Inter-governmental organisations sharing and linking open and real-time data for inclusive governance: development effectiveness and protection of privacy and security

Keisha Taylor, Trinidad and Tobago

Abstract

The rapid rise of the Internet has encouraged the use of open, real-time, and linked data to help understand and improve development processes. This has gained prominence in the public, private, and civil society sectors, as each one independently and collaboratively examines ways in which the vast amounts of data and information generated online can be mapped and linked to help with research and development in all fields, including economics, sustainable development, education, health, agriculture, science, and humanitarian and disaster relief, at local, national, regional, and international levels. The availability of data online is also generating increased possibilities for interdisciplinary study and cross boundary research and analysis. Organisations are not only making data available online for reuse by others but are also using data generated actively and passively by the public to inform business and government decisions. Moreover, individuals are using data for day to day decisions about issues that are of importance to them, their families, and their communities. Figure 1 shows datasets that have been published in Linked Data format, by contributors to the Linking Open Data community project and other individuals and organisations. The advancement of data use for development without an Internet governance framework, however, raises the importance of inclusion of the most marginalized, as well as privacy and security. This paper will examine such issues, as well as the role inter- governmental organisations can play in helping to encourage the use of data while supporting the protection of privacy and security.

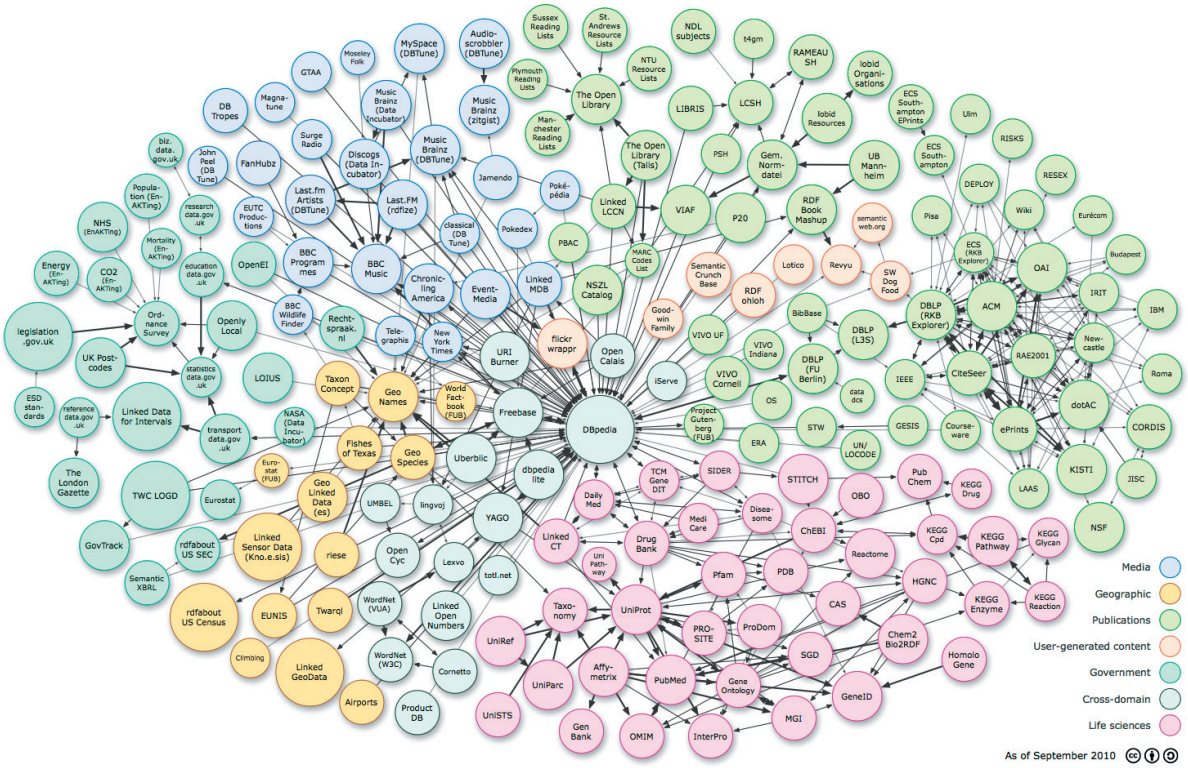
Keywords: open data; real-time data; linked data; inclusive development; privacy and security

Introduction

The growth of the Internet has led to increased data availability, and now, arguably, there has been a worldwide culture shift towards transparency and accountability even though a few decades ago this was a most unlikely scenario. Today, we live in a world that is constantly generating data, which is increasingly being made available. Archived data is coming alive online as open data becomes mainstream, while real-time data on our opinions, how we interact, and where we go, can also be mined to provide very valuable information. Technology has ushered in unpredictable, yet exciting times, as all of this data can also be linked. It can be aggregated, compared, mashed up, and analysed, resulting in an unprecedented opportunity to forecast future trends, outbreaks, measure political and social consciousness, and understand what the public wants and needs and when they want and need it. There is, therefore, tremendous opportunity to develop the use of data to improve development effectiveness and aid overall community, national, and international well-being. This, however, leaves us at a cross roads, for what are the harms of realising the benefits of this data in comparison to keeping it private. Inter-governmental organisations (IGOs) have a role to play in ensuring that open, real-time, and linked data can be used for development in a way that protects the human rights to privacy and security.

This document has been produced by DiploFoundation with the financial assistance of the European Union. The contents of this document are the sole responsibility of the author and can under no circumstances be regarded as reflecting the position of the European Union or DiploFoundation.

Figure 1. 'This image shows datasets that have been published in Linked Data format, by contributors to the Linking Open Data community project and other individuals and organisations. It is based on metadata collected and curated by contributors to the CKAN directory' (Cyganiak and Jentzsch, 2010).



What is open data?

The Open Knowledge Foundation (OKN) offers this definition of open data: A piece of content or data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and share-alike (OKN, no date a).' The OKN also provides useful and overarching guidelines to help determine when data is open. Please see Box 1. which highlights OKN guidelines on what constitutes open data.

Open data is also 'non-personal data' (Local Government Improvement and Development, 2011), and it is usually data that is already in the public domain and is therefore eligible to be open. There is a growing call by civil society organisations to make much more government information openly available to the public, while an increasing number of Freedom of Information Acts (Wikipedia, 2011) demand that governments make information available to citizens who ask for it.

The potential for the use of open government data, arguably, took on renewed emphasis when the United States launched data.gov in May 2009. The next government to launch an open data website was Australia with data.australia.gov.au, followed by New Zealand with data.govt.nz. The United Kingdom's (UK), data.gov.uk, was then launched in January 2010 with support from the founder of the World Wide Web, Tim Berners Lee. Cities[1] in Canada, the UK, and the USA (Sherman, 2011) are also doing the same, and government departments in other countries are exploring the possibilities of open access on a departmental level. Governments are also challenging other stakeholders to make use of their data and share it with others in a way that can help them to better understand its complexities and make better informed decisions. Open data can also be institutional; for example, research institutions are increasingly making their academic information public.

Of course, it is technology that is aiding this open data drive, as governments and other institutions digitise their information, and the filing cabi-

1 Data websites currently available include ca.gov/data data.octo.dc.gov toronto.ca/open. Visit for a list of all open government data sites. CTIC, Open data@ctic, http://datos.fundacionctic.org/sandbox/catalog/faceted/ [accessed 25 February 2011].

Box 1. OKN guidelines

1. **Access**
   The work shall be available as a whole, and at no more than a reasonable reproduction cost, preferably downloading via the Internet without charge. The work must also be available in a convenient and modifiable form.
2. **Redistribution**
   The license shall not restrict any party from selling or giving away the work either on its own or as part of a package made from works from many different sources. The license shall not require a royalty or other fee for such sale or distribution.
3. **Reuse**
   The license must allow for modifications and derivative works and must allow them to be distributed under the terms of the original work.
4. **Absence of Technological Restriction**
   The work must be provided in such a form that there are no technological obstacles to the performance of the above activities. This can be achieved by the provision of the work in an open data format, i.e. one whose specification is publicly and freely available and which places no restrictions monetary or otherwise upon its use.
5. **Attribution**
   The license may require as a condition for redistribution and re-use the attribution of the contributors and creators to the work. If this condition is imposed it must not be onerous. For example if attribution is required a list of those requiring attribution should accompany the work.
6. **Integrity**
   The license may require as a condition for the work being distributed in modified form that the resulting work carry a different name or version number from the original work.
7. **No Discrimination Against Persons or Groups**
   The license must not discriminate against any person or group of persons.
8. **No Discrimination Against Fields of Endeavor**
   The license must not restrict anyone from making use of the work in a specific field of endeavor. For example, it may not restrict the work from being used in a business, or from being used for genetic research.
9. **Distribution of License**
   The rights attached to the work must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.
10. **License Must Not Be Specific to a Package**
    The rights attached to the work must not depend on the work being part of a particular package. If the work is extracted from that package and used or distributed within the terms of the work's license, all parties to whom the work is redistributed should have the same rights as those that are granted in conjunction with the original package.
11. **License Must Not Restrict the Distribution of Other Works**
    The license must not place restrictions on other works that are distributed along with the licensed work. For example, the license must not insist that all other works distributed on the same medium are open.

(OKN, no date b)

net increasingly becomes a thing of the past. In the face of a Web 2.0 digitised environment and the emerging cloud computing marketplace, the opportunities to make data open is on the rise.

What is real-time data?

Real-time data refers to data that is delivered to a third party immediately after it is generated. Data feeds generated via social media sites like Facebook and Twitter, or information generated by machines and related information and communication technology (ICT) devices, such as mobile phones, also provide real-time data. The value of tracking real time information is increasingly being recognised around the world, and the use of real-time analytics is being used by the private sector to help them compete in various industries. Much of this data is also being linked by the use of geolocation or geomapping technology.

Crowdsourcing, which enables the outsourcing of data gathering and other microtasks to the public to help find answers to problems and utilise a remote workforce has also been acknowledged as a useful tool. Real-time data is increasingly being explored as a way of gaining access to vital information needed to develop public services, like health, education, and agriculture, but also to monitor political unrest.

Box 2. The four Linked Data principles advocate the use of:

1. URI references to identify, not just Web documents and digital content, but also real world objects and abstract concepts. These may include tangible things such as people, places and cars, or those that are more abstract, such as the relationship type of *knowing somebody*, the set of all green cars in the world, or the color green itself. This principle can be seen as extending the scope of the Web from online resources to encompass any object or concept in the world.
2. HTTP URIs to identify objects and abstract concepts, enabling these URIs to be *dereferenced* (i.e., looked up) over the HTTP protocol into a description of the identified object or concept. The HTTP protocol is the Web's universal access mechanism. In the classic Web, HTTP URIs are used to combine globally unique identification with a simple, well-understood retrieval mechanism. Thus, In order to enable a wide range of different applications to process Web content, it is important to agree on standardized content formats. The agreement on HTML as a dominant document format was an important factor that made the Web scale.
3. A single data model for publishing structured data on the Web – the Resource Description Framework (RDF), a simple graph-based data model that has been designed for use in the context of the Web. (The underlying structure of any expression in RDF is a collection of triples, each consisting of a subject, a predicate and an object. Please see figure 2 for an illustration of RDF triplets.)
4. Hyperlinks to connect not only Web documents, but any type of thing. For example, a hyperlink may be set between a person and a place, or between a place and a company. In contrast to the classic Web where hyperlinks are largely untyped, hyperlinks that connect things in a Linked Data context have types which describe the relationship between the things. For example, a hyperlink of the type *friend of* may be set between two people, or a hyperlink of the type *based near* may be set between a person and a place. Hyperlinks in the Linked Data context are called *RDF links* in order to distinguish them from hyperlinks between classic Web documents.

   In summary, the Linked Data principles lay the foundations for extending the Web with a global data space based on the same architectural principles as the classic document Web.

   (Heath and Bizer, 2011).

For instance, civil-society-based organisation Ushahadi develops free and open source software for collecting and crowdsourcing real-time data, which is then visualised and mapped interactively. One journalist reflecting on the importance of the Ushahadi-type platform during the 2010 Haitian earthquake said:

'The most notable innovations to emerge from Haiti were: the translation of crowd-sourced data to actionable information; the use of SMS message broadcasting in a crisis; and crowdsourcing of open maps for humanitarian application . . . A dizzying array of new media and information technology groups, Haitian diaspora networks and media development partners were involved in these initiatives. Although these innovations had varying levels of impact in Haiti, they showcased the potential for use in future crisis' (Nelson, 2011).

Real-time data is an invaluable way of understanding trends which can aid the best estimation about where resources should be allocated at a specific place and time, as well as identify what is most important for the public, not by the day or the hour but by the fraction of a second.

What is linked data?

Linked data refers to 'using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods' (Linked Data, no date). In other words, linking data helps to find other related data from across the web. When successful, it could also help to improve the reliability and validity of data and information generated.

Tim Berners-Lee put forward a set of 'Linked Data Principles' for publishing Web data to help develop a 'single global data space' (Berners-Lee, 2000).

These include:

1. Use of Uniform Resource Identifiers (URIs)[2] as names for things.

2 According to Bizer *et al.,* (2009) 'In computing, a Uniform Resource Identifier (URI) is a string of characters used to identify a name or a resource on the Internet. Such identification enables interaction with representations of the resource over a network (typically the World_Wide_Web) using specific protocols. Schemes specifying a concrete syntax and associated protocols define each URI.'

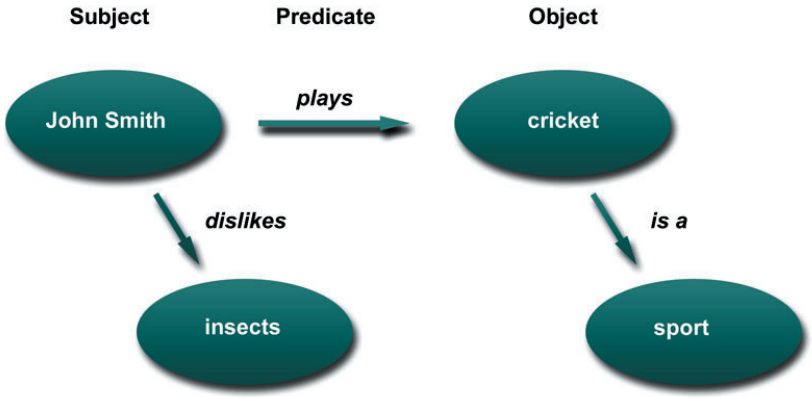


Figure 2. An illustration of RDF triplets (Bradley, 2011).

2. Use of HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).
4. Include links to other URIs, so that they can discover more things.

Linked data is also a part of the semantic web, another term coined by Tim Berners- Lee. He defines the semantic web as 'a web of data that can be processed directly and indirectly by machines' (Berners-Lee, 2000). The semantic web is therefore made up of linked data (Berners-Lee, 1998). The Food and Agricultural Organisation (FAO) is one IGO championing the use of linked data, although many projects seem to have not yet entered the mainstream. Open government data is providing great impetus for the linked data movement.

How does open data, real-time data and linked data connect?

How do these data models relate to the use of data for development? If governments, IGOs and other institutions, commercial companies, and individuals make quantitative and qualitative data available in a form which allows open data to be linked to real-time data from multiple sources, they can provide information about what is needed in different fields for development. The technology exists for this to happen, and in a world which seems to be heralding the development of personalised computing, the public increasingly needs to adapt to the demands of today's world by understanding the data presented and using the developing technology at hand in a way which can help them and their communities. However, although this is developing, as of yet there is no overarching structure that exists to monitor and help solve data protection and security issues that will arise with the use of such data.

Why pursue the use of this data?

From time immemorial, information on public, private, and charitable services, along with geographic and demographic information and other related information, have been important sources of data. As far back as 1854, John Snow was able to prove that cholera was a waterborne disease by using a map to show deaths in relation to water pumps.[3] Micro and macro datasets are both valuable. Micro datasets may be specific to a community, and the results of its analysis may not be applicable at a national or international level. It may, however, provide very useful case studies for groups facing similar issues elsewhere. Macro datasets apply to a much bigger group.

Individuals may not be aware that they need data until a circumstance arises where they look for it and cannot find it. The Internet has made those

3 'At the time, London was supplied its water by two water companies. One of these companies pulled its water out of the Thames River upstream of the main city while the second pulled its water from the river downstream from the city. A higher concentration of Cholera was found in the region of town supplied by the water company that drew its water from the downstream location. Water from this source could have been contaminated by the city's sewage… After the panic-stricken officials followed Snow's advice to remove the handle of the Broad Street Pump that supplied the water to this neighborhood, the epidemic was contained' (Crosier, 2009).

with easy access to it, take for granted that information can be just a search engine away; however, it is important to remember that many do not have access to the Internet and other ICT services to provide them with the data they need. Libraries were built before most people knew how to read, so people from all walks of life should have data at their disposal, for this is a primary way to encourage them to engage with and utilise it. According to blogger Rolf Kleef:

'We stimulate literacy by offering the stakeholders ways to get open data, publish their own, experiment with applications, and learn from each other. And we circumvent the tendency towards over-standardisation by doing this in parallel with other initiatives with sometimes overlapping goals and often different agendas' (Kleef, 2010).

On the other hand, offline activities can be used to gather data which can be utilised online. This should be explored, where technology is not easily accessible and it must be adapted to communities' needs.

IGOs use of open, real-time and linked data

IGOs, like the Organization for Economic Cooperation and Development (OECD) and the World Bank, are providing information on overseas development assistance and are intent on making many more diverse datasets available. The World Bank, one of the largest sources of development information (Zoellick, 2010), launched data.worldbank.org in April 2010 to provide information on development data and has been reaching out to civil society organisations, technologists, other development agencies, and the public to help organise and make sense of the vast sets of development data. To date, the value of the information provided in such a short time has been recognised as valuable, and in some quarters the World Bank is also acknowledged as a sort of data bank.

The World Bank, in collaboration with the OECD, launched Aid Flows in 2010 to make global data on aid funding more easily accessible. This demonstrates how much development aid is provided and received around the world and offers more transparency about the flow of development funds from countries providing aid to countries receiving the aid (The World Bank Ground and OECD, 2011). The World Bank is answering the call for transparency and open data and is making much more of the information it collects on countries freely available in the hope that resources can be better allocated.

AidData, a web portal, (a Development Gateway programme) was launched in March 2010 and provides information on aid statistics from the OECD and other aid sources. AidData recently teamed up with the World Bank Institute to complete the first stage of a groundbreaking Geocoding project, Mapping for Results Initiative. Working collaboratively, a team of 13 interns successfully geocoded 1216 World Bank projects in 7 weeks, despite the odds (Stern, 2010).

Robert B. Zoellick, President of the World Bank, is reported to have said, 'researchers and other professionals in the field should be able to examine the data and assumptions behind World Bank reports, and interrogate this information to draw their own conclusions' (cited in Fahey, 2010). As the World Bank is making more data available and is reaching out to various stakeholders, they are increasingly being recognised as a dominant player in open data for development landscapes.

The United Nations (UN) is also falling in line with the open/real-time data trend. It has made much of its statistical information freely available online on UNdata, a database that was launched in February 2008. Furthermore, the World Food Programme (WFP) is collaborating with other UN partners and governments to collect food security baseline data using mobile phones.They are trialing FrontlineSMS and RapidSMS, two mobile data collection software tools as well as using custom Personal Digital Assistant (PDA) to collect data (Verclas, 2011).

In light of the rapidly changing technology environment, and the need for the UN system to adapt, the UN launched UN Global Pulse in January 2010, which seeks to move beyond the provision of statistical information and gather real-time information to help understand what is

needed for development in today's ICT age. The Global Pulse monitoring system, which will be implemented in 2011, aims to help better track the impact of crises on the most marginalised and vulnerable in the world. This initiative has the potential to be revolutionary for development efforts, and, if done collaboratively, may be able to also address privacy and security concerns.

The UN's FAO is one IGO that has already acknowledged the potential of linked data for development of food and agriculture industries. They developed an Agrovoc thesaurus, which uses URIs with the RDF concept, and which contains approximately 40000 concepts in over 20 languages on various food and agriculture-related themes (FAO, 2010a)[4]. They have also introduced an AGRIS search engine, which includes almost three million structured bibliographical records on agricultural science and technology (FAO, 2011). This database provides information dating back to 1975 and is freely accessible online. AgroTagger, which uses Agrovoc as a controlled vocabulary, and a Drupal platform is also being developed to manage all types of information resources (FAO, 2010b). The linked data concept is driving these projects.

Ushahadi has led to other African initiatives. For example, JamiiX, a cloud computing platform that was recently launched in South Africa and now uses various channels to distribute information throughout the country. Following several earthquakes, the World Health Organisation (WHO) deployed JamiiX in Indonesia to aid communications.[5] Such platforms have provided evidence of the value of such data and have led to calls for policy research and development to include not only government data but also crowdsourced and real-time data in order to help everyone contribute to solving the development dilemmas confronted.

4 According to FAO, AGROVOC contains close to 40 000 concepts in over 20 languages, covering subject fields in agriculture, forestry, and fishery, together with cross-cutting themes, such as land use, rural livelihoods, and food security.

5 According to Wmworia (2011), 'While Ushahidi uses social media and mobile communications very cleverly to collect information from the people at the scene of significant events, JamiiX uses the channels to share vital information and advice with people on the ground.'

Mutually beneficial partnerships are being forged with Google, the heavyweight of online data, to help each party access and understand data. Together with the United Nations Economic Commission for Africa (UNECA), the Google Map Maker team introduced a training programme in Africa to help strengthen the national capacity in making statistics more accessible, useful, and interesting to the wider African public. This was done by linking technologies like Google Map Maker, mobile applications, Google Earth and Maps, Google fusion tables, and Public Data Explorer (Graham, 2011) to support data collection, sharing, visualisation, and publishing (Google, 2010). Google also utilised data from the World Bank to display graphs for queries on 17 world development indicators through the use of the World Bank's public API and the introduction of a World Bank data visualiser.

Zoellick has spoken about the 'need to democratize and demystify development economics' (Zoellick, 2010). He emphasised the World Bank's intention to open up its datasets to contribute to the *'open research revolution'* as no one institution has all the answers. He added:

'We need to recognize that development knowledge is no longer the sole province of the researcher, the scholar, or the ivory tower. It's about the health-care worker in Chiapas recording her results; it's about the local official posting the school budget on the classroom door so that parents can complain when their children are shortchanged; it's about the Minister, the academician, the statistician, and the entrepreneur comparing notes on the impact of incentives' (Zoellick, 2010).

Furthermore, UN Global Pulse acknowledges that 'designing a system to bring together the people, tools, and information needed to ensure achievement of the Millennium Development Goals can only be accomplished by harnessing grassroots innovation, particularly in the Global South' (Global Pulse, 2010). It follows that only inclusion of all data will have the potential to revolutionise the way development projects are run for the better from the community to the international level.

A lack of regulatory structures for data use

Civil society organisations that need data and also have access to data are increasingly collaborating with IGOs and the private sector. In some cases, they are attempting to provide a regulatory framework and overarching structure where none exists at the IGO and industry levels. For example, OKN is a non-profit organisation (NGO) based in the UK that is engaged in a wide range of open data activities. They provide tools and infrastructure as they work on a wide variety of initiatives, such as CKAN, and Open Data Commons, among many others, that are starting to form common standards for accumulation and distribution of data, where none previously existed (OKN, no date c). OKN provides an outlet for those interested in data issues and involved in the civil society sector, technology sector, or governmental sector to interact. CKAN, which is an open source system, currently used in data.gov.uk and in over 20 catalogues around the world, is 'connecting different catalogues to more easily pull data together from lots of different countries with a few clicks of the mouse' (Pollock, 2010).

Creative Commons is also a non-profit organisation that provides licenses to help 'provide a free, public, and standardized infrastructure that creates a balance between the reality of the Internet and the reality of copyright laws' (Creative Commons, 2011) in the absence of a regulatory infrastructure.

Where is non-IGO data coming from and how is it being generated?: The need for multiculturalism

Ways can be found to make data relevant to groups and communities that need it. One of the key ways is through engaging and educating communities about the value and potential uses of, not only information, but of raw data as well. This could help in gathering data needed to help support communities and citizens, and the resulting benefits can rebound to the rest of society. There are many examples that illustrate how open data is being used to find out what is needed in a region or community, although many of these initiatives are concentrated in Europe and the USA. For instance, in Nantes, France, the mobilisation of local residents and the local government combined with education and activism is facilitating the opening up of the city's data to aid its development. This followed the release of data by Rennes, in 2010, because of private activism and in Paris as a result of local government support. Local open data policies are now emerging in different cities, such as Besancon, Toulouse, Lille, and Dijon (Libertic, 2010).

This illustrates a knock-on effect. Once certain types of data are available in one place, others acknowledge an information deficit and want more information available to them. It follows that what happens in one region or place is copied and adapted elsewhere, which is building not only micro datasets but regional and national datasets. These may be of value on their own, as well as when mashed up with data elsewhere.

Many open data days and events are also being championed around the world. Berlin is hosting open data days (Open Data Berlin, 2011), The World Bank launched an Apps for Development competition, and other Apps projects have been introduced, which aim to link technologists with other stakeholders. For instance, the Apps for Amsterdam project recently took off, following in the footsteps of another Dutch city, Rotterdam, which was inspired by Apps for Democracy in Washington DC (Dammers, 2011). The Dutch government also launched an open data business initiative to help connect entrepreneurs who need data with programmers. The Italian city of Turin launched an open data competition (Crea, 2011), while Leida city in Spain is also opening up its data to the public (Leida, 2011). Open data initiatives are spreading throughout Europe at the micro dataset level and the macro dataset level.

IGOs, governments, and other stakeholders are revealing information they want to make public, so more citizen engagement is necessary to get data that is needed locally, and, in this globalised world, internationally. This, however, must be done in a way which can still protect each individual, despite the odds. Data should be generated not only from the top down but from the bottom up, and infrastructure must be

in place to ensure that it is not used improperly. However, for people to feel secure enough to provide information there needs to be more regulations that can protect them. Some argue that regulation will restrict the provision of necessary data, although it can also be argued that the reverse is true. A lack of regulation can reduce confidence in the Internet and in the integrity of the Internet, two vital ingredients that are necessary for generating voluntary data, as well as increasing the likelihood that data being used for development is as valid and reliable as can be.

Individuals make democratic decisions based on data that is available to them. The public also values data provided, not only by the government and other stakeholders, but also by other ordinary members of the general public and their communities. Therefore, access to information from the most marginalised in society and from the general public is of vital importance for inclusive governance. There is a dominance of the open data debate in the western world. In some countries in other parts of the world, open access to information is not necessarily the norm since many infrastructural and social constraints exist, which influence the extent to which data is available to them, the extent to which the public are aware of the available data, and whether or not they can use it.

Gurstein (2011) has written about what an individual needs in order to access data (Box 3).

Also of importance within an open data and linked data model are ontologies and thesauri, which is something that all data protagonists including those working in IGOs must pay special attention to. Wallack and Srinivasan (2009) highlighted how mismatched ontology hindered an Indian community's ability to provide information to the Indian Karnataka State, as well as thwarted the state's ability to understand the community's needs. According to the authors,

'Bus stop conversations about water logging, and reported complaints about pipes and drains refer to the same occurrence. Yet the commu-

Box 3. What is needed for an individual to access data

1. Internet — having an available telecommunications/Internet access service infrastructure sufficient to support making the data available to all users. Issues here would include:

I. The affordability of Internet access: a major issue for many particularly in the developing world.

II. The availability of sufficient bandwidth for the range of uses to which the data might be effectively put, for example, whether the data access has been designed on the basis that for example, broadband is necessary for the use of the data being made available.

III. The accessibility of the network, for example, where access to the network or to connectivity is restricted for political or other reasons.

IV. Physical accessibility/usability of access sites as for example for the physically disabled.

2. Computers and software — having access to machines/computers/software to access and process the available data and machines that are sufficiently powerful to do various analyses; having sufficient time on the equipment to do the analyses (many people need to share computers); knowledge of how to operate the equipment sufficient to access and analyze the data and so on. Does the use of the data require more powerful (and expensive) computers or software than might be generally available, for example?

3. Computer/software skills — having sufficient knowledge/skill to use the software required for the analyses/making the mashups/doing crosstabs, etc. Techies know how to do visualization, university trained persons and professionals know how to use the analytical software but ordinary community people might not know how to do either and getting that expertise/support might be either difficult or expensive or both.

4. Content and formatting — having the data available in a format (language, coding for display, appropriate geo–coding) to allow for effective use at a variety of levels of linguistic and computer literacy. What are the language, computer literacy, data analytic literacy levels that are required for an effective use of the 'open data'? Does the use of the data presume that it is being used by a professional and are there means through which those professionals might be available to those who can't afford expensive fees?

5. Interpretation/Sense making — sufficient knowledge and skill to see what data uses make sense (and which don't) and to add local value (interpretation and sense making); being able to identify the worthwhile information and to figure out how to put the data into the right format or context so that what might otherwise be numbers on a page becomes something that can change people's lives.

6. Advocacy — having supportive individual or community resources sufficient for translating data into activities for local benefit. Availability of skills and local resources, community infrastructures, training, the means for advocacy and representation all are required to enable effective local interventions based on the open or other data.

7. Governance — the required financing, legal, regulatory or policy regime, required to enable the use to which the data would be put.

nity's understanding of the situation and the Karnataka government record of the event label, catalogue and interpret the event in divergent manners (Wallack and Srinivasan, 2009).

This, along with language translation issues, needs to be understood within any overarching data standardized system.

Research undertaken by Becky Hogge (2010) and commissioned under the umbrella of the Transparency and Accountability Initiative explored the feasibility of applying the UK and US government approach to open data in several middle income and developing countries. Hogge's research revealed that three groups of actors were crucial to the projects' success:

- Civil society, and in particular a small and motivated group of 'civic hackers';
- An engaged and well-resourced 'middle layer' of skilled government bureaucrats; and
- A top-level mandate, motivated by either an outside force (in the case of the UK) or a refreshed political administration hungry for change (in the US).

As Tim Berners-Lee observed 'It has to start at the top, it has to start in the middle and it has to start at the bottom'. (Hogge, 2010) The conclusion to this report strengthens that assertion, arguing that developing countries should not copy UK and US initiatives but that they must pay attention to each of the three spheres of influence. One can add that in different counties there may also be additional groups that may be equally influential as well.

This study lends support to the argument that a multistakeholder approach is essential to not only gathering data, but also to promoting its reuse and educating the public about its benefits. This study also highlights the importance of legislation in ensuring that information is made available by the government. In countries without enforceable freedom of information laws, getting data can prove much more challenging. This raises the issue of whether IGOs can play a role in helping to ensure that a certain level of information is made available to the public.

The need for representation in the use of real-time data is also vital, for failure to have an inclusive dataset will result in skewed results. This could be detrimental for development efforts. As initiatives that utilise ICT and real-time data, including data gained from social media sites, are developed to help measure happiness,[6] flu outbreaks,[7] political sensitivities, and the needs of communities, it must be understood that without inclusive representation, findings may not be applicable to all groups in all societies. They may ignore those that do not have access to ICT or those that do not use the social networking sites.

A 2005 Ugandan community-based case study, which predates what is currently being achieved in Nantes in 2011, examined how access to data could improve communities' access to services. Ritva Reinikka and Jakob Svensson launched a newspaper campaign in Uganda to decrease the amount of money being rerouted from education. Armed with data, parents monitored how funds from a large educational grant were spent. They found that only 20% of the grant was given to schools. National newspapers publicised this, and community members monitored amounts received locally. This community-based approach meant that members of the community held their local authorities to account for their actions and as a result increased the amount of grant money that schools received by 60%, thereby improving educational services (Beech, 2011).

Data and information is valuable to everyone from all countries and communities, including those that do not have access to the technology that is being used to gather data. Novel approaches can be devised offline, which can be then utilised online, as more information is digitised and e-government is increasingly prioritised. The agriculture and rural development sector has acknowledged the importance of mobile phone technologies to enhance the flow of information to farmers and enable them to access more accurate and timely information on market prices at different stages in the value chain. This

6 Mappiness is an iPhone application which enables users to provide information about their feelings throughout the day.
7 Google Flu Trends uses aggregated Google search data to estimate flu activity.

improves their market position and their profitability. The importance of mobile technology for this open, real-time, and linked data movement will prove crucial in many developing countries that still lack reliable access to broadband.

Privacy and security – can opening up data be a Pandora's box?

Previous discussions all lead to an issue that is often ignored when highlighting the benefits of data. This is the issue of privacy and security. There may be no privacy and security issue with open government data, as the laws of the country place this information in the public domain, and institutions can choose whether to make their data open. Instead, these governments and IGOs are encouraging others to reuse it. However, if data that was not freely provided by an individual is made available, this data may be difficult to contain, and it may become part of the public domain or be used by third parties without express authorisation.

Furthermore, many remain ignorant of their data rights, decide to pay little attention to the fine print, and freely give up information in exchange for a service without due consideration of the risks. This also makes implementing formulating and implementing policies somewhat challenging.

'Some consumers are troubled by the collection and sharing of their information. Others have no idea that any of this information collection and sharing is taking place. Still others may be aware of this collection and use of their personal information but view it as a worthwhile trade-off for innovative products and services, convenience, and personalization. And some consumers – some teens for example – may be aware of the sharing that takes place, but may not appreciate the risks it poses. In addition, consumers' level of comfort might depend on the context and amount of sharing that is occurring.' (Federal Trade Commission, 2010)

This is part of the tangled web of data being weaved in today's technological world, for many of these services have become not only luxuries, but necessities, and an increasing amount of data, including real-time data, is being linked with other open datasets to provide more insightful and holistic pictures for multiple stakeholders.

Complaints about privacy usually feature Google and social media sites like Facebook, but as more and more data is generated and utilised by individuals and other third parties, many will use it in a way which invades the user and non-user's privacy and threatens their security. Numerous reports of computer hacking only serve to emphasise that security and privacy merits attention in any discussion about the openness of data. The Internet and Smartphone users usually have no idea who controls and owns their data, how it is used, or what their privacy rights are. Many times, it is not the individual consumer that owns their own data, but the business, government, or ISP. A recent survey by Opera Software revealed that consumers in the USA, Japan, and Russia are more concerned about Internet privacy than terrorist attacks, home attacks, or bankruptcy, yet they still do not do enough to protect their online privacy and security (Opera, 2011).

Today, privacy and security protection is in large part driven by businesses, although national governments also have their own respective privacy legislation, which tends to play catch-up with technology. Large private companies invest in privacy and security protection because if consumers are not confident that their information will be safe, they will be less likely to utilise a service. However, this may not necessarily apply to 'free' services, as there is no guarantee that data generated from the use of social media sites will remain private or even solely the property of the free service provider.

Governments also request information about users without informing them. For instance, journalist Shi Tao claimed that Yahoo provided the Chinese government with information about him and other dissidents which led to their arrest (PEN, no date), Twitter faced a US government subpoena for information on a number of individuals connected to Wikileaks (O'Dell, 2011), and Google's revelation of the number of times Google data has been requested by different governments reveals the extent to which data provided in exchange for services is not private

at all (Google, 2011b). Furthermore, if medical data, for instance, is shared anonymously with researchers online to aid the scientific process, there is also no guarantee that the information will remain confidential (Carnegie Melon, 2006). For open and real-time data to be linked, shared, contributed to, and appreciated, privacy and security issues must be addressed.

It is also necessary to examine geomapping and other geolocation technology that are integrated within the linked data model. Google street view caused a furor, as visual information about a person's residence could be found on its maps (Van Camp, 2010). Furthermore, geolocation data that is integrated with other data can sometimes make it much easier to identify a person or a placeThe extent to which this can jeopardise the privacy and security of those who can be identified via such easily available maps also merits attention.

However, it is not just large technology companies that spy on others for it has become much easier for others to gather data online. Cost effective yet powerful and sophisticated gadgets and applications can now be easily purchased to gather personal data and spy on others online (Henn, 2011). According to Henn,

'The way the high-tech economy works, with gadgets getting cheaper even as they're getting more sophisticated, maybe it's not Big Brother we should be worrying about at all. If the prospect of pretty much anybody being able to get their hands on the electronic version of you is an issue, Little Brother may be more of a problem.'

More must therefore be done to protect data privacy as data is used not only for development but also placed on cyber data black markets, with negative consequences for the uninformed user. When databases are hacked and private data is misused, the said user may be unlikely to want to provide data of any kind for development efforts as confidence in the Internet wanes.

Crowdsourced data also goes to those who crowdsource it. There is no common place where it resides, so it can either be productively used or abused by those that possess it. This is a situation that often goes ignored as individuals engage in 'slacktivism' or other activities online. Nadim Mahmud of Medic Mobile, argues that there is no efficient system to manage online crowd data and that those organisations that try to implement such a system are met with apprehension and rivalry (cited in Lloyd, 2011). This only serves to support the need for an overarching body that can help to ensure that the information and data is used responsibly, as well as in a way that provides maximum benefits to individuals. As IGOs attempt to follow in the footsteps of civil society organisations and mash up their data with crowdsourced information, they will need to examine these issues.

Given that the industry is led by commercial interests, it is important to have other stakeholders involved who can ensure that some level of oversight and standards are maintained. As Swire and Litan (1998) put it:

'Consider the incentives of a company that acquires private information. The company gains the full benefit of using the information in its own marketing efforts or in the fee it receives when it sells the information to third parties. The company does not, however, suffer losses from the disclosure of private information. Because customers often will not learn of the overdisclosure, they may not be able to discipline the company effectively. In economic terms, the company internalizes the gains from using the information but can externalize some of the losses and so has a systematic incentive to overuse it.

This market failure is made worse by the costs of bargaining for the desired level of privacy. It can be daunting for an individual consumer to bargain with a distant Internet merchant … about a desired level of privacy. To be successful, bargaining might take time, effort and considerable expertise in privacy issues' (cited in Samuelson, no date).

Although this may be true, it can be further explored to ensure that the use of data by the market place may not only help secure privacy and security but lead to a win-win situation for the ISP, company, consumers, and the wider public. Just as offline rules for privacy exist online, rules of communications and engagement must evolve which respect such rights too. Working with governments

and civil society organisations, IGOs can help to develop such international norms.

However, the ability to manage data privacy is further complicated by the international nature of the Web because different countries take different approaches to protecting consumer privacy. The 1998 EU's Directive on Data Privacy prohibits the transfer of personal data to non-EU countries that do not meet its guidelines for privacy protection. The directive also provides for the creation of government data protection agencies that will oversee the registration, and in some cases the approval, of databases containing personal information. If a company conducting business with EU citizens online does not comply, they are in breach of EU data protection laws (cited in Rappa, 2010). Most recently, the EU mandated that under the European e-privacy directive, businesses within the EU must limit the use of cookies for targeted behavioral advertising by 25 May 2011 (BBC, 2011). The use of government oversight, as well as demands for industry self-regulation, can provide a useful example of how Internet data privacy policies can be approached at the IGO level.

Moreover, digital natives are growing up in this new digital data landscape, and they too are providing information that can be utilised. However, in most countries, there is not enough safeguarding of data of the youth. 'In a world where authenticating who is at the other end of a web connection is never an absolute certainty' (Rappa, 2010) children must be protected. This is important and must be addressed at the international level. This is not only because data that youth innocently give today may be detrimental to them in the future, but also because in the same way that institutions have been set up in the past to assist with issues that transcend borders, some sort of international regulation and standardisation may prove vital in helping to protect the world's youth. Moreover, problematic issues arise when dealing with data on young people. Though such data is useful, security issues usually take precedent.

Privacy is also culturally specific. Information that may be deemed private in one country or culture may not be private in another. As a result, while information made available in a local context may be acceptable, it may prove detrimental if accessed by the rest of the world. The marginalised and those facing political prosecution often feature in the work of civil society organisations and UN organisations. Care needs to be taken to ensure that sensitive information on vulnerable individuals, such as the stateless and victims of abuse, are not made publicly available, as this could lead to victimisation, which runs counter to success.

The same holds true in today's social media world. According to Llyod (2011),

'Real-time data is now being collated to create dozens of tool types that go by seductive names such as human sensor networks, personal censuses, social-sentiment measurements, mobile communities, see-something/text-something, instant mapping, context cartography, timeline narratives, intelligent infrastructure and data democracy.'

Individual protests captured on video or Facebook can lead to prosecution by governments and rogue bodies, and there must be a way for them to be able to get data privacy and security if necessary. Gregory (2011) believes:

'Facebook's insistence that its users use their 'real identity' when signing up – and deleting accounts and groups that do not comply – makes it difficult for human rights activists needing to work anonymously or pseudonymously. And it makes it easier for governments to track not only individuals but also their networks.'

There are competing organisations, governments, and even private interests for the data. As Cablegate showed, once information is made available online, it is difficult to contain its dispersal (Infosthetics, 2010). Today's data-possessed world must also be aware of this and find ways to deal with the consequences.

The right for an individual to forget and delete information posted online, as well as the ways in which communication online can be done anonymously must also be explored, as should the effects of this on the availability and use of data. Today, anonymous online interaction 'focus[es] on options for encryption or

for using proxy server(s) and circumvention approaches like Tor to conceal both the person communicating and the data being transmitted' (Shane, 2011). However, as Mark Zuckerberg of Facebook claims, local and international social norms are moving away from an emphasis on privacy and anonymity, and there it is necessary for the data movement to seriously examine this phenomenon, rather than take it as a fact (cited in Horn, 2010). Indeed, even if this phenomenon is a fact, it reveals a lot of problems which may require regulating to help solve them. As support for a 'property-rights approach to the information privacy problem' (Samuelson, no date) circulates, these and other scenarios can be examined at the international level.

The renewed emphasis on Cloud computing introduces another aspect to the data privacy issue. Inexpensive computing power, 3G mobile roll, improved Internet, and electricity infrastructure, is leading to an increase in the number of Cloud services. The Cloud relates not only to paid services, but to free services as well, like email, Google docs, Facebook, Twitter, and other social media sites. It has tremendous potential to contribute to development in developing countries. However, privacy and security concerns must also be addressed there. Individuals need to acknowledge that data provided online can be lost and that there are no guarantees that it will be kept far from prying eyes.

The EU recently released a warning to EU member states about the use of Cloud computing (Shane, 2011), emphasising that at the regional level it has been clearly acknowledged that data is not necessarily safe and secure in the cloud. There is as of yet no international standardised regulation of Cloud computing providers, although such information may be used in a linked data environment. A non-binding, informal Open Cloud Manifesto that aims to be 'dedicated to the belief that the Cloud should be open' has been developed, but this in no way provides any guarantee of privacy and security.

The use of personal technologies is increasing, such as GPS, mobile phones, and laptops all of which are equipped with sensors and collect geographically tagged data as people conduct daily routine activities. Such data is even referred to as 'information exhaust' (PSFK, 2011). For instance, The Copenhagen Wheel, a commercial initiative, intends to allow cyclists to capture the energy dispersed while cycling and use it later when they need it (The Copenhagen Wheel, no date). It will also map pollution levels, traffic congestion, and road conditions in real-time and is controlled through a smartphone.[8]

Other initiatives include Grassroots mapping, which utilises balloons and kites to help map areas affected by disasters. However, such information must only be used with permission, and users should be able to opt out whenever they wish. As more companies gather this type of information, and given that people may not always be aware of the way in which they are giving data, strict laws will be necessary to ensure that such sensitive information on a person's whereabouts and activities remain private.

Real-time data is available online, and many academic, governmental, and civil society organisations are making it available. Commercial companies, like Recorded Future, are using data to predict the future, while there are other initiatives, like Voyurl, which was recently launched in beta format and takes advantage of the voyeur's need to peer into others' activities by allowing individuals to share their click stream and 'find out who is looking at what, when and where' (Voyurl, 2011). Analytical tools are also being introduced to help understand integrated data to know what is needed to realise the millennium goals and what may be needed to address environmental and health concerns. Despite the indisputable fact that the data use can benefit communities and the world, as more ways of gathering and understanding data are developed for governmental, commercial, civil society, and private use, the importance of privacy and security within such an environment is becoming essential. The more data is made available the more valuable will be the data analy-

8 According to the Copenhagen Wheel website, 'A person will be able to use their phone to unlock and lock their bike, change gears and select how much the motor assists you. As a person cycles, the wheel's sensing unit is also capturing effort level and information about surroundings, including road conditions, carbon monoxide, NOx, noise, ambient temperature and relative humidity. The user can also access this data through their phone or the web and use it to plan healthier bike routes, to achieve your exercise goals or to meet up with friends on the go. Their data can be shared with friends, or with their city - anonymously if they wish.'

sis, which is why privacy concerns are sometimes pushed aside when discussing the use of data for development. However, some balance between data use and privacy must be reached and IGOs in collaboration with other stakeholders can help to ensure that such a balance is achieved.

Conclusion

This paper examined the concepts of linked, open, and real-time data, as well as the privacy and security constraints that may affect their extensive usage. It finds that, while there is a lot of potential for the use and linkage of such data for development, privacy and security issues have still largely been ignored. However, as the use of such data grows, it will become of paramount importance. Today, civil society organisations, like the OKF, are playing a big role in trying to reach some type of standardisation of data use for development, while other IGOs, like the World Bank, OECD, and the UN, have already recognised the importance of the data use for development, as they collaborate and encourage the linking of their data with other datasets. They are reaching out to each other, civil society organisations, private companies, and technologists to understand the data and make it more widely available to the public.

Certain important issues are emerging within this data movement, which may at some point threaten its growth and ultimate success. These relate to issues of privacy and security. Though some social network proponents may argue that privacy is dead, this should not thwart an investigation of the ways in which privacy and security can be protected within an open, real-time, linked data movement. The marginalised and the vulnerable will be some of the most affected, and they must be helped in order for development effectiveness to become a reality. There needs to be a better examination of how technology can be adapted to ensure that data privacy and security can be protected. Data and information terrorism may indeed be, in today's technological world, more potent than bombs.

The issue, however, of who should manage such a linked data model arises. Civil society organisations and the technology community have been leading the way in innovative use of data for development, and governments and IGOs are now collaboratively following in their footsteps. Though each stakeholder has a part to play, international entities may be able to serve an important role in ensuring that data is used for development and is not manipulated by those in power. The use of data must be inclusive of the most marginalised and vulnerable, but it must also protect their rights to privacy and security. While acknowledging the difficulty of international Internet governance, as well as the bureaucratic constraints that lie within international organizations, a code, as law,[9] or 'privacy by design'[10] may prove useful when combined with IGO regulation. It may help address not only data for development issues, but also those related to privacy and security. Furthermore, the EU Protection Directive also provides some insight into how an intergovernmental organisation can protect privacy, and this may aid understanding of how such an international organisation can simultaneously use data for development. There will be a need for more collaboration and exchange of ideas from the bottom, middle, and the top, as well as from various stakeholders, to ensure that this is achieved.

Today, to a large extent, as Llyod (2011) put forward, 'Developers of such (data) applications and services must be stewards of that data … You can't just make this stuff open… there are a number of competing factions in the open source world. There should be a collaborative dynamic. There isn't always one.' Open and real-time data are increasingly being used to help communities, nations, and regions. It is only a matter of time before privacy and security issues take centre stage. As linked data is highlighted as a way forward to help with the validation of data, so too will security and privacy also help to ensure validity of the results

9 In this case the code is 'the software and hardware that make cyberspace as it is. This code, or architecture, sets the terms on which life in cyberspace is experienced. It determines how easy it is to protect privacy, or how easy it is to censor speech. It determines whether access to information is general or whether information is zoned. It affects who sees what, or what is monitored' (Lessing, 2000).

10 Privacy by Design ensures that "privacy and data protection compliance is designed into systems holding information right from the start, rather than being bolted on afterwards or ignored" (Information Commissioners Office UK, (no date).

that are revealed from its analysis. Undoubtedly, more study needs to be done on the way in which this could be addressed as development paradigms shift and open data increasingly helps to influence related changes However, existing initiatives may provide clues that can help individuals and institutions to both use data and provide data, confident in the knowledge that their right to privacy and security will be protected.

References

1. Apps for Amsterdam (2011) Available at http://www.appsforamsterdam.nl/ [accessed 25 February 2011].
2. *BBC* (2011) New net rules set to make cookies crumble. 8 March. Available at http://www.bbc.co.uk/news/technology-12668552 [accessed 25 February 2011].
3. Beech A (2011) Open data, and how it can improve development outcomes. *Devex*,8 February. Available at http://www.devex.com/en/blogs/full-disclosure/open-data-and-how-it-can-improve-development-outcomes [accessed 25 February 2011].
4. Berners-Lee T (1998) Semantic Web Road Map. Available at http://www.w3.org/DesignIssues/Semantic.html [accessed 25 February 2011].
5. Berners-Lee T (2000) *Weaving the Web: The Past, Present and Future of the World Wide Web by its Inventor.* London: Texere.
6. Bizer C *et al.*, (2009) Linked Data - The Story So Far. Preprint of paper in Heath T *et al.*, [eds.] Special Issue on Linked Data. *International Journal on Semantic Web and Information Systems*, January. Available at http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf [accessed 20 November 2010].
7. Bradley A (2011) SEO, the semantic web and information discovery. *SEO skeptic* blog, 6 January. Available at http://www.seoskeptic.com/seo-semantic-web/ [accessed 25 February 2011].
8. Carnegie Mellon Data Privacy Lab (2006) Available at http://privacy.cs.cmu.edu/ [accessed 25 February 2011].
9. Crea A (2011) Open Data Contest: un concorso per democratizzare i dati della Pubblica Amministrazione. *PNTuner.net*, 16 February. Available at http://www.pctuner.net/blogwp/2011/02/open-data-contest-un-concorso-per-democratizzare-i-dati-della-pubblica-amministrazione/?utm_source=twitterfeed&utm_medium=twitter [accessed 25 February 2011].
10. Creative Commons (2011) Available at http://creativecommons.org/ [accessed 25 February 2011].
11. Crosier S (2009) John Snow: The London Cholera Epidemic of 1854, Center for Spatially Integrated Social Science. Available at http://www.csiss.org/classics/content/8 [accessed 7 March 2011.
12. CTIC (2011) *Open data*. Available at http://datos.fundacionctic.org/sandbox/catalog/faceted/ [accessed 25 February 2011].
13. Cyganiak R and Jentzsch A (2010) *Linking open data cloud diagram*. Available at http://richard.cyganiak.de/2007/10/lod/ [accessed 5 March 2011].
14. Dammers W (2011) Amsterdam stimuleert gebruik van open data. *SOLV*, 16 February. Available at http://www.solv.nl/weblog/amsterdam-stimuleert-gebruik-van-open-data/17843 [accessed 25 February 2011].
15. Fahey R (2010) Democratizing Development Economics through Open Data. *Talkin' bout a revolution: Change we can believe in* blog, 10 April. Available at http://www.rfahey.org/2010/10/04/democratizing-development-economics-through-open-data/ [accessed 25 February 2011].
16. FAO (2011) AGRIS: International information system for the agricultural sciences and technology. Available at http://agris.fao.org/ [accessed 25 February 2011].
17. FAO (2010a) Agricultural Information Management Standards: AGROVOC. Available at http://aims.fao.org/website/AGROVOC-Thesaurus/sub [accessed 25 February 2011].
18. FAO (2010b) Status of the AgroTagger service and its integration into Drupal and DSpace, Available at http://aims.fao.org/news/status-agrotagger-service-and-its-integration-drupal-and-dspace Agriculture Information Management Standards (AIMs), FAO, 2010 [accessed 5 March 2011].
19. Global Pulse (2010) Strategy and Roadmap. Available at http://www.unglobalpulse.org/strategy-and-roadmap [accessed 25 February 2011].
20. Google (2010) Helping to unveil the power of statistical data. *Google Africa blog* blog, 17 June. Available at http://google-africa.blogspot.com/2010/06/helping-to-unveil-power-of-statistical.html [accessed 25 February 2011].
21. Google (2011a) Flu Trends. Available at http://www.google.org/flutrends/ [accessed 5 March 2011].
22. Google (2011b) Transparency Report. Available at http://www.google.com/transparencyreport/governmentrequests/ [accessed 25 February 2011].
23. Graham N (2011) Future of Real-Time: Data Democracy. *PSFK* 3 February. Available at http://www.psfk.com/2011/02/future-of-real-time-data-democracy.html [accessed 23 October 2010].
24. Grassroots mapping (no date) *About*. Available at http://grassrootsmapping.org/about/ [accessed 5 March 2011].
25. Gregory S (2011) Human Rights Video, Privacy and Visual Anonymity in the Facebook Age. *Witness*, 16 February. Available at http://blog.witness.org/2011/02/human-rights-video-privacy-and-visual-anonymity-in-the-facebook-age/ [accessed

5 March 2011].
26. Gurstein M (2011) Open Data: Empowering the empowered or effective data use by everyone? *First Monday* 16(2). Available at http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/3316/2764 [accessed 25 February 2011].
27. Heath T and Bizer C (2011) *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1. Available at http://linkeddatabook.com/editions/1.0/ [accessed 20 November 2010].
28. Henn S (2011), Little Brother is watching, Marketplace, Wednesday, April 27, 2011 http://marketplace.publicradio.org/display/web/2011/04/27/pm-little-brother-is-watching/ [accessed 9 May 2011].
29. Higher Level.nl (2011) Een nieuw board: conversaties over het nieuwe overheidsinitiatief OpenData, 1st February. Available at http://www.higherlevel.nl/forum/index.php?board=59;action=display;threadid=35459 [accessed 25 February 2011].
30. Hogge B (2010) Open Data Study. Commissioned by the Transparency and Accountability Initiative. Available at http://unpan1.un.org/intradoc/groups/public/documents/un-dpadm/unpan042939.pdf [accessed 28 April, 2011].
31. Horn H (2011) Facebook Loses Face Over Self-Serving Privacy Policy. *The Atlantic Wire,* 12 January. Available at http://www.theatlanticwire.com/opinions/view/opinion/Facebook-Loses-Face-Over-Self-Serving-Privacy-Policy-2164 [accessed 25 February 2011].
32. Infosthetics (2010) WikiLeaks Cable Gate: the Visualizations and the Infographics, 29 November 2010. Available at http://infosthetics.com/archives/2010/11/wikileaks_us_embassy_cables_the_visualizations.html [accessed 5 March 2011].
33. (Information Commissioners Office UK, (no date), Privacy by Design, http://www.ico.gov.uk/news/current_topics/privacy_by_design.aspx [accessed 9 May 2011].
34. Kleef R (2010) Let's build a 'Debian for Development Data'. *Rolf's blog* blog, 22 November. Available at http://www.drostan.org/en/2010/11/lets-build-a-%E2%80%9Cdebian-development-data%E2%80%9D [accessed 25 February 2011].
35. Leida (2011) L'Ajuntament de Lleida posa en marxa l'Open Data. 16 February. Available at http://www.paeria.es/cat/ajuntament/noticies.asp?Detall=True&IdNoticia=12636&Dia=16&Mes=2&Any=2011&IdDepartament=-1&TextCerca&Consulta=False&PaginaAnterior=%2Fcat%2Fajuntament%2Fnoticies.asp&Pagina=1 [accessed 25 February 2011].
36. Lessig L (2000) Code is Law: On liberty in cyberspace. *Harvard Magazine*. Available at http://harvardmagazine.com/2000/01/code-is-law.html [accessed 5 March 2011].
37. Libertic (2011) The Role of Citizens in Nantes Open Data, Nantes, (translation by Aurora Crane-Pearson) 16 February 2011, Available at http://www.epsiplus.net/news/news/the_role_of_citizens_in_nantes_open_data/ [accessed 25 February 2011].
38. LinkedData (no date) *Linked Data.* Available at http://linkeddata.org/ [accessed 25 February 2011].
39. Lloyd R (2011) Crowd-sourced data hold potential for positive change and human rights abuses. *Scientific American Observations* blog, 18 February. http://www.scientificamerican.com/blog/post.cfm?id=crowd-sourced-data-hold-potential-f-2011-02-18 [accessed 25 February 2011].
40. Local Government Improvement and Development (2011) *Open data*. Available at http://www.idea.gov.uk/idk/core/page.do?pageId=18562122 [accessed 25 February 2011].
41. Mappiness (2011) Available at http://www.mappiness.org.uk/ [accessed 5 March 2011].
42. Medic Mobile (2011) Available at http://medicmobile.org/ [accessed 25 February 2011].
43. Nelson A (2011) How Mapping, SMS Platforms Saved Lives in Haiti Earthquake. *Media Shift,* 11 January. Available at http://www.pbs.org/mediashift/2011/01/how-mapping-sms-platforms-saved-lives-in-haiti-earthquake011.html [accessed 25 February 2011].
44. O'Dell J (2011) Twitter Subpoenaed by U.S. Government for WikiLeaks Accounts. *Mashable*, 8 January. Available at http://mashable.com/2011/01/08/twitter-subpoenaed-by-u-s-government-for-wikileaks-accounts/ [accessed 25 February 2011].
45. Open Cloud Manifesto (no date) *Open cloud manifesto.* Available at http://www.opencloudmanifesto.org/index.htm [accessed 25 February 2011].
46. Open Data Berlin (2011) Save the date / Einladung zum Berlin Open Data Day 2011. *OpendataBerlin's blog* blog, 16 February. Available at http://networkedblogs.com/emkvQ [accessed 25 February 2011].
47. Open Knowledge Foundation [OKN] (no date a) *Open Definition provides criteria for openness in relation to data, content, and software services.*Available at http://www.opendefinition.org/ [accessed 25 February 2011].
48. Open Knowledge Foundation [OKN] (no date b) *Open Definition*. Available at http://www.opendefinition.org/okd/ [accessed 25 February 2011].
49. Open Knowledge Foundation [OKN] (no date c) Projects. Available at http://okfn.org/projects/ [accessed 28 April 2011].
50. Opera (2011) Who's watching you? Data Privacy Day survey reveals your fears online. Available at http://www.opera.com/press/releases/2011/01/28/ [accessed 5 March 2011].
51. Pen American Center [PEN] (no date) China: Shi Tao. Available at http://www.pen.org/view-

media.php/prmMID/764/prmID/172 [accessed 25 February 2011].
52. Pollock R (2010) Open government data goes global. Keynote at Open Government Data Camp, University of London Union, 18–19 November. Available at http://rufuspollock.org/2010/11/24/open-government-data-goes-global-ogdcamp-keynote/ [accessed 25 February 2011].
53. PSFK (2011) Future of Real-Time [video]. Available at http://www.slideshare.net/PSFK/psfk-presents-future-of-realtime?from=ss_embed [accessed 7 March 2011].
54. Rappa M (2010) Managing the digital enterprise: Data privacy. Available at http://digitalenterprise.org/privacy/privacy.html [accessed 25 February 2011].
55. Recorded Future (2011) Available at https://www.recordedfuture.com/ [accessed 5 March 2011].
56. Samuelson P (no date) Privacy as Intellectual Property. Available at http://people.ischool.berkeley.edu/~pam/papers/privasip_draft.pdf [accessed 29 April 2011].
57. Shane D (2011) Don't use cloud for sensitive data, EU warns members. *Information Age*, 18 January. Available at http://www.information-age.com/channels/data-centre-and-it-infrastructure/news/1514033/dont-use-cloud-for-sensitive-data-eu-warns-members.thtml [accessed 25 February 2011].
58. Sherman A (2011) How Open Data Initiatives Can Improve City Life. *Mashable,* 15 February. Available at http://mashable.com/2011/02/15/how-open-data-initiatives-can-improve-city-life/ [accessed 25 February 2011].
59. Stern A (2010) Mapping For Results. *AidData* blog, 8 August. Available at http://blog.aiddata.org/2010/08/mapping-for-results.html [accessed 5 March 2011].
60. The Copenhagen Wheel (no date) *About*. Available at http://senseable.mit.edu/copenhagenwheel/ [accessed 5 March 2011].
61. Tor (no date) Available at https://www.torproject.org/ [accessed 25 February 2011].
62. Ushahadi (2011) Available at http://www.ushahidi.com/ [accessed 25 February 2011].
63. Van Camp J (2010) Google Street View: more privacy, more problems. *Digital Trends* 26 October. Available at http://www.digitaltrends.com/computing/google-street-view-more-privacy-more-problems/ [accessed 25 February 2011].
64. Verclas K (2010) Mobile Data for Early Warning. *MobileActive.org blog* blog, 11 May. Available at http://mobileactive.org/mobile-data-early-warning [accessed 25 February 2011].
65. Voyurl (2011) Available at http://www.voyurl.com/welcome [accessed 5 March 2011].
66. Wallack J and Srinivasan TN (2009) Local-Global: Reconciling Mismatched Ontologies in Development Information Systems. Proceedings at the 42nd Hawaii International Conference on system science. Available at http://ekarine.org/wp-content/uploads/lib/LocalGlobal.pdf [accessed 28 April 2011].
67. Wikipedia (2011) *Freedom of information legislation*. Available at http://en.wikipedia.org/wiki/Freedom_of_information_legislation [accessed 25 February 2011].
68. Wmworia (2011), South Africa's JamiiX set to be the next Ushahidi.
69. Afrinnovator, 23 February. Available at http://afrinnovator.com/mobile/south-africa%E2%80%99s-jamii-set-to-be-the-next-ushahidi?utm_source=feedburner&utm_medium=twitter&utm_campaign=Feed%3A+Afrinnovatorcom+%28Afrinnovator.com%29 [accessed 5 March 2011].
70. The World Bank Group and The Organisation for Economic Co-operation and Development [OECD] (2011_*AidFlows*. Available at http://siteresources.worldbank.org/CFPEXT/Resources/299947-1266002444164/index.html [accessed 25 February 2011].
71. World Bank (2010) Data Visualizer. Available at http://devdata.worldbank.org/DataVisualizer/ [accessed 7 March 2011].
72. Zoellick R (2010) *Democratizing Development Economics*. Paper presented at The World Bank Group, Georgetown University, 29 September. Available at http://web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:22716997~pagePK:64257043~piPK:437376~theSitePK:4607,00.html [accessed 25 February 2011].