty.maxwell.syr.edu/iamergel/files/Summer2010LR.pdf (accessed on 2 May 2014).
17. Suri, M.V. From Crowdsourcing Potholes to Community Policing: Applying Interoperability Theory to Analyze the Expansion of "Open311". Available online: http://works.bepress.com/manik_suri/20/ (accessed on 2 May 2014).
18. O'Reilly, T. Government as a Platform. In *Open Government: Collaboration, Transparency, and Participation in Practice*; Lathrop, D., Ruma, L., Eds.; O'Reilly: Sebastopol, CA, USA, 2010; pp. 11–39.
19. Benkler, Y. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*; Yale University Press: New Haven, CT, USA, 2006.
20. Krumm, J.; Davies, N.; Narayanaswami, C. User-Generated Content. *Perv. Comput.* **2008**, *7*, 10–11.
21. Goodchild, M.F. Citizens as voluntary sensors: Spatial data infrastructure in the World of Web 2.0. *Int. J. Spat. Data Infrastruct. Res.* **2007**, *2*, 24–34.
22. Sunstein, C.R. *Infotopia: How Many Minds Produce Knowledge*; Oxford University Press: Oxford, UK, 2008.

23. Utani, A.; Mizumoto, T.; Oumura, T. How Geeks Responded to a Catastrophic Disaster of a High-Tech Country. In Proceedings of the Special Workshop on Internet and Disasters, SWID'11, Tokyo, Japan, 6–9 December 2011. Available online: http://dl.acm.org/citation.cfm?id=2079369 (accessed on 2 May 2014).
24. Finnegan, S. #OpThunderbird: Anonymous engagement to combat violence against aboriginal women in Canada. Available online: http://www.giswatch.org/sites/default/files/canada_gisw13.pdf (accessed on 2 May 2014).
25. Edward Snowden. Available online: http://www.theguardian.com/world/edward-snowden (accessed on 2 May 2014).
26. Solove, D. *Understanding Privacy*; Harvard University Press: Cambridge, MA, USA, 2009.
27. Warren, S.D.; Brandeis, L.D. The Right to Privacy. *Harv. L. Rev*. **1890**, *4*, 193–220.
28. Cohen, J. Examined lives: Informational privacy and the subject as object. *Stanf. L. Rev.* **2000**, *52*, 1373–1438.
29. Bennett, C.J.; Raab, C. *The Governance of Privacy: Policy Instruments in Global Perspective*; Massachusetts Institute of Technology Press: Cambridge, MA, USA, 2006.
30. Westin, A. *Privacy and Freedom*; Atheneum: New York, NY, USA, 1967.
31. Kulk, S.; van Loenen, B. Brave new open data world? *Int. J. Spat. Data* **2012**, *7*, 196–206.
32. Hoffman, J.M. By the course of the law: The origins of the open courts clause. *Or. Law Rev.* **1995**, *74*, 1279–1318.
33. Maness, A. Does the First Amendment's "Right of Access" Require Court Proceedings to be Televised? A Constitutional and Practical Discussion". *Pepperdine Law Rev*. **2006**, *34*, 123–184.
34. Vancouver Sun (Re), 2004 SCC 43, [2004] 2 SCR 332. Canadian Legal Information Institute: Ottawa, ON, Canada. Available online: http://www.canlii.org/en/ca/scc/doc/2004/2004scc43/2004scc43.html (accessed on 26 May 2014).
35. Berzins, C. Personal information in the adjudicative decisions of administrative tribunals: An argument for limits. *Advocates Q.* **2008**, *34*, 261–284.
36. Berzins, C. Freedom of information, privacy and adjudicative agencies in Ontario: Unresolved issues and emerging concerns. *Advocates Q.* **2006**, *31*, 125–152.
37. Roberts, A. *Blacked Out: Government Secrecy in the Information Age*; Cambridge University Press: Cambridge, UK, 2006.
38. Office of the Privacy Commissioner of Canada. Electronic Disclosure of Personal Information in the Decisions of Administrative Tribunals. Available online: http://www.priv.gc.ca/information/pub/gd_trib_201002_e.cfm (accessed on 2 May 2014).
39. Office of the Information Commissioner of Canada, Annual Report 2012–2013. Available online: http://www.oic-ci.gc.ca/eng/rp-pr_ar-ra.aspx (accessed on 24 February 2013).
40. Haughney, C. After Pinpointing Gun Owners, Paper is a Target, *New York Times*, 6 January 2013. Available online: http://www.nytimes.com/2013/01/07/nyregion/after-pinpointing-gun-owners-journal-news-is-a-target.html?pagewanted=all&_r=0 (accessed on 2 November 2014).
41. Worley, D.R. The gun owner next door: What you don't know about the weapons in your neighborhood. Available online: http://www.lohud.com/apps/pbcs.dll/article?AID=2012312230056&nclick_check=1 (accessed on 2 May 2014).

42. Incalcaterra, L. Many Rockland handgun permits have outdated, inaccurate data, 26 January 2013. Available online: http://www.lohud.com/article/20130127/NEWS03/301270041/Many-Rockland-handgun-permits-outdated-inaccurate-data?nclick_check=1 (accessed on 24 February 2013).
43. Matter of New York Times Co. v. City of New York Police Dept. 2013 NY Slip Op 00686 [103 AD 3d 405]. New York State Law Reporting Bureau: New York, NY, USA, 5 February 2013. Available online: http://www.nycourts.gov/reporter/3dseries/2013/2013_00686.htm (accessed on 30 April 2014).
44. Stone, B. Prop 8 Donor Website Shows Disclosure Law is 2-edged Sword. *New York Times*, 7 February 2009. Available online: http://www.nytimes.com/2009/02/08/business/08stream.html?_r=0 (accessed on 2 May 2014).
45. Lourie, D. Rethinking Donor Disclosure after the Proposition 8 Campaign. *S. Cal. L. Rev*. **2009**, *83*, 133–172.
46. Scassa, T. Journalistic Purposes and Private Sector Data Protection Legislation: Blogs, Tweets, and Information Maps. *Queen's Law J.* **2010**, *35*, 733–781.
47. Scassa, T. Privacy and Publicly Available Information. *Can. J. Law Tech*. **2013**, *11*, 1–23.
48. McDonald, A.M.; Cranor, L.F. The Cost of Reading Privacy Policies. *I/S J. Law Policy Inf. Soc.* **2009**, 4, 543–568.
49. Garber, D.J. COPPA: Protecting children's personal information on the internet. *J. Law Policy* **2001**, *10*, 129–188.
50. Bailey, J. Systematic government access to private sector data in Canada. *Int. Data Priv. Law* **2012**, *2*, 207–219.
51. Cate, F.H. Government data mining: The need for a legal framework. *Harv. Civ. Right Civ. Lib. Law Rev*. **2008**, *43*, 435–490.
52. Hoofnagle, C.J. Big brother's little helpers: How choice point and other commercial data brokers collect and package your data for law enforcement. *N. Car. J. Int Law Commer. Regul.* **2004**, *29*, 595–638.
53. Coalition of Major Internet Companies and Advocates Rallies around Surveillance Transparency Legislation. Available online: https://www.cdt.org/pr_statement/coalition-major-internet-companies-and-advocates-rallies around-surveillance-transparen (accessed on 2 May 2014).
54. Transcript of Tweet Chat. Government of Canada: Ottawa, ON, Canada, 11 December 2012. Available online: http://data.gc.ca/eng/transcript-tweet-chat (accessed on 2 May 2014).
55. Thompson, E. Clement Tweets His Way into Canadian History with Twitter Town Hall. 15 December 2011. Available online: http://www.ipolitics.ca/2011/12/15/clement-tweets-his-way-into-canadian-history-with-twitter-town-hall/ (accessed on 2 May 2014).
56. Goel, V.; Wyatt, E. Facebook Privacy Change is Subject of F.T.C. Inquiry. *New York Times*, 11 September 2013. Available online: http://www.nytimes.com/2013/09/12/technology/personaltech/ftc-looking-into-facebook-privacy-policy.html?_r=0 (accessed on 2 May 2014).
57. Denham, E. Report of Findings into the Complaint Filed by the Canadian Internet Policy and Public Interest Clinic (CIPPIC) against Facebook Inc. under the Personal Information protection and Electronic Documents Act; PIPEDA Case Summary #2009–008; Office of the Privacy Commissioner of Canada.: Ottawa, ON, Canada, 16 July 2009. Available online: https://www.priv.gc.ca/cf-dc/2009/2009_008_0716_e.asp (accessed on 2 May 2014).

58. Citron, D.K. Fulfilling Government 2.0's Promise with Robust Privacy Protections. *Geo. Wash. L. Rev.* **2009**, *78*, 822–845.
59. Bertot, J.C.; Jaeger, P.T.; Hansen, D. The impact of policies on government social media usage: Issues, challenges, and recommendations. *Gov. Inf. Q.* **2012**, *29*, 30–40.
60. Investigation Report P98-011, Office of the Information and Privacy Commissioner for British Columbia: Victoria, BC, Canada, 31 March 1998. Available online: http://www.oipc.bc.ca/report/investigation-reports.aspx (accessed on 2 May 2014).
61. Manyika, J.; Chui, M.; Brown, B.; Bughin, J.; Dobbs, R.; Roxburgh, C.; Byers, A.H. Big Data: The Next Frontier for Innovation, Competition and Productivity. Available online: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (accessed on 2 May 2014).
62. Kitchin, R.; Lauriault, T.P. Small Data, Data Infrastructures and Big Data. The Programmable City Working Paper 1. Available online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376148 (accessed on 2 May 2014).
63. On the Data Trail: How Detailed Information about You Gets into the Hands of Organisations with Whom you Have No Relationship. Canadian Internet Policy and Public Interest Clinic (CIPPIC), University of Ottawa: Ottawa, ON, Canada, 2006. Available online: https://www.cippic.ca/sites/default/files/May1–06/DatabrokerReport.pdf (accessed on 2 May 2014).
64. Ohm, P. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Rev.* **2010**, *57*, 1701–1764.
65. Porter, C.C. De-Identified Data and Third Party Data Mining: The Risk of Re-Identification of Personal Information. *Shidler J.L. Com. Tech.* **2008**, *5*, 1–8.
66. Graux, H. Open Government Data: Reconciling PSI Re-use Rights and Privacy Concerns. European Public Sector Information Platform Topic Report No. 2011/3. Available online: http://www.epsiplatform.eu/sites/default/files/Topic_Report_Privacy.pdf (accessed on 2 May 2014).
67. Scassa, T. Geographic Information as Personal Information. *Oxf. Univ. Commonw. Law J.* **2010**, *10*, 185–214.