

The Information Spring



Kieron O'Hara • University of Southampton

A (Nobel-prizewinning) wag once said that you can see the computer age everywhere except the productivity statistics.¹ If that's true of the private sector, how much sharper a barb is it against governments, who strive to do so much, and do it so terribly badly?

When we look at the sort of cutting-edge stuff reported in journals such as this, it seems a no-brainer to apply semantic technologies and linked data to public administration. After all, government is massive, unwieldy, distributed, and siloed. Its giant troves of data are rendered in myriad representational formats and standards. It struggles to coordinate its various departments and layers – health, social services, policing, local, state, and national bodies, and quasi-independent agencies bash their heads together more often than provide “joined up government.” Organizations that provide services, using information from heterogeneous sources, expressed via diverse ontologies – the classic Semantic Web use case.

The great sociologist Max Weber conceived government as a sort of giant information processor, while Herbert Simon, doyen of information scientists, argued as early as 1973 that “If we examine the kind of information that executives use, we find that a large proportion of it is simply natural language text. ...[Computers could be] initial filters for most of the information that enters the organisation from outside.”² Yet despite great strides and many successes in this area, semantically enabled public administration is still something of a minority sport; it would be 30 years before anyone acted on Simon's insight. Is this something to depress us pointy-headed techies?

I would argue not, because semantic technologies could revolutionize not just governments'

information processing practices, but the purpose and scope of government itself. We have to come to terms with not only the machinery of administration, but also the wider question of the relationship between citizens and Leviathan. This will take some wrestling with, and is obviously not merely a technical question.

Liberation has been a theme in modern politics, from the Prague Spring to the Arab Spring. We are seeing the emergence of an Information Spring, which could set data and information free to serve the people – if we understand its implications in the right way.

Semantic E-Government: Challenges and Slow Progress

Given the problems of coordinating government, it's unsurprising that governments around the globe fell upon technology with delight. Given the challenges of implementing digital versions of legacy analogue systems, it's equally unsurprising how little has happened beyond rhetoric.

E-government involves digitizing governments' interactions, spawning phrases such as G2C, G2B, G2G, B2G, and C2G – indicative for their implication of a one-way information transfer. However, some tried to push further, via what became known as “transformational government,” the use of IT plus business process reengineering (BPPE) to improve delivery of public services.

In this spirit, essays in semantic e-government began to appear. The challenges were many. Government has many drivers – unions, taxpayers, party members, big donors, and interest groups need to be kept happy, and efficiency isn't always the main aim. The perception of the Semantic Web as a complex and difficult technology combined with the enormous difficulties

of change management in the giant government machine present a daunting prospect.

The science still pushed forward. The EU sponsored interesting projects with names such as Access e-Gov, OntoGov, and SemanticGov to develop and roll out the technology. Practical ideas such as life-event ontologies,³ which created a unified set of terms relating to significant events in citizens' lives — moving house, registering to vote, registering a death — raised the hope that citizen-centric services might appear irrespective of the combination of agencies collaborating behind the interface. Complex architectures were designed, such as that of SemanticGov, where five layers connected stakeholders, applications, and service providers via a semantically enabled middleware environment.⁴

Pragmatics, Natch

Such ambitious programs could potentially implement a whole-hearted transformation of government processes, but, as so often with the Semantic Web, pragmatics is the Achilles heel. A lot of BPRE would seem to be on the cards — should this be managed as a risky big bang, or should the focus be on incremental change instead, with the potential for loss of momentum, complexity, and compromises with legacy systems? Tech suppliers, which benefit from current arrangements, lack expertise in semantic technologies and incentives to change, while the trend of outsourcing IT provision has devastated government's in-house ability to manage major upgrades.⁵ One perennial problem is first-mover disadvantage — the corollary of Metcalfe's law is that the later users benefit most, so how will early adopters find partners and build networks?

Furthermore, although semantic e-government is more citizen-centric, it's still prescriptive. Standards are

determined and services are specified by governments, whether produced in-house by government agencies, outsourced (provided by the private sector and paid for by government) or privatized (private sector services paid for by users). Life events might make it easier for me to find the services I need at a particular juncture, but I still have to tailor my life to a life-event ontology designed by someone else. James Scott's brilliant *Seeing Like a State* details the state's need to render us, its citizens, legible to it.⁶ Our choices and expressions are narrowed whenever they are rendered in standardized forms as pieces of data. Is this the right solution for a pluralistic society?

However, the e-government context has been changing; transparency is becoming a major driver of political change. The idea that "sunlight is the best disinfectant," in the words of Louis Brandeis — that citizens' access to information would facilitate understanding of democracy and decision-making, hold governments to account, and reduce opportunities for corruption — had been honored more in the breach than the observance. Yet as the ICT revolution and the Web flourished, the spread of information helped foster good governance. As more information-sharing technology moved into place — not only the Web to link documents, but linked data, big data, massive number-crunching capabilities, and the democratization of analytic tools — it combined with the ideology of serendipitous reuse to turn the perception of government transparency from medicine to opportunity.

Open Data

Enter open data. A few early pioneering exercises (including AKTive PSI, in which your columnist had a small involvement)⁷ demonstrated the value of reusing data in new contexts, and the need for pragmatic development methods.⁸ The need for

legitimacy for the economic stimulus following the financial crisis meant that each cent from the US taxpayer needed to be accounted for, leading to the development of data.gov (launched May 2009), now the open repository for nonsensitive US government information.

I won't go into enormous detail on open data, which Nigel Shadbolt and I have discussed in the Linked Data department in this magazine.⁹ Suffice it to say that data is open if it is machine-readable and online under an open license, so its use and reuse aren't constrained by terms and conditions, or access control mechanisms. Ideally, the data would be represented in nonproprietary formats, such as CSV or RDF. Open data are best expressed using open standards — doubly exploiting the power of open.

Utilitarian arguments for open data will eventually be shown to be good or bad by sophisticated econometrics. But more powerful, in the Digital Citizen's humble opinion, is the argument for a right to government data that is nonpersonal and nonsensitive. Government is empowered (or empowers itself) to collect data (as we know from Edward Snowden's revelations, quite large quantities of it). It can do this morally because of the legitimacy that citizens provide via the ballot box, and economically because of the taxes we pay. Citizens are central to the data-gathering model, so why shouldn't citizens reap some of the benefits, rather than being passive recipients of government-defined services?

Democratic governments have been generally, if sclerotically, moving toward freedom of information (FoI) for some time. The US Freedom of Information Act became law under President Lyndon Johnson on Independence Day 1966, while data protection — a complex hybrid balancing data sharing with personal privacy protection — emerged in the 1970s

and was codified in principle by the Organization for Economic Cooperation and Development in 1980 (www.oecd.org/internet/ieconomy/oecdguidelinesontheProtectionofPrivacyandTransborderFlowsofPersonalData.htm). More recent innovations include the provision of data in a form suitable for the user's purposes, rather than the provider's. Some governments were swept along by a wave of transparency they hadn't asked to surf – the UK's Tony Blair, a man not hitherto noted for self-criticism, berated himself in remarkable terms for introducing FoI: "You idiot. You naïve, foolish, irresponsible nincompoop. There is really no description of stupidity, no matter how vivid, that is adequate. I quake at the imbecility of it."¹⁰

Blair's regrets notwithstanding, the major step forward was the realization that governments could provide information routinely, rather than insisting that people ask for it (assuming they knew it existed in the first place). By 2003, the EU Re-Use of Public Sector Information Directive stated that "Member States shall ensure that, where the re-use of documents held by public sector bodies is allowed, these documents shall be re-usable for commercial or non-commercial purposes... Where possible, documents shall be made available through electronic means" (<http://ec.europa.eu/digital-agenda/en/european-legislation-reuse-public-sector-information>).

This is a win-win situation. Governments have more information than they can handle; openness lets others use it. Indeed, government agencies actually find it easier to consume their open, and especially linked, data than to query their own heterogeneous databases. Quality can be crowd-sourced as more people see the data produced; it's hard for governments to know everything, but everyone knows something: knowledge is distributed.

Furthermore, initiatives are emerging to give citizens some

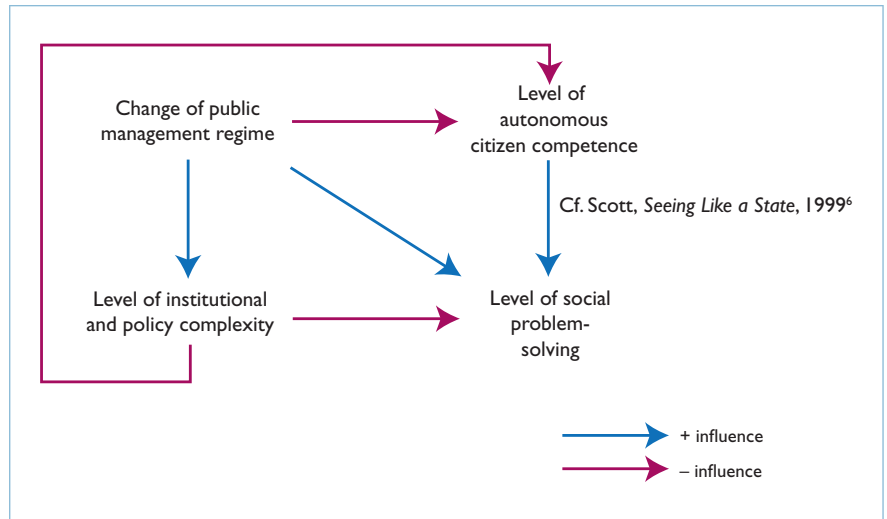


Figure 1. The effects of changes in public management regimes. Blue arrows signal a positive influence (increase), whereas magenta arrows indicate a negative influence (decrease). Figure adapted from Digital Era Governance: IT Corporations, the State, and E-Government.⁵

access to personal data about them. In the US, the Blue Button initiative lets military veterans download health information, and an expansion aims to add demographics, lab results, pathology reports, and more (www.va.gov/bluebutton/). The Mesinfos group in France has launched a pilot study that makes private-sector data available for new applications and services (<http://fing.org/?-MesInfos-les-donnees-personnelles-&lang=en>). In the UK, midata coordinates routes for private-sector data back to the consumer in a safe environment, with data owners' cooperation.¹¹

Such initiatives let citizens/consumers combine their own personal data with rich data sources from governments to increase their understanding of their own environments. This vision — the Information Spring — has a degree of contrast with the application of more traditional Semantic Web methods to government information processing, in that it's bottom-up rather than top-down, and lightweight rather than demanding complex architectures. Yet it opens the door to the application of semantic technology via open standards for linked data. Open data isn't necessarily linked, but

link technology sits very well with the open data ideology.^{9,12}

Making Things Better in the Information Spring

In an interesting analysis of governance, Patrick Dunleavy and his colleagues show how changes in government have unintended consequences for social problem-solving (Figure 1).⁵ In the diagram, blue arrows signal a positive influence (that is, an increase), whereas magenta arrows indicate a negative influence (decrease). The point of a change in public management is to solve some perceived social problem (for example, Medicare in the US guarantees access to healthcare for older people, while Bolsa Família in Brazil helps relieve poverty and increase access to education); the change in management tends to increase the level of social problem-solving (shown by the central diagonal arrow).

However, other factors influence social problem-solving. Citizen competence, conceived autonomously from government – self-help – is also important for addressing social problems.⁶ But citizen competence is highly context-dependent, so the change in

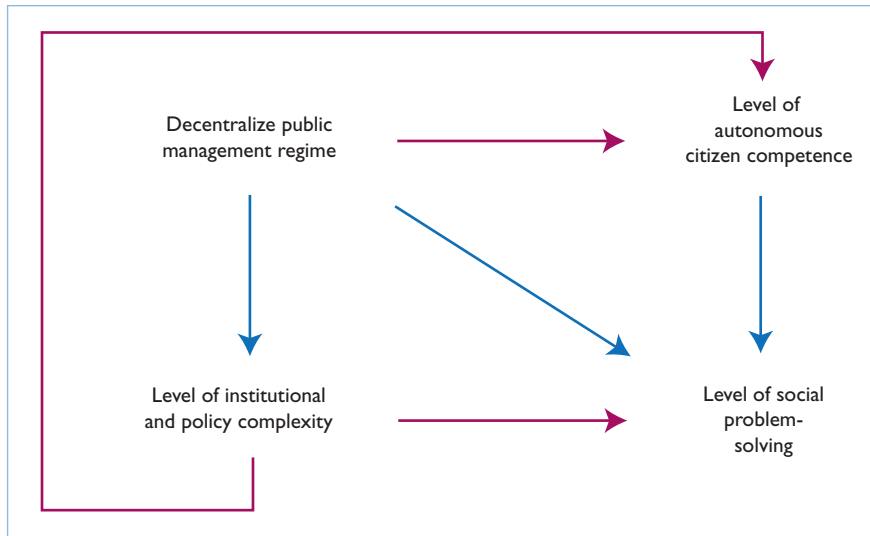


Figure 2. Decentralizing public management. Blue arrows signal a positive influence (increase), whereas magenta arrows indicate a negative influence (decrease). Decentralization will tend to increase citizen competence, which will boost its positive influence on problem-solving.

public management decreases competence, thereby decreasing its positive influence on problem-solving.

Furthermore, changes in management increase institutions' complexity,¹³ which makes them less effective in problem-solving (another magenta arrow). Finally, the complexity of institutions makes them more difficult to work with, which nullifies citizens' competence, rendering them less able problem-solvers. The change in management could have a positive, direct effect on social problem-solving, but it also produces three negative indirect effects, which in the worst cases will offset all the gains and deliver worse outcomes.

The Information Spring, liberating information from the drudgery of serving only the task for which it was gathered, could transform not only the use of personal data, but government's purpose as a whole. It is truly transformative.

This isn't simply a technocratic choice about delivery systems. Ultimately, it harks back to a deep dispute in political philosophy between Baron Montesquieu (1689–1755), who understood politics as an ongoing

process of triangulation between competing interests, and Jean-Jacques Rousseau (1712–1778), who theorized the “general will” – the idea that a mass of people share a common set of interests to promote their well-being.¹⁴ Montesquieu's politics were designed to shackle powerful groups by separating powers; Rousseau's aimed to clear the obstacles hindering governments' pursuit of the general will. Were they around today, Montesquieu would be all for giving data back to the people, whereas Rousseau would surely prefer a more top-down approach, because “when particular interests begin to make themselves felt and sectional societies begin to exert an influence over the greater society, the common interest then becomes corrupted and meets opposition.”¹⁵

In politics generally, Rousseau currently has the upper hand. For instance, FoI-hating Tony Blair is on Rousseau's side – he once called his party “the political wing of the British people,” an expression of a generalized national will if ever there was one. Such universalizing sentiments are common: for example, the stellar group of authors who constitute the

Oxford Martin Commission for Future Generations recently argued that “The ability to address today's global challenges is undermined by the absence of a collective vision for society. To remedy this, the Commission urges renewed dialogue on an updated set of shared global values around which a unified and enduring pathway for society can be built.”¹⁶ Against this, the still small voice of Montesquieu protests that individuals are the best judges of their own interests, and sometimes their values will not coincide – and that's okay.

Under the Information Spring, the general will evaporates, and individual interest reasserts itself. Services can be defined and provided by governments, the private sector, or non-profits (this isn't an anti-government initiative that says free markets are best). They all have the same data to work with – no monopolies, no rent-seeking. Citizens have at least some access to the data that describes them and their behavior, allowing increased personalization. The possibility of varied services depending on the same data enables leveraging the citizen competence to improve social problem-solving, as we can see if we adapt the Dunleavy diagram for the Information Spring (Figure 2).

In such a world, public management doesn't change to respond to social problems, but is rather decentralized, flattened, and maybe even shrunk. Let's assume, for the sake of argument, that this tends directly to reduce the level of social problem-solving (a big assumption, actually), so the central blue arrow turns magenta. However, the indirect effects are different. Decentralization will tend to increase citizen competence, which will boost its positive influence on problem-solving. It will also decrease institutional complexity, reducing its negative effects on social problem-solving and citizen competence.

Is the Information Spring with us? Not quite — this is the political economy of a world that does not yet exist. Under such a regime, new injustices, inefficiencies, and inequalities will no doubt be uncovered. Yet a new landscape is gradually emerging, and institutional structures are beginning to create frameworks for open data.

Montesquieu and Rousseau each inspired a revolution. The Founding Fathers in the US read Montesquieu, and created a lasting settlement with a bicameral system and separation of powers. Robespierre in France was a devotee of Rousseau, who streamlined government and removed the restraints on it; the result was a bloodbath that turned the republic into a predatory empire within 15 years.

Of course, the current gridlock in Washington doesn't put the separation of powers in a very good light. But what if government saw its job as making sure interested groups got access to the right data? Could that be a more realistic model for social problem-solving at a time when the role of the state itself is in bitter dispute?

Just a thought.



Acknowledgments

This work is supported by SOCIAM: The Theory and Practice of Social Machines, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1.

References

1. R. Solow, "We'd Better Watch Out," *New York Times Book Rev.*, 12 July 1987, p. 36.
2. H.A. Simon, "Applying Information Technology to Organization Design," *Public Administration Rev.*, vol. 33, 1973, pp. 268–278.
3. I. Trochadis, E. Tambouris, and K. Tarabanis, "An Ontology for Modeling Life Events," *IEEE Int'l Conf. Services Computing*, 2007, pp. 719–720.
4. T. Vitvar et al., "Pan-European E-Government Services Architecture," *Semantic Technologies for E-Government*, T. Vitvar, V. Peristeras, and K. Tarabanis, eds., Springer, 2010, pp. 25–46.
5. P. Dunleavy et al., *Digital Era Governance: IT Corporations, the State, and E-Government*, Oxford Univ. Press, 2006.
6. J.C. Scott, *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*, Yale Univ. Press, 1998.
7. H. Alani et al., "Unlocking the Potential of Public Sector Information with Semantic Web Technology," *Proc. 6th Int'l Semantic Web Conf. (ISWC 07)*, 2007, pp. 708–721.
8. H. Alani et al., "Building a Pragmatic Semantic Web," *IEEE Intelligent Systems*, vol. 23, no. 3, 2008, pp. 61–68.
9. N. Shadbolt and K. O'Hara, "Linked Data in Government," *IEEE Internet Computing*, vol. 17, no. 4, 2013, pp. 72–77.
10. T. Blair, *A Journey*, Hutchinson, 2010, p. 516.
11. N. Shadbolt, "Midata: Towards a Personal Information Revolution," *Digital Enlightenment Yearbook 2013: The Value of Personal Data*, M. Hildebrandt, K. O'Hara, and M. Waidner, eds., IOS Press, 2013, pp. 202–224.
12. L. Ding, V. Peristeras, and M. Hausenblas, eds., *IEEE Intelligent Systems*, special issue on Linked Open Government Data, vol. 27, no. 3, 2012.
13. J.A. Tainter, *The Collapse of Complex Societies*, Cambridge Univ. Press, 1988.
14. P. Riley, "Rousseau's General Will," *The Cambridge Companion to Rousseau*, P. Riley, ed., Cambridge Univ. Press, 2001, pp. 124–153.
15. J.J. Rousseau, *The Social Contract*, IV:5.
16. P. Lamy, ed., *Now for the Long Term*, Univ. of Oxford, 2013; www.oxfordmartin.ox.ac.uk/downloads/commission/Oxford_Martin_Now_for_the_Long_Term_Executive%20Summary.pdf.

Kieron O'Hara is a senior research fellow in the Web and Internet Science Group in the Electronics and Computer Science Department at the University of Southampton. His research interests include trust, privacy, open data, and Web science. O'Hara has a DPhil in philosophy from the University of Oxford. Contact him at kmo@ecs.soton.ac.uk.

New in 2014



IEEE CLOUD COMPUTING

IEEE Computer Society's newest magazine
tackles the emerging technology of cloud computing.
Subscribe today!

computer.org/cloudcomputing

IEEE  computer society

