
An Introduction to the Open Data User Group

Paul Malyon

*Head of B2B Transactional & Marketing
Products at Experian Ltd and Member of
the Open Data User Group*

DOI: [10.5033/ifosslr.v6i1.99](https://doi.org/10.5033/ifosslr.v6i1.99)

Abstract

This article gives a succinct introduction to the work of the Open Data User Group and the demand driven process being employed by the United Kingdom to boost the release of publicly owned data assets and how this is creating greater transparency and economic benefit to the country. Examples given also show how international collaboration will enhance these benefits as more formal structures and processes are created to manage the Open Data explosion.

Keywords

Open data; demand driven; transparency; open data institute; open data user group; addresses

What is Open Data?

Open Data represents a fundamental shift in the way people communicate ideas and information. Inextricably tied to the World Wide Web, the purpose of Open Data is to facilitate interoperability and intermixing of data sets. It is seen by many web scientists as the key ‘language’ as a further 3 billion people gain access to the Web over the next decade.

The Open Data Handbook defines Open Data as “data that can be freely used, reused and redistributed by anyone subject only, at most, to the requirement to attribute and sharealike.”¹ In this sense, it is very similar to Open Source software in the way it can be obtained, combined with other information and used to create value. A more complete explanation is available via the Open Data Handbook.² Today, Open Data is primarily non-personal information as this avoids the risks posed to privacy and personal security.

Open Data as a movement is still in its infancy with work focussing on the release of Government statistics and ‘base level’ data such as maps, company listings and other data of social, environmental or economic importance. This article describes some of the early work

¹ <http://opendatahandbook.org/en/what-is-open-data/>

² <http://opendatahandbook.org/en/what-is-open-data/>

in the UK and further afield to bring Open Data to the masses as part of a drive for greater transparency and sharing of information, as well as how this ‘public sector first’ approach is beginning to influence the world of big business.

The Open Data Movement

The growth of Open Data in the UK is linked inexorably to the growth of the web by several things, with the most obvious being the name Sir Tim Berners-Lee.³ This article will come back to him later. However, there are some clear parallels to be drawn immediately between the Web and Open Data. The Web was intended for use by anyone, for free and for any purpose. Open Data is exactly that: information released for Free (or at most with a cost-recovery charge) which can be used, re-used and combined with other information or software tools to create a service, product or dataset that can then be shared or sold. Open Data is often available as raw information in a spreadsheet and increasingly as an API (or direct software link) for easy use by developers and data scientists. Importantly, there should also be no control on what the data can be used for (other than the normal lawful purposes governed by state personal data protection and other laws).

Open Data in its most simple and common guise is often made up of the information collected or created by public bodies in the course of providing services to tax payers, for example: maps, addresses, rail stations, weather forecasts and tax receipts. Such bodies often build up vast quantities of reference data and transactional results (e.g. whether or not your child’s school had a good pass rate in GCSE Geography). That data may be used to improve those services, recoup the cost from other agencies or publish statistics every now and then, such as the Census or GDP.

In the past, this data was either locked away within the agency that collected it or sold to the open market for rather high prices with very complex and restrictive licenses designed to prevent onward sale and reuse (or at least make some more money from any reuse!). Good examples of this practise include Ordnance Survey⁴ maps or Royal Mail postcode data⁵ – all of this data built by government agencies in order to serve their citizens. This data is paid for by taxes but then sold back to those same citizens for a profit.

Open Data is different – the agencies releasing it have realised that you’ve already paid for it, it’s not ‘secret’ or personal and you have a right to see the information and use it for the benefit of your business or community. It also serves the notion of ‘transparent government’ as it shows you (or more commonly, investigative journalists) what your tax money is spent on; giving you the opportunity to hold those in power to account and make an informed choice on election day. As an added bonus, this extra scrutiny makes it easier for public bodies to find efficiencies and improve services.

The UK Open Data Structure

As with everything involving Government, the number of bodies involved in Open Data weaves a complex web. I’ll try to simplify this as much as possible and identify only the key

³ <http://www.w3.org/People/Berners-Lee/>

⁴ <http://www.ordnancesurvey.co.uk/business-and-government/products/index.html>

⁵ <http://www.poweredbypaf.com/>

people, bodies and processes and consider the process which applies when a person or body seeks to have some data held by the public sector released as open data.

The body tasked with collecting requests for data is the Open Data User Group⁶ (ODUG). The ODUG is an independent body set up by the Cabinet Office to be the voice of the users of public sector information. For example, citizens using apps to locate the nearest bus stop and the businesses building those apps. Requests for access to public sector data may be made through a process administered by the Cabinet Office called the Data.gov.uk Data Request Mechanism. The Cabinet Office reviews such requests in conjunction with ODUG as discussed below.

Made up of representatives from commercial data users large and small, academia, the public sector and so on, ODUG, in addition to collecting requests, is tasked with receiving feedback on existing open datasets (including licensing, access & quality) and taking this to government to encourage the release of quality open data that will benefit the widest possible audience. Requests are made via the data.gov.uk website and are usually public – this transparent method allows for the community to be involved early and support requests or point the requester in the right direction if the data is already available or is not a government-held dataset. At time of writing, the data.gov.uk portal had received 789 requests since September 2012⁷.

The requests are then filtered down by ODUG. Many requests can be dealt with simply and quickly by the Cabinet Office's Transparency Team (a group of civil servants tasked with encouraging other departments to open up their data). More complex requests are worked into a business case (often involving the original requestor and other interested parties) which can then be taken to the next group of stakeholders – the Public Sector Transparency Board⁸.

The Public Sector Transparency Board is responsible for the wider transparency agenda (of which Open Data is only one part) with membership ranging from the likes of Sir Tim Berners-Lee (inventor of the World Wide Web) and Sir Nigel Shadbolt (Chairman of the Open Data Institute and Professor of Artificial Intelligence at Southampton University) to Dr Rufus Pollock (Open Knowledge Foundation⁹) and various business leaders. This group has a wider remit advising government on greater transparency to create more interest in the political process from the general public, to encourage more collaboration between the public and private sectors, and to create the environment for economic growth.

The Public Sector Transparency Board will review the business cases and (where required) liaise directly with the relevant body to further the request or ask ODUG to continue their work to strengthen the case in specific areas.

In many instances though, business cases are taken directly to government bodies by the Cabinet Office's own Transparency Team who work closely with ODUG. This can lead to 'quick wins' in areas of quality assurance or licensing, or simply speed up release schedules.

While the above may sound complex, the Data Request Mechanism can be simplified as follows.

1. A request for data is made on data.gov.uk (either for new data or improvements to

⁶ <http://data.gov.uk/odug>

⁷ <http://data.gov.uk/odug-roadmap>

⁸ <https://www.gov.uk/government/groups/public-sector-transparency-board>

⁹ <https://okfn.org/>

-
- current data).
2. Requests are assessed by ODUG and Cabinet Office.
 - a. Complex requests go into an ODUG Business Case template.
 - b. Simple requests are processed by Cabinet Office and the relevant Department (e.g. fixing a broken link or updating a dataset).
 3. Business Case is published on data.gov.uk and discussed by Public Sector Transparency Board.
 4. Data owner (Department) responds to the request.
 5. Data is released (often following further consultation) or a reason for non-release is given.

The UK is a leading light in Open Data. All of the work going on with the Transparency Board, ODUG and various other sector or department-specific groups contributes to the Transparency Agenda. This thought leadership can be highlighted best via two other groups (while not connected directly to the data request mechanism they are still of great importance) – the Open Data Institute and the Open Government Partnership.

The Open Data Institute

Set up in 2012 by Sir Tim Berners-Lee & Sir Nigel Shadbolt with £10m of government funding (to be matched by private sponsorship over 5 years); the Open Data Institute¹⁰ (ODI) was born from the work of the two Knights pushing successive governments to invest in technology and data innovation via the Public Sector Transparency Board¹¹, Technology Strategy Board¹² (now known as Innovate UK – a group set up to fund and champion economic growth based around tech industries) and other outreach work.

The ODI has set out to be the hub of open data innovation in the UK – training data scientists with universities and schools, mentoring start-ups, training public servants and existing businesses and liaising with other groups to improve the infrastructure that supports innovation (for example, working with ODUG on business cases). The ODI have extended this role by offering Open Data Certificates to public sector bodies (and private organisations) who publish Open data to signify the quality, usefulness and openness of the data.

While a UK-only organisation, the ODI has garnered international attention with over 30 governments from around the world visiting their Shoreditch base to talk about setting up their own franchise (with more than a dozen ‘nodes’ now active in locations such as Dubai, Gothenburg & Chicago). This influence on global open data initiatives leads us nicely onto the other significant piece of the jigsaw.

¹⁰ <http://theodi.org/>

¹¹ <https://www.gov.uk/government/groups/public-sector-transparency-board>

¹² <https://www.gov.uk/government/organisations/innovate-uk>

The Open Government Partnership

This group of around 60 nations has committed to the principles of transparent government and working with civil society to encourage the release of data to provide informed choices and strong accountability. Countries involved range from Mexico to Sweden to Kenya and beyond.

The UK co-chaired the group (alongside Brazil) in 2013 and used the OGP¹³ summit in October that year to release the first draft of the UK National Information Infrastructure (a plan for a single store of key datasets – mainly open – that are required for the efficient running of UK society, such as public service locations, company identifiers and transport data)¹⁴ and announce a ‘Beneficial Ownership’ database¹⁵ of companies. This database identifies those people who have beneficial ownership of a company. This assists with the more efficient collection of business taxes and helps tackle certain tax avoidance techniques. Many OGP members announced significant new commitments to open government and many other countries announced their intention to join.

The OGP could be regarded as something of a talking shop but aims to work across global borders to build opportunities for experiences to be shared and progress to be encouraged. The effectiveness of the OGP will be easier to judge once all of the commitments made in 2013 are enacted.

So what does all of this mean?

The Open Data movement is a global process that is linked to the open government agenda to offer greater transparency, accountability, choice and ultimately opportunities for efficiency within government as well as economic / social opportunities outside of it.

The UK has been at the vanguard of making government more transparent and looking for opportunities to enable economic growth using open data.

While progress has been mixed, a few good examples help to explain the opportunity.

Case Study 1 - The NHS – saving money through data sharing

Work by Mastodon C (one of the start-ups mentored at the ODI), Health Care UK and the writer of Bad Science (Dr Ben Goldacre) demonstrated the potential size of savings available to the NHS if prescribing doctors switched from branded to cheaper, non-branded alternatives of common drugs. The specific example examined statins (used to prevent cardiovascular problems) and highlighted that even though doctors were advised to use the cheapest available product (from 81p per prescription) versus more expensive (up to £20), branded versions, this was not what was actually happening on the ground.

¹³ <http://www.opengovpartnership.org/>

¹⁴ <http://data.gov.uk/dataset/national-information-infrastructure>

¹⁵ <http://www.theguardian.com/politics/2014/apr/21/vince-cable-public-register-company-ownership>

It's important to note that studies found all versions of the drug were equally safe and effective for patients so there really was no good reason to prescribe the expensive alternatives.

The study examined 37 million rows of data and found potential monthly savings of £27m if prescriptions of the two branded versions had been switched to cheaper alternatives¹⁶.

The team noted that previous studies had estimated potential savings of over £1 billion per year across a number of drug types. This specific example shows how simple it would be to save money on the NHS budget to direct to new drugs, hospital buildings or other services.

If more NHS data (including anonymous outcome data) is made available, there are many potential uses for savings, faster new drug studies and informed choice for patients across the board.

In 2013 the UK government announced the Care.data¹⁷ initiative. This met with public outcry because of the perceived lack of consultation on plans to share patient outcome data from GPs with other public bodies and private businesses. The Care.data experience shows that there is still work to be done to educate society (and government) on the risks and benefits of sharing data (although sharing data is not the same as Open Data, the two are closely linked).

Case Study 2 - The Birmingham Civic Dashboard

Between August 2011 and May 2013 the Birmingham Civic Dashboard ran in order to study the way people interacted with Birmingham City Council, what services they wanted, when and where. Using open mapping data from the Ordnance Survey, the team from developer Mudlark made a simple to use tool for council workers and the public to view the kinds of requests at different times of day to get a better understanding of the issues affecting people and how the council responded to them.

All of the resulting data was also made available in an Open form for others to download¹⁸.

This kind of engagement could be rolled out further within local authorities or even nationally to help tax payers understand where their money goes, the kinds of services available to them and then choose what they need, where and when. For government bodies, they would be able to direct their resources at what was needed most in much shorter timescales based upon actual data rather than long term estimates.

Case Study 3 - Data.gov.uk Data Request Mechanism

The Cabinet Office manages a mechanism, called the Data.gov.uk Data Request Mechanism, which allows the wider 'data community' to actively influence the release of data by the UK public sector.

¹⁶ <http://www.economist.com/news/britain/21567980-how-scrutiny-freely-available-data-might-save-nhs-money-beggar-thy-neighbour>

¹⁷ <http://www.nhs.uk/NHSEngland/thenhs/records/healthrecords/Pages/care-data.aspx>

¹⁸ <http://civicedashboard.org.uk/>

Requests are regularly reviewed with the ODUG to identify quick wins and those requests that require a full benefits case to demonstrate the value to the relevant data owner.

Since the current process was introduced in autumn 2012, datasets released (or scheduled for release) include historic house prices from the Land Registry¹⁹ and the Charities register. Benefits cases published in the first year included the VAT register²⁰ (which has led to a key announcement in the UK 2014 budget on the release of this data), an open national address register (now being actively pursued by the ODI), Energy Performance Certificates (still under consideration), data on vehicle registrations and stolen vehicles (currently under consultation) among several others.

The key examples here have changed attitudes in the relevant departments leading to open consultations on how best to release the data rather than reasons not to release it. The ODUG and its group of stakeholders are also helping the government respond to key issues such as the flooding in England in the winter of 2013/14 via the release of data and a #FloodHack event. This led to the historic announcement from the Environment Agency to become a completely 'Open Data' agency.²¹

In short, the work of ODUG in representing the economic and social needs of the data community is contributing to shift the position of government on open data from *one of those things that had to be done to tick a box for the Minister* to *a real opportunity to innovate, change behaviours and create significant benefits*.

Building for the future

As everything above hopefully explains, the journey towards total Open Data and transparent governance is progressing nicely but there is still more to do. The UK Public Administration Select Committee recently published a report pointing out that while progress is being made, some serious mistakes must never be repeated²² (with specific reference to the privatisation of the Royal Mail Postcode Address File).

In the example of the Postcode Address File (PAF), there is a clear need for a single, accurate and Open address register for purposes beyond delivering mail. The Census requires addresses to ensure that everyone can take part (with the 2011 Census having to spend £7 million on creating their own address file due to restrictions on PAF). There are also growing requirements for addresses in fields such as navigation, mobile application development and the delivery of crucial public services (everything from getting an ambulance to the right place to planning where to build much needed social housing).

Without an Open address register, much of the benefit of Open Data is lost. This is why ODUG used its Release of Data Fund to help the ODI pilot the first phase of a program to create an Open Address Register²³ with further funding confirmed in November 2014.

Evidence presented to the committee indicates a great deal of positivity but a real need to

¹⁹ <http://www.landregistry.gov.uk/market-trend-data/public-data/price-paid-data>

²⁰ <http://data.gov.uk/benefits-of-releasing-an-open-vat-register>

²¹ <http://us6.campaign-archive2.com/?u=e7311d49e9ac144a359ee2a96&id=206953c7ad>

²² <http://www.parliament.uk/business/committees/committees-a-z/commons-select/public-administration-select-committee/news/open-data-substantive/>

²³ <http://theodi.org/news/383k-government-grant-released-to-create-uk-open-address-list>

change attitudes in certain sectors of Government from one of caution to one of support.

Common themes in the submissions made include short term costs to release data, loss of revenue where data is currently sold (for example, Royal Mail PAF or OS MasterMap) and fears from privacy lobbyists and sections of the media (for example, how anonymous can you make medical records so that they are useful but not an invasion of privacy?). However, this final point is more to do with Data Sharing (which can be the sharing of private, personal or sensitive data for specific uses rather than free and unrestricted use) than Open Data – another area of confusion that needs to be addressed.

Recent moves by the EU to introduce new Data Protection regulations across the bloc (with added impetus provided courtesy of the NSA hacking scandal) could create tensions between transparency advocates and the privacy lobby. Careful education of citizens and politicians is required.

Closer to home, ODUG will focus on the following key areas:

1. The National Information Infrastructure:²⁴ Announced in October 2013, the NII was an attempt by the UK government to ‘codify’ what should be included in a basic toolbox of data assets that could be used as the backbone for all data-based decisions. For example, this ‘infrastructure’ could be obvious datasets on transport networks, costs and timetables or more fundamental, such as the basic locations of all public services in the country. After some slow progress, the ODUG and Cabinet Office are working to re-invigorate the programme to define what should form the basis of our national Open Data toolkit.

2. Open Address Register:²⁵ Thanks to a ‘Release of Data Fund’ managed by ODUG, work is beginning to create an Open and free address register to form the backbone of the NII and all other open data opportunities. This is being led by the ODI.

This doesn’t mean that work on data requests and other funding requests will stop. ODUG are regularly supporting²⁶ agencies and NGOs to release and make use of Open Data²⁷ and will continue to hold the public sector to account whenever the call to ‘give us our data’ is resisted.

The real task now is for businesses, academia and public sector experts to unite to put across the strongest case for more anonymous open data to be released but with the strongest possible sanctions against its use for nefarious means. The UK is lucky; ODUG, ODI, Transparency Board and large network of experts means we are well placed to make the case to do more and take balanced risks to benefit society.

²⁴ <http://data.gov.uk/blog/the-national-information-infrastructure-where-are-we>

²⁵ <http://theodi.org/blog/open-addresses-discovery-phase>

²⁶ <http://data.gov.uk/blog/funding-agreed-important-new-open-data-projects>

²⁷ <http://data.gov.uk/blog/release-data-fund-update>

About the author

Paul Malyon is a Product Manager for the global information services organisation, Experian. He is also a member of the Open Data User Group and is working to support businesses and society with the release of new sources of Open Data. He is a regular speaker on the topic and blogs at <http://dataisaproduct.blogspot.com>

Licence and Attribution

This paper was published in the International Free and Open Source Software Law Review, Volume 6, Issue 1 (December 2014). It originally appeared online at <http://www.ifosslr.org>.

This article should be cited as follows:

Malyon, Paul (2014) 'An Introduction to the Open Data User Group', *International Free and Open Source Software Law Review*, 6(1), pp 51 – 60

DOI: 10.3366/ifosslr.v6i1.99

Copyright © 2014 Paul Malyon.

This article is licensed under a Creative Commons UK (England and Wales) 2.0 licence, no derivative works, attribution, CC-BY-ND available at

<http://creativecommons.org/licenses/by-nd/2.0/uk/>

As a special exception, the author expressly permits faithful translations of the entire document into any language, provided that the resulting translation (which may include an attribution to the translator) is shared alike. This paragraph is part of the paper, and must be included when copying or translating the paper.



