

Linked Open Government Data: Lessons from Data.gov.uk

Nigel Shadbolt, Kieron O'Hara, Tim Berners-Lee, Nicholas Gibbins, Hugh Glaser, Wendy Hall and m.c. schraefel, *University of Southampton*

Services require data. In a top-down political culture where the state is the service provider of first resort, the state becomes a powerful data monopoly, able to structure and homogenize the interactions between itself and its citizens. Such one-sided interactions are expensive and unresponsive

to citizens' needs. Even when governments have exposed service provision to market disciplines, they haven't succeeded in presenting data to citizens in innovative ways to create new value streams. Privatized service providers have preserved monopolies of service design and provision.

As technology has increased the power of data by facilitating linking and sharing, and political thinkers have embraced transparency and citizens' right to data, this top-down culture is being challenged. Many governments now release large quantities of data into the public domain, often free of charge and without administrative overhead. This allows citizen-centered service delivery and design and improves accountability of public services, leading to better public-service outcomes.

In the United Kingdom, transparency is focused on Data.gov.uk, the public data catalogue that points to thousands of datasets downloadable under a permissive open government license. The datasets are often

in comma-separated value (CSV) format or spreadsheets, but there is potential for increasing their utility by linking them using structured machine-processable formats. Resource Description Framework (RDF) is the format most integrated into current thinking about future generations of the Web, as its use of URIs allows data to be identified by reference and linked with other relevant data by subject, predicate, or object. The use of Semantic Web standards in open government data (OGD) was pioneered by Advanced Knowledge Technologies (AKT) in a precursor to the work described here, and was reported to the UK Parliament in 2007.¹

We refer to this vision as the linked-data Web (LDW). The LDW is already well populated through initiatives such as DBpedia, the DBLP Computer Science Bibliography, the *London Gazette*, the *New York Times*, and the Comprehensive Knowledge Archive Network (CKAN). The formalisms and infrastructure are appearing according

A project to extract value from open government data contributes to the population of the linked data Web with high-value data of good provenance.

to linked data principles set out by Tim Berners-Lee some time ago (www.w3.org/designissues/linkedata.html), but vital research issues still need to be addressed.

First, we need to understand how to build or reuse ontologies easily and appropriately for particular applications. Second, we want query methods that scale across the unbounded Web, not just within small islands of well-structured data. Third, we need visualization and browsing tools, and fourth, we need to populate the LDW to increase the network effects of large-scale linking. These objectives drive the fundamental research of the EnAKTing project (www.enakt.org), funded by the UK's Engineering and Physical Sciences Research Council (EPSRC).

OGD will make an important contribution to the LDW. Its quantity will help deliver the network effects expected from the LDW, its provenance is clearer than that of many other types of data, and it is often seen as high quality, trustworthy, and neutral.

Representing OGD in RDF and linking to other datasets presents important research challenges, including

- discovering appropriate datasets for applications,
- integrating OGD into the LDW,
- understanding the best join points for diverse datasets—that is, the points of reference the databases share, which are extremely valuable for linking—and
- building client applications to consume the data, including interfacing with real-world users.

In this article, we use the EnAKTing approach to develop an integrated account of how to bring OGD into the LDW. EnAKTing's focus is the LDW as a whole, but here we focus

on the population of the LDW with OGD from Data.gov.uk, looking in turn at these four issues.

Discovering and Migrating Data

The adoption of OGD for use in the LDW will depend on its availability, and a necessary first step into expanding the LDW with OGD is the data discovery process. There have been a number of services supporting the location of public sector information (PSI), including Data.gov in the US and Data.gov.uk in the UK. Tools research and development is permitting the translation of PSI datasets into RDF and the generation of RDFa (RDF with attributes) catalogs, while the UK government is exposing linked-data endpoints of available PSI for reuse.²

However, innovative uses of PSI transcend borders; meteorology or transport applications, for example, need data from more than one nation. The LDW will be an important mechanism for data convergence, as shown by the Open Knowledge Foundation's CKAN—a registry of open data available for public use with a common cataloging schema built on a few metadata terms—and by the European statistical service Eurostat, which has amalgamated thousands of datasets with their metadata for download from its website. However, there is no single facility for retrieving related resources from the portals of the various nations, or for searching intelligently across regional, national, and supranational sources.

EnAKTing has proposed an architecture—not yet fully implemented—for integrating PSI catalogs via the activities and components essential for discovery. Architectures of this type allow the presentation of catalogs in a standardized form, facilitating search and retrieval across resources.

The first phase of this architecture involves downloading and transforming catalogs with retrievable records into a common schema language format, whereas the second addresses semantic heterogeneity with schema matching and statistical analyses of ontology structures. Once common ontologies are in place, the search engine layer can be developed, allowing distributed querying and federated search and retrieval.

Initial work has tested this architecture, using approximately 7,000 records taken from Data.gov.uk, the US site Data.gov, and the Australian national PSI catalog. Records were converted from native format into RDF, each detailing some 14 to 25 metadata fields, and stored in an RDF triple-store.

The initial translation was intentionally minimal, reflecting the catalogs' original contents and preserving the underlying arrangement of data. This reveals the need for data normalization. For example, temporal data such as release or modification dates were not always represented with a universal standard. With thousands of ambiguous dates, classifiers need to be applied to the data before evaluation and comparison of the resources referred to in the catalogs are possible.

Integrating OGD into the Web of Linked Data

Once datasets have been discovered, they must still be integrated into the linked data cloud. An application we developed for EnAKTing provides an example of such integration, bringing together six government datasets, covering the work of individual MPs (members of Parliament) and of Parliament as a whole, crime, mortality and health statistics, and geographical data from the Ordnance Survey. This application lets users investigate

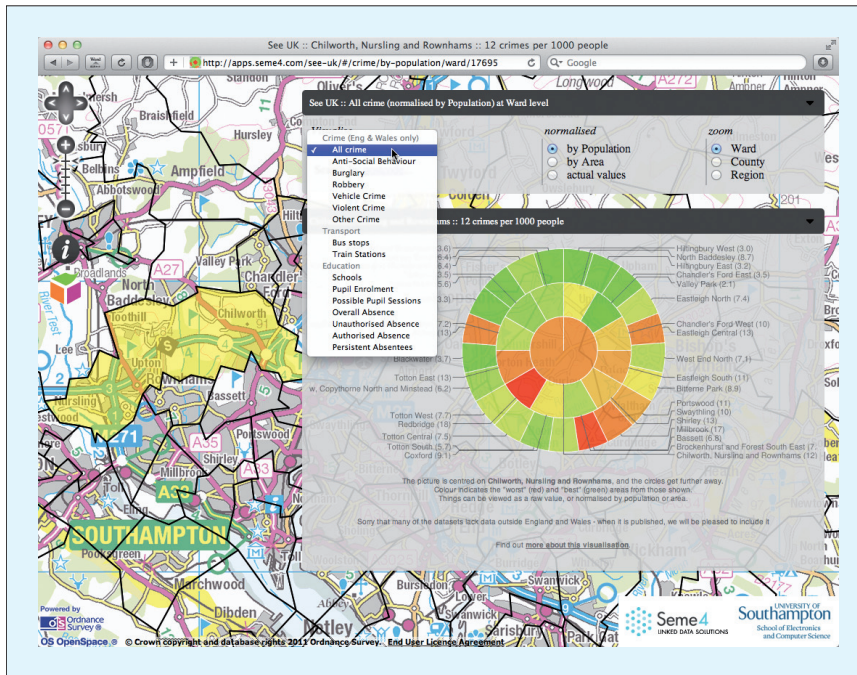


Figure 1. See UK, showing relative crime figures for a ward in Southamton. The pie chart shows comparisons between it and neighboring wards normalized by population, and the user can select figures and comparisons for particular classes of crime from the drop-down menu.

a particular geographical region.² Only the Ordnance Survey material was in RDF.

Publication and Consumption of the Datasets

Using well-known ontologies such as Dublin Core, Friend of a Friend (FOAF), and the Statistical Core Vocabulary (Scovo) eased the modeling overhead. Scripts were written to convert data from spreadsheets into RDF, and the Jena Semantic Web Framework was used to convert the HTML and XML, making data linkable without determining the semantics. For instance, data commonly contains terms that make perfect sense to experts in the field but are opaque to the rest of us; the health datasets used in our application included the codes *SHA Code* and *Org Code*, which can only be understood by someone *au fait* with UK National Health Service (NHS) administration. Such problems multiply across datasets, requiring an ontological alignment stage.

In our application, this involved the correct identification of `owl:sameAs` relations across a dimension linking the datasets. Administrative geography provided the link, via MPs' constituencies, NHS trusts, and so on. The alignment can be complex. For instance, to align the health statistics, we needed to use the Google Maps API to get the coordinates of NHS administrative units, and then query the Ordnance Survey data manually using string matching for the corresponding Parliamentary constituencies. The time dimension adds further complication to administrative geography. Parliamentary constituencies are regularly redrawn in response to demographic change, and different data sources deal with this in different ways; the Ordnance Survey administrative geography stores only the latest classification.

However, when issues such as changes of semantics do not occur, our techniques allow incremental on-the-fly updating for data consumption.

Many of the applications discussed in this article visualize a single store, using data harvested and processed into RDF by EnAKTing researchers. This data, along with the associated visualization based on the current contents of the store, can be refreshed at any point. Other applications query Data.gov.uk in real time. So, for example, See UK (<http://apps.seme4.com/see-uk/>) imports the UK's monthly crime data into its store, and the view is always of the latest figures (see Figure 1).

The Value of Place for Linking Datasets

Geography provides an intuitive way to align datasets—no surprise, as governments generate PSI about the territory over which they have jurisdiction, so the data has an implicitly geographical dimension. The LDW is well-stocked with geographical data; the Geonames service manages eight million URIs for geographical resources. Therefore, where there is an authoritative geographical knowledge base available, as in the UK, geography is an irresistible join point for datasets.³ (See the sidebar, "Related Work on the Linked Data Web.")

In our application for EnAKTing, the region gives context for the displayed data and is the central point from which we link to the LDW. New views of the data or concepts generate new searches and presentations on the basis of aggregations that make sense in the new contexts; for instance, having moved up the geographical hierarchy from a constituency to a county, the application can present the statistics (such as crimes committed) relevant to the county as a whole.

This approach will not work with some types of territory, such as Parliamentary constituencies, which don't map easily onto the administrative geography of the UK. Yet, if we can

Related Work on the Linked Data Web

Open government data (OGD) is becoming increasingly important across the globe, although currently most initiatives involve making data in proprietary formats downloadable. Surveys have shown that there are relatively few attempts to combine OGD with the linked-data web (LDW),¹ and that Data.gov.uk and Data.gov are unusual in their commitment to the LDW vision. Many other important and interesting developments have been more opportunistic, including the creative use of Open Street Map data in the aftermath of the Haiti earthquake in 2010, while initiatives such as the Open Government Partnership (www.opengovpartnership.org) have begun to spread best practices even further.

The work closest to our project aims to migrate Data.gov to the LDW. The Tetherless World Constellation (TWC) Linked OGD portal² also recognizes data-publishing stages for OGD on the LDW:

- the catalog stage, where an inventory of datasets is created;
- the retrieval stage, where a snapshot of the dataset's online data file at a point is input to a Linked OGD converter; and
- the conversion stage, where the data is converted to RDF in a layered manner that allows many of the conversion issues and bottlenecks to be sidestepped.

An initial automatic conversion is done by the portal, and enhancements such as mapping *ad hoc* database column names to common properties can be done by users. Many of its linking strategies were anticipated by the Data-gov wiki.³ The strategy of the LOGD portal has been to foster an LOGD community by actively engaging users with demos.

The Data-gov wiki is a social Semantic Web platform that has produced more than 5 billion triples, covering topics such as government spending, environmental records, and statistics on the cost and usage of public services. It goes through a series of steps similar to those just outlined, including conversion of data into RDF, enhancement and linking by declaratively associating URIs in related contexts (done both automatically and by hand), and designing applications and demos to address the important issue of data consumption. The Data-gov wiki limited its efforts to well-formed comma-separated value (CSV) files, and so was able to sidestep several conversion issues. It also took a lightweight

approach, with a minimal and extensible conversion to preserve the structure and content of the raw data and no more. The TWC team did not use properties from existing ontologies, to avoid manual moderation, but properties used in converted RDF data were dereferenceable (that is, accessible from their URIs) to terms in well-known ontologies (such as Friend of a Friend and Dublin Core) or RDF and XML pages generated by Semantic MediaWiki. The Data-gov wiki also focused on provenance, and was able to use this as a join point, linking by derivation- and version-based provenance associations.

Evangelos Kalampokis and his colleagues have also exploited the social Web, using OGD to enrich data mined from social networking and microblogging sites—for example, linking tweets from high-crime areas in the UK to the crime data from <http://police.uk> for those regions.⁴ The aim of this work is to allow policy makers to assess public opinion and predict public reaction. Kalampokis's team's linked data architecture integrates OGD with data mined from the social Web, to enable the collection of OGD related to a specific set of criteria that the decision-maker provides. The integrator integrates and stores as RDF the social data with objective data related to the specified target group, as well as the variables related to social data and real-world objective facts coming from government data. The improvement of the OGD comes via augmentation from the social Web, rather than from the integration processes used by EnAKTing and TWC LOGD. Crowdsourcing (obtaining data from a distributed group of citizens) is clearly an important way forward.

References

1. E. Kalampokis, E. Tambouris, and K. Tarabanis, "A Classification Scheme for Open Government Data: Towards Linking Decentralized Data," *Int'l J. Web Engineering and Technology*, vol. 6, no. 3, Inderscience, 2011, pp. 266–285.
2. L. Ding et al., "TWC LOGD: A Portal for Linked Open Government Data Ecosystems," *J. Web Semantics*, vol. 9, no. 3, Elsevier, 2011, pp. 325–333.
3. L. Ding et al., "Data-gov Wiki: Towards Linking Government Data," *Proc. AAAI Spring Symp. Linked Data Meets Artificial Intelligence*, AAAI, 2010, pp. 38–43.
4. E. Kalampokis, M. Hausenblas, and K. Tarabanis, "Combining Social and Government Open Data for Participatory Decision-Making," *Proc. 3rd Int'l Conf. eParticipation (ePart 11)*, LNCS6847, Springer, 2011, pp. 36–47.

establish that one entity is completely contained within another (for example, the Parliamentary constituency of Fareham within the county of Hampshire), we can discover relevant data and present it to the user, ideally exploiting existing LDW resources or bringing more geographical resources onto the LDW (see Figure 2).

To help with this kind of reasoning, EnAKTing has developed a service

(<http://geoservice.psi.enakting.org>) to support the discovery of geographical resources pertaining to the UK on the LDW by querying containment relations.³ This service exploits knowledge about instance equivalence that is already available via coreference systems such as SameAs (<http://sameas.org>). It normalizes the data, translating the `os:contains` relation into two statements, a `has-part` and

a `part-of`, to produce a structure such as that in Figure 2, allowing the service to infer containment via resources from difference datasets using `owl:sameAs`.

Reasoning Services

Geolinking services are only one kind of reasoning needed to enrich linked data. As another example, EnAKTing has developed a *backlinking* service

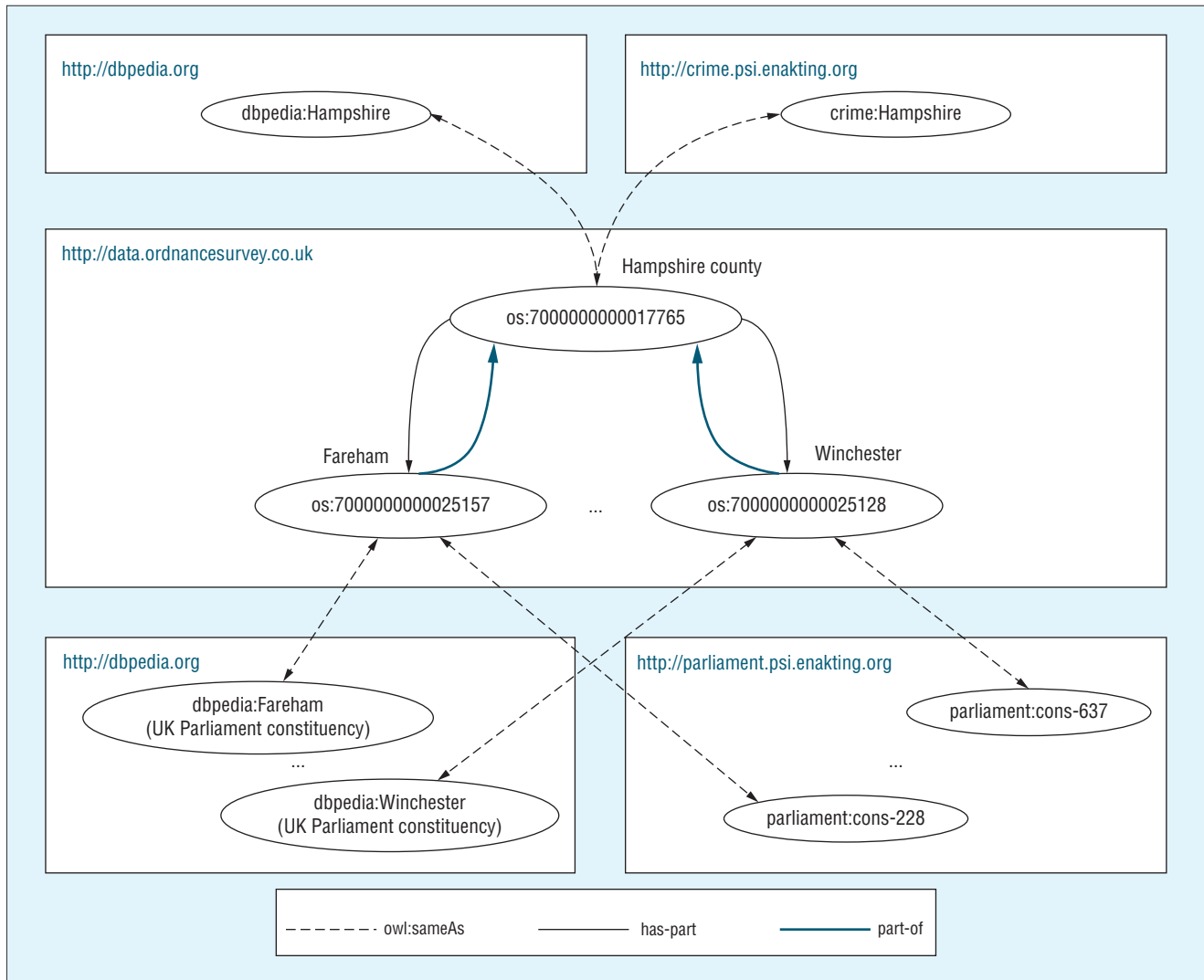


Figure 2. Inferring geographical containment with the EnAKTing Geoservice. The service can use `has-part` and `part-of` relations and `owl:sameAs` to infer that parliamentary constituencies Winchester and Fareham are in Hampshire, thereby giving vital context for the linking of datasets.

(<http://backlinks.psi.enacting.org/>),⁴ a generic architecture component to support the discovery of useful links between items across highly connected data sets (directed graphs) that direct URI resolution cannot find. The service discovers foreign URIs—that is, URIs X that appear in RDF triples of the form $\langle s, p, X \rangle$ in an RDF graph G , where domain $(X) \neq \text{domain}(G)$. A Foreign URI pattern discovery component crawls the LDW, retrieving all foreign URIs found in the datasets under consideration, and then asserts new URIs (generated using an `rdfs:seeAlso`

statement with the foreign URI in the subject position) into a backlinking knowledge base.

When backlinking is integrated with geolinking, the number of URIs discovered increases by orders of magnitude. For instance, although backlinking on its own discovers only a handful of URIs linking to `dbpedia:Hampshire` or equivalents from `owl:sameAs`, with the geoservice it retrieves thousands of resources representing such entities as schools in the area, CO₂ emissions, and census details, and provides hundreds of millions of extra links

between datasets such as DBpedia, Geonames, and OpenlyLocal, as well as the specific PSI datasets on which we tested it.

The backlinking service also exploits a coreference evaluation service developed within the Resilience for Survivability in Information Society Technologies (Resist) project, SameAs, which finds URIs that identify identical things within the scope of an application and then stores and publishes them. Note the context-relativity of such judgments; in contrast, the global scope of `owl:sameAs` implies a globally valid identity. Instead,



Figure 3. Multiple source link-sliding and custom charting for combined data with Geordi. The user has developed this view of the data and is able to examine specific expenses from aggregated spending data.

SameAs bundles equivalent URIs together and separates knowledge about equivalence from the main datasets, thereby enabling different applications to use different coreference services (and thereby different bundles) in different contexts.

Other EnAKTing services include temporal reasoning, catalog aggregation, determining whether a URI identifies a real-world object or a digital resource, and an explorer based on the Sparql query language (see <http://www.enakting.org/services.php>).

Browsing and Exploring

USAspending.gov is an effective visualization that lets citizens query data about US government spending, but it's a closed data system analogous to a Web 2.0 mashup: it doesn't let the user change the questions or link up with other data. In politics, this is a serious issue, as it concentrates power in the hands of application developers.

EnAKTing's Geometric Rich Data Interface (Geordi) browser (www.geordi.enakting.org) lets citizens convert

data in, say, a spreadsheet to RDF, and survey, explore, and link data without demanding an unrealistic level of engagement or programming skills.⁵ Uploading data with Geordi requires merely spreadsheet skills. (There are approximately 500 million Excel users worldwide). It uses predefined templates for translation into RDF, so that users must transform the spreadsheet into a template and then add metadata such as a dataset name, description, and URL.

Data discovery in Geordi also lowers skill demands (unlike other discovery services involving Sparql queries) by presenting users with a catalog of datasets to cue choice and exploration. If users want to go further, Geordi presents collections of resources of a particular type (such as constituencies, persons, and mortality statistics) that can lead to exploration of the data proper. Browsing uses link-sliding: refocusing on a specific set of items by looking at common properties. Figure 3 shows an entirely user-driven view of data in,

which the user has moved from NHS spending data to an examination of suppliers, expense areas, and dates on which expenses were incurred. With custom charting enabled, the user can examine spending against time for individual suppliers.

Linking is also possible via link-sliding; currently Geordi supports finding and asserting `owl:sameAs`, but future work will extend the range of properties that can be asserted.

The geographical focus of OGD has also prompted the development of See UK, another visualization of UK data (see Figure 1). See UK uses data from Data.gov.uk (for example, crime, transport, and education datasets) that has been processed into linked data, although the general-purpose design can also incorporate other datasets. The exploitation of the possibilities of linked data vastly increases the power of See UK's visualisations, for example (as in Figure 1) if data is enriched by calculating area statistics from point data, inferring aggregate values for regions that

don't have explicit data, and establishing links between the datasets.

The visualization provides a view centered on a chosen region of the specified granularity and presents a pie chart showing how that region compares with similar regions around it. Color indicates the worst (red) and best (green) areas. The user can navigate by looking and clicking on either the chart or the map.

See UK can present cross-dataset correlation naturally, because the data can be normalized by population or area. The user can therefore see how regions compare in terms of, for example, crime density by population or area, as opposed to having to guess that the level of crime is low because the population is low.

Lessons Learned, Questions Remaining

These projects, including EnAKTing, have taught us many lessons, some anticipated and others unexpected, some of which we've addressed while others pose open questions. Some of the more prominent relate to governments, the technical community, and citizens. Broadly, these lessons constitute a roadmap showing the issues in going from raw government data to data that is linked into the LDW and regularly consumed by citizens.

Lessons for Governments

There are several lessons we have learned for government. First, although hard-to-link formats are a problem, the regulatory setting of reusability is crucial: data needs an open license to begin to count as open data. Data managers don't have the authority to determine release modes, and policy makers need to understand that licensing restrictions are the biggest obstacle to OGD.

Second, few systems other than Data.gov, Data.gov.uk, and Eurostat

provide direct access to catalogs in raw formats; usually, they have to be discovered by HTML search. National PSI portals, where they exist, provide powerful and rich metadata, but regional data is less well described. Third, metadata standards tend to be tailored to the information needs of the primary intended users of the information assets.⁶ Metadata standards should be more widely adopted, and in the context of LDW they need to be directed toward all data users, rather than only the immediate communities who generated the data.

Fourth, there are often temporal or geographical gaps between datasets that limit the usefulness of linking them, or even methodological gaps that render the data incommensurable. Versioning on the LDW is a vital topic. Pan-national interpretation and comparison is particularly challenging. Where URIs are minted by governments, they and others should seek to reuse them; that way, they can become the essential join points of the OGD digital infrastructure.

Lessons for the Technical Community

There are also several lessons for the technical community. First, our experience shows that there are many steps in the creation of linked data from OGD that would benefit from better tools and, where possible, automation. The perennial user interface issue is critical in this politicized context, as the ease with which ordinary citizens can access and query the data is a crucial factor for OGD's value. If few people are able to interrogate the data, then OGD's contribution to democratic debate will be correspondingly less.

Second, work is needed to identify join points that let the data consumer assemble a coherent picture across

datasets, and that facilitate data normalization to detect and correct variations in nomenclature. The EnAKTing project has identified geography as one of these; others are time and provenance.⁷ The idea of *life events* from work in semantic e-government could be another valuable structuring principle. Identifying join points will be an important factor for helping governments identify resources to prioritize.

Third, the EnAKTing work is deliberately lightweight and pragmatic—for example, in the way it relies on well-known ontologies.⁸ We have used Scovo for statistical data because of its relative simplicity and widespread adoption. This design decision is pragmatic and should be made on a case-by-case basis; for example, Data Cube could easily be used instead if it supported data consumption scenarios that Scovo didn't.

Fourth, coreference resolution remains a serious problem. Rather than use `owl:sameAs`, we adopt the SameAs context-relative approach, which suits a data- and citizen-centric approach. Fifth, we must not neglect the consumption of linked data. Migrating OGD to the LDW should be accompanied by quick development of websites supporting data-centric interactions.

Lessons for Citizens

Finally, there are a few lessons for citizens. First, transparency will create opportunities for citizens to hold their governments accountable, and to develop and use innovative services. Citizens should use these new powers. The more they do, the more likely it is that OGD will continue to be released. Second, in an ideal world of OGD 2.0, the citizen would have rights and powers to correct the gaps, errors, and omissions of OGD, which is about things that people understand

or interact with (schools, bus stops, potholes, roads, and so on). With the right tools, citizens can discover and fix incorrect data. Moreover, infrastructure is needed to enable people to create their own linked data applications, to ensure that the full benefits of OGD are available. This raises complementary issues about control and stewardship.

Bottlenecks in Exporting OGD to the LDW

Much discussion about transparency has focused on the unwillingness of public service providers to surrender control of their data. However, we would argue that the pace of export of OGD to the LDW is also a significant factor, and EnAKTing has shown us several important bottlenecks in the process, such as the following:

- *Discovery of OGD.* More international and sector-based portals are needed rather than patchy national coverage, as well as integration tools to bring the data together.
- *Ontological alignment.* This could be addressed with better international standards (which are likely to emerge as transparency increases, since most datasets were not developed to be linked).
- *Interfaces.* Geordi has lowered the skill demands on anyone who wants to query OGD in a flexible and personalized way, and there needs to be more work in this direction.
- *Consumption.* It's important to increase the number of applications and intermediaries prepared to present information to citizens in creative and valuable ways. The intermediary ecosystem will be a vital determiner of how the technology develops. This includes not only the recognition of opportunities and

THE AUTHORS

Nigel Shadbolt is a professor of artificial intelligence and head of the Web and Internet Science Group in electronics and computer science at the University of Southampton, and will be co-director of the Open Data Institute in London. He is also an adviser to the UK government on open data and a member of the UK Public Sector Transparency Board. His research interests include the Semantic Web, Web science, and knowledge technologies. Shadbolt has a PhD in AI from the University of Edinburgh. Contact him at nrs@ecs.soton.ac.uk.

Kieron O'Hara is a senior research fellow in electronics and computer science at the University of Southampton. His research interests include the politics and philosophy of technology, particularly transparency, privacy, and trust on the World Wide Web and the Semantic Web. O'Hara has a DPhil in the philosophy of AI from the University of Oxford. Contact him at kmo@ecs.soton.ac.uk.

Tim Berners-Lee is the director of the World Wide Web Consortium, the 3Com Founders Professor of Engineering at the Massachusetts Institute of Technology, a professor of computer science at the University of Southampton, and will be co-director of the Open Data Institute in London. He is also an adviser to the UK government on open data and a member of the UK Public Sector Transparency Board. His research interests include the Web, the Semantic Web, Web science, and linked data. Contact him at timbl+ieee@w3.org.

Nicholas Gibbins is a lecturer at the University of Southampton. His research interests include scalable infrastructures for the Semantic Web, hypertext and hypermedia, and large-scale multiagent systems. Gibbins has a PhD in computer science from the University of Southampton. Contact him at nmg@ecs.soton.ac.uk.

Hugh Glaser is the chief architect at Seme4 and a visiting research fellow in electronics and computer science at the University of Southampton. His research focuses on building systems around Semantic Web technologies that are scalable and fit for purpose. Contact him at hugh.glaser@seme4.com.

Wendy Hall is dean of the faculty of physical and applied sciences and a professor of computer science at the University of Southampton. Her research interests include Web science, the Semantic Web, hypermedia, and multimedia. She has a PhD in pure mathematics from the University of Southampton. Contact her at wh@ecs.soton.ac.uk.

m.c. schraefel is a reader in the Agents, Interaction and Complexity Group in electronics and computer science at the University of Southampton. Her research focuses on human systems interaction. Contact her at mc+ieee@ecs.soton.ac.uk.

information gaps, but also the development of business models.

OGD needs to create economic value alongside its other merits. Until the data on the LDW makes it out into the world and we can understand the demand side, crucial feedback for local, regional, and national governments to identify high-value data sources will be missing.

OGD is an opportunity and a challenge for the LDW. The opportunity is to grow by linking with PSI—real-world, useful information with good provenance. The challenge is to manage the sudden influx of

heterogeneous data, often with minimal semantics and structure, tailored to highly specific task contexts. The EnAKTing project is intended to promote the LDW by developing techniques such as those described here to integrate OGD.

Releasing OGD is not solely a technical problem, though it presents technical challenges. OGD is not a rigid government IT specification, but it demands productive dialogue between data providers, users, and developers. We should expect a “perpetual beta,” in which best practice, technical development, innovative use of data, and citizen-centric politics combine to drive data-release programs.

Integrating OGD onto the LDW will vastly increase the scope and richness of the LDW. A reciprocal benefit is that the LDW will provide additional resources and context to enrich OGD. Here, we see the network effect in action, with resources mutually adding value to one another. ■

Acknowledgments

The work reported in this article was funded by Engineering and Physical Science Research Council Grant EP/G008493/1. We are indebted to EnAKTing researchers Gianluca Correndo, Costas Koumenides, Dan Smith, Max van Kleek, Ian Millard, Igor Popov, and Tope Omitola, and to the three anonymous reviewers of this article.

References

1. H. Alani et al., “Unlocking the Potential of Public Sector Information with

Semantic Web Technology,” *Proc. 6th Int’l Semantic Web Conf. (ISWC 07)*, Springer, 2007, pp. 708–721 .

2. T. Omitola et al., “Put in Your Postcode, Out Comes the Data: A Case Study,” *Proc. 7th Extended Semantic Web Conf.*, Springer, 2010, pp. 318–332.
3. G. Correndo et al., “Geographical Service: A Compass for the Web of Data,” *Proc. Linked Data on the Web (LDOW 10)*, CEUR Workshop Proceedings, 2010; http://CEUR-WS.org/Vol-628/ldow2010_paper15.pdf.
4. M. Salvadores et al., “Domain-Specific Backlinking Services in the Web of Data,” *Proc. 2010 IEEE/WIC/ACM Int’l Conf. Web Intelligence*, IEEE, 2010, pp. 318–323.
5. I. Popov et al., “GEORDi: Supporting Lightweight End-User Authoring and Exploration of Linked Data,” tech. report,

Electronics and Computer Science, University of Southampton, 2010.

6. C. Koumenides et al., “Global Integration of Public Sector Information,” *Proc. Web Science Conf. 2010 (WebSci 10)*, 2010; http://journal.webscience.org/3032/websci10_submission_38.pdf.
7. L. Ding et al., “Data-gov Wiki: Towards Linking Government Data,” *Proc. AAAI Spring Symp. Linked Data Meets Artificial Intelligence*, AAAI, 2010, pp. 38–43.
8. L. Ding et al., “TWC LOGD: A Portal for Linked Open Government Data Ecosystems,” *J. Web Semantics*, vol. 9, no. 3, Elsevier, 2011, pp. 325–333.



Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.