

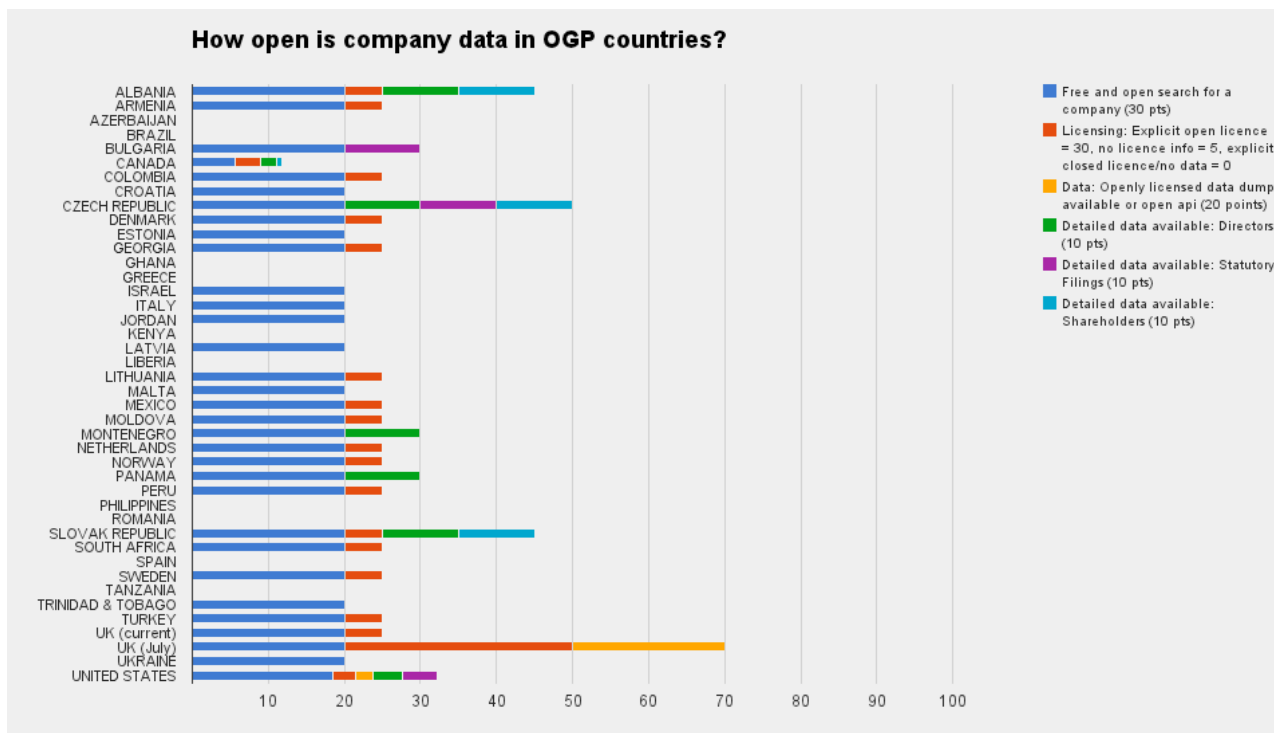
The Closed World of Company Data

An examination of how open company data is in Open Government Partnership countries

Executive Summary

Open Government Partnership countries score badly for open access to company data

- “Increasing Corporate Accountability” is one of the Open Government Partnership’s 5 Grand Challenges
- Accountability requires access to information, and in a complex corporate world that means access to it as machine-readable data, with a licence to reuse, remix, combine with other data
- The information on companies is collected for a statutory purpose, as companies are by definition artificial entities that have been created by the state for the wider benefit of society, but is often restricted to only those who have the funds to buy the data, usually companies using it to enhance proprietary datasets, thus doubly undermining access.
- OGP countries score badly for open access to company data, with an overall average of 21 (out of a possible 100 points), with several important countries (Spain, Greece, Brazil) scoring 0 points



Virtually all OGP countries score very badly for openness of company data, with several – including countries such as Spain, Greece and Brazil – effectively closed for the public, civil society and the wider world, undermining corporate governance, and providing a fertile ground for corruption, money laundering, organised crime, and tax evasion

Background

Why open company data matters

Equal access to core information is a fundamental principle of both free markets and an informed, active, functioning democracy, and in a world where our lives are not just governed by data, but increasingly are data, this means free and open access to the core datasets that affect our lives, and the ability to use that data, to analyse it and combine it with other data.

The world is also now increasingly a corporate one, dominated by highly interconnected global companies that have exploded in complexity and power, and the ability to affect our lives for good and bad. Yet in tandem with this growth, fueled by globalisation and the internet, corporate governance has decreased, as the company stakeholders (shareholders, workers, suppliers, customers, society at large) struggle to understand companies, still less influence them. This problem is accentuated for those supra-national corporations that are both everywhere and nowhere (as witnessed by the question of whether BP is a British or US company following the Gulf oil disaster), and thus effectively beyond the scope of any one country.

In addition, as the World Bank's "Puppet Masters" report¹ makes clear, this complex corporate world, and lack of access to good data about corporate entities, is a key enabler of corruption, organised crime and money laundering.

In fact these problems were recognized by the inclusion of "Increasing Corporate Accountability" as one of the Open Government Partnership's five grand challenges². But accountability is not possible without access to information, and, in this world where increasingly power comes from the the ability to analyze, reuse and combine large datasets with other large datasets, that means access to the information as data in a machine-readable form under an open licence that allows reuse.

Yet despite this, this survey and analysis of access to core company data in the Open Government Partnership countries shows that almost without exception they perform badly when scored on access to core company data. In addition, anecdotally few OGP countries seem to include increasing corporate transparency in their country plans/commitments for future action.

¹ http://www1.worldbank.org/finance/star_site/documents/Puppet%20Masters%20Report.pdf

² http://www.opengovpartnership.org/ogp-participation#Grand_Challenges

What is 'core company data'?

Companies are artificial entities (as opposed to 'Natural Persons', i.e. people) given legal personality by the state for the wider benefit of society, but in order for this to work it's important that the public – employees, suppliers, customers, government – can answer the question: does this company exist, and what are the core, statutory facts that define it. This is what we mean by 'core company data': the minimal set of data about a company that is needed to for free, open, fair and stable markets to exist, and for society to understand, influence, and, where necessary, regulate the behaviour of companies.

We would suggest that 'core company data' should including the following:

1. Verification that the company exists, its current legal name, the jurisdiction in which it is legally situated, the ID given by that jurisdiction (sometimes called a company number), the current status of the company (e.g. active, dissolved, etc), the type of entity (e.g. PLC, SA, LLC, etc), the incorporation date, and the registered address (at which legal service can be made).
2. Directors and officers of the company – these are the individuals and sometimes other corporate entities who direct and run the company, and who are legally responsible for it.
3. Statutory filings – the official filings that must be made, for example, to report significant events in the life of a company, e.g. annual reports, annual accounts, change of directors, change in name, increase in share capital. Increasingly these are collected as data, including annual accounts (e.g. as XBRL), but even where these are paper- or image-based these should be considered as core company data
4. Significant shareholdings and corporate hierarchy relationships. This information is essential to understand the company – for example, what's the parent company or companies, and, potentially, the beneficial owner, critical in anti-corruption and anti-money-laundering.

Detailed results: the average, the bad, & the completely closed

Overview

The detailed figures for OGP countries are overleaf, but one look at the overall average score (20.8 out of 100) shows how closed OGP countries' company registers are. None currently scores over 50, although the UK will do in July when it starts publishing a regular dump of a limited dataset under the Open Government Licence. Many are effectively completely closed, not even allowing searches for companies without payment.

The average

No OGP country company register scores well in this survey, the best being the UK (from July), ironically followed by countries (the Czech Republic, Albania, Slovak Republic) that have a much shorter recent history of democratic government than many of the other countries with much lower scores.

The bad

The next group of countries, including the US (*see below*), all score 32 or less, in general allowing little more than free online searching. Little extra information is made available, there is no licence that allows for reuse, and often an explicit closed licence that prevents reuse to some degree. Within this group the level of company data returned by search varies from basic (generally lacking essential data such as directors, statutory filings and significant shareholders) to very sparse, sometimes with just a name, company type and status.

The completely closed

Finally, we get to those countries that are completely closed, preventing even searches of companies without charge and without registration. These countries, which include Spain, Greece and Brazil³, appear to treat the registers, not for the public benefit, but as an extension of the global business information services (Dun & Bradstreet, Bureau van Dijk, Experian, etc), who buy the data in bulk and resell to their corporate customers.

The justification – that they can find entities who will pay for this data – misunderstands both the role of data and companies in the modern world, and the nature of the public good. There are many things the state could charge for (the police force, street lights, even the right to vote), but doing so would undermine both the state's legitimacy and the wider society. Similarly with company data – this is data collected for a statutory purpose for the public benefit and should be treated accordingly – and to bring the old truism up to date, in the 21st century “Data Is The Currency Of Democracy”.

³ Brazil has a central register that aggregates the regional registers, and it's on this register we have scored Brazil

	Free & open search (20 points)	Licensing Open = 30, no licence = 5, closed = 0	Available as (open) data (20 points)	Directors information	Statutory Filings	Shareholders information	Total (out of 100 points)
ALBANIA	20	5	0	10	0	10	45
ARMENIA	20	5	0	0	0	0	25
AZERBAIJAN	0	0	0	0	0	0	0
BRAZIL	0	0	0	0	0	0	0
BULGARIA	20	0	0	0	10	0	30
CANADA	6	3	0	2	0	1	12
COLOMBIA	20	5	0	0	0	0	25
CROATIA	20	0	0	0	0	0	20
CZECH REPUB.	20	0	0	10	10	10	50
DENMARK	20	5	0	0	0	0	25
ESTONIA	20	0	0	0	0	0	20
GEORGIA	20	5	0	0	0	0	25
GHANA	0	0	0	0	0	0	0
GREECE	0	0	0	0	0	0	0
ISRAEL	20	0	0	0	0	0	20
ITALY	20	0	0	0	0	0	20
JORDAN	20	0	0	0	0	0	20
KENYA	0	0	0	0	0	0	0
LATVIA	20	0	0	0	0	0	20
LIBERIA	0	0	0	0	0	0	0
LITHUANIA	20	5	0	0	0	0	25
MALTA	20	0	0	0	0	0	0
MEXICO	20	5	0	0	0	0	25
MOLDOVA	20	5	0	0	0	0	25
MONTENEGRO	20	0	0	10	0	0	30
NETHERLANDS	20	5	0	0	0	0	25
NORWAY	20	5	0	0	0	0	25
PANAMA	20	0	0	10	0	0	30
PERU	20	5	0	0	0	0	25
PHILIPPINES	0	0	0	0	0	0	0
ROMANIA	0	0	0	0	0	0	0
SLOVAK REPUBLIC	20	5	0	10	0	10	45
SOUTH AFRICA	20	5	0	0	0	0	25
SPAIN	0	0	0	0	0	0	0
SWEDEN	20	5	0	0	0	0	25
TANZANIA	0	0	0	0	0	0	0
TRINIDAD & TOBAGO	20	0	0	0	0	0	20
TURKEY	20	5	0	0	0	0	25
UK (current)	20	5	0	0	0	0	25
UK (July)	20	30	20	0	0	0	70
UKRAINE	20	0	0	0	0	0	20
UNITED STATES	19	3	2	4	5	0	33
OVERALL AVERAGE	14.9	2.8	0.5	1.3	0.6	0.7	20.8

A couple of countries – Kenya and the Philippines – in theory have online company registers, but both have been non-functioning for some time, in the case of Kenya, for at least six months.

The US and Canada

In the United States companies are registered by the individual states, and each state has its own register, and indeed company legislation. This has led to competition between the registers to attract companies, reducing the regulation and reporting requirements in some states to a lower level than some traditional tax havens. This ‘race to the bottom’ is a well-known effect of the ability of companies to jurisdiction-shop, and could easily result in access to company data, and with it the ability to influence and supervise companies, being significantly reduced in the coming years from its already low level.

As it stands, access to company data from the state registers is poor, with relatively meagre data being collected, restrictive licences being common (Illinois actually has a specific law to prevent company data being distributed for free), little open access to data, and the public benefit being forgotten for the sake of selling to data to commercial buyers. Even the District of Columbia, which has a cultivated a reputation for open government, scores badly, restricting searches to registered users. Texas, alone of the US states, restricts even basic searches to paying customers.

Canada also has local registers, many of which are not even online, making search practically impossible. There is also a federal company register, for national companies, and here the situation is rather better, with free access to both basic company data and company directors.

The United Kingdom

The United Kingdom has made a public commitment to regularly publish the most basic information from its company register (i.e. name, company number, company type, registered address, status) as an openly licensed bulk download from July 2012. We have therefore rated the UK twice, once for the current situation, and once for that in July. It has also said that it is looking at publishing directors’ information as open data later in the year, but has as yet made no commitment to this, nor to statutory filings, shareholdings or to company accounts (which it is now collecting as data).

In March the UK became the junior chair of the Open Government Partnership, and in September the senior chair. It remains to be seen whether it uses this as an opportunity to lead by example, extending its open data to directors, filings and shareholders, particularly given its important position in the financial and corporate world.

Depth and quality of information available

As described above, company registers often return very little data for searches for companies, making it difficult for customers, suppliers and other interested parties to even understand the nature and status of the company. In addition, some of the registers have problems with quality and timeliness of the data – Mexico’s register, for example, is only updated every year by the companies on it, and seems far from authoritative.

Finally, as with any large dataset, there are bound to be errors, either due to bad data being supplied by companies, errors in data entry, or systematic issues. The best way of identifying such errors is to reuse the data and combine with other datasets – of course this requires the information to be available for reuse by a wider variety of users, ideally in a machine-readable form, and with a licence that allows for reuse.

The role of company registers

The role and legal position of the company register varies from country to country, in most being a government agency that has a statutory role to create and incorporate legal entities, and keep a record of those entities in a

register. However, this is not universal, and some are registers of legal entities that are created by another arm of the state (for example, courts), which can create problems of accuracy or timeliness. Finally, there are issues of governance, 'business models', incentives and legal status: some countries use non-departmental public bodies (e.g. the Netherlands' Chambers of Commerce), others quasi-publicly-owned businesses (UK Companies House is a Trading Fund⁴), and at least one country (Latvia) outsources the role of company register to a private company.

How should this be done?

An example of a different approach to company data comes from New Zealand, where all the information on the company register is open for all, without charge, without registration, and without significant restrictions. In addition, there is a free API which allows access to and reuse of the underlying data (registration is required, in part because it can be used for making filings for companies, which are charged for). In addition, the entries on the register are much more detailed than on all the OGP countries we have examined, including directors, statutory filings and significant shareholders.

Corporate relationships and the myth of competitive advantage

A modern corporation is no longer a single legal entity, nor even a simple hierarchy with a PLC, Inc, or SA at its head, but a complex web of sometimes thousands of different corporate legal entities. This has very significant implications, from the ability to understand the overall behaviour of a corporation, to issues of liability and responsibility (which parts of the corporation are liable and/or responsible for an action, and does that liability pass on to the other parts of the corporation), to the ability to tax corporations on profits, to understanding and regulating systematic risk (as seen with the collapse of Lehman Brothers).

We can think of no legitimate reason why a corporation's constituent parts should not be part of the public record. According to a paper by Linda Powell et al of the Federal Reserve Board⁵, "A common reason provided for why a relationship between entities should not be public is that releasing the ownership data to the public could harm the competitive advantage of an entity."

In fact, this provides a powerful reason for why corporate relationships should be made public.

First, this is not about legitimate commercial confidentiality but about protecting incumbents' power, position and, often, legacy business models.

Second, this is often as not about evading scrutiny – from society, from shareholders, and from regulators. Make it difficult enough to understand the hundreds or thousands of complex legal entities making up a global corporation, and they can do what they want.

Third, competitive advantage should be about coming up with a better way of doing something, taking a risk on a new product and reaping the benefits, having a closer relationships with your customers, and so on. Not obscuring the public record so that it's impossible for the wider world to understand who they're doing business with, competing with, investing in or regulating.

Fourth, this lack of transparency, enabled in part by company registers restricting information about companies to those who can afford the high prices, forces all companies to play the dodge the regulator/tax/public scrutiny game. This reduces the attention spent by good companies on real-world innovation and being a good

⁴ http://en.wikipedia.org/wiki/Trading_fund

⁵ http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1956664

corporate citizen, and forces all into a game of regulatory arbitrage, playing jurisdictions and regulators against each other to force them to a race to the bottom that leaves companies able to act outside the law.

Other global initiatives

There are a number of initiatives around the world looking at the question of core company data, from different angles. Some are looking at identifiers for corporate entities, some looking at beneficial ownership (the individuals that ultimately own companies), some at corruption, or the activities of companies in developing countries, particularly in the extractive industries, and others normalising the information about companies to allow data to be more easily combined.

G20/Financial Stability Board global legal entity identifier

The Financial Stability Board has a programme⁶ for a global legal entity identifier (an id that can represent a company anywhere in the world) on behalf of the G20, and this is due to be ready to announced by the next G20 meeting in June. This is particularly focused on understanding the interconnectedness of companies, to reduce systematic risk, particularly following the financial crisis and the collapse of Lehman's. *[Disclosure: OpenCorporates is on the Advisory Panel for this programme].*

EU/W3c ISA Business Vocabulary

The EU has several programs around company data, including the creation of a Business Vocabulary⁷, jointly with the W3c to make it easier to share company data. This specifies standard terms for core company attributes (in particular how to identify a company in a register), an important step for being able to combine information about companies *[Disclosure: OpenCorporates helped develop this vocabulary].*

World Bank

The World Bank has a number of programmes in this area, including its international corruption and Stolen Asset Recovery teams, this last one publishing a comprehensive and influential report last year on how the lack of transparency of company data is fueling corruption, organised crime and undemocratic regimes. It is also involved in the IATI programme⁸ for international aid transparency.

Beneficial owners

There is also work going on in the US to pass a bill to publish the beneficial owners of companies, important in reducing corruption, tackling money-laundering, and make the subversion of campaign contribution rules more difficult.⁹ This also has implications for money flows to and from the developing world and is supported by the

⁶ http://www.financialstabilityboard.org/list/fsb_publications/tid_156/index.htm

⁷ https://joinup.ec.europa.eu/asset/core_business/release/02

⁸ <http://www.aidtransparency.net/>

⁹

<http://www.levin.senate.gov/newsroom/press/release/levin-grassley-introduce-bill-to-combat-us-corporations-with-hidden-owners>

Task Force on Financial Integrity and Economic Development, a consortium of governments and research and advocacy organizations working for increased transparency in financial transactions with the developing world¹⁰.

Global Initiative for Fiscal Transparency

The Global Initiative for Fiscal Transparency¹¹ is being formally launched at the Open Government Partnership meeting in April, and is a multi-stakeholder network working to improve fiscal transparency, participation, and accountability in countries around the world. Although corporate information and corporate identifiers isn't a core goal or metric, there are clear overlaps with this initiative, not least because a critical question when looking at government spending is: who got the money?

¹⁰ <http://www.financialtaskforce.org/>

¹¹ <http://fiscaltransparency.net/>

OpenCorporates: a global open data success story

From 3 million to 40 million companies in just 15 months

[OpenCorporates](#) is by far the largest openly licensed database of companies in the world, with over 40 million distinct corporate legal entities in over 50 jurisdictions. It has its origins in the UK open data movement, and as a direct result of the UK government's open data policy, when one of its co-founders, Chris Taggart, faced the problem of matching the companies in the newly published local council spending data to legal entities.

He realised this was not just a UK problem, but a global one, and he and co-founder Rob McKinnon, another open-data pioneer, set out to build an an openly licensed database with one core feature: an entry for every corporate legal entity in the world, and an associated open URI identifier in which OpenCorporates claimed no IP rights, allowing it to be freely and openly used around the world. Since then it has grown from the original 3 million companies in 3 jurisdictions to over 40 million in over 50 jurisdictions. It is also increasingly adding other public data and matching it to those legal entities, from government spending and trademarks to official government supplier lists and health & safety violations.

OpenCorporates is now an integral part of the open data world and has been praised by European Vice-President Neelie Kroes as “the kind of resource the (Digital) Single Market needs”. It has also has been heavily involved in many governmental efforts to increase the quality and openness of company data, including helping design the new UK Companies House URIs, helping the EU with its Business Vocabulary, and is on the advisory panel for the Financial Stability Board's global legal entity identifier programme for the G20.

Its co-founder, Chris Taggart, is on the Access To Information panel at the OGP meeting in Brasilia, and will be talking about this subject, and this report.

Notes on results/methodology

Overview

One of the challenges in producing this report is that good information about company registries is short supply – there's not even a comprehensive directory of company registries¹². In addition, company registers, despite their importance (they are or at least are supposed to be the canonical reference source for the legal information about companies) the registers themselves are very difficult to find either from central government sites or from search engines. Once you've found them, they are often difficult to use, unclear about the depth, quality, and legal status of the information, with much and sometimes all the data hidden from view from visitors.

However, despite these obstacles, with the help of the open data community, we visited over 110 individual company registers over the course of approximately 10 days in April 2012 to examine access to the data and the ability to reuse it.

Scoring

We have taken the view that company registers are public registers, the public record of artificial entities given legal personality by the state for the benefit of society. In a free and open society this important information should be free to use and reuse for all, without charge and without restrictive licence conditions. While we believe this material should be in the public domain, with no licence restrictions at all for this survey we have scored based on whether the licence complies with the generally accepted Open Knowledge definition¹³.

In addition in a world where power comes from the ability to access to data and reuse it, the underlying data should be available to all without fee and without restrictions as machine-readable data – otherwise you are restricting access to this important and empowering dataset only to those with the resources to pay, thus undermining scrutiny, innovation, and a free and open market.

In some countries – specifically US and Canada – companies are registered under the laws of individual states or provinces (Canada also has federal companies, registered by the federal register). In this case the individual registers were scored and the data averaged. For this report we are showing the average figures, but the individual scores for the US states and Canadian provinces and territories are published online¹⁴. We use a mean when calculating the average, although arguably, given that understanding of a country's companies requires full and equal access to all the areas the lowest ranking scores should be given greater weight. We would be happy to work with anyone who would like to pursue this area of study.

¹² There are online directories but all have considerable gaps in their information. Notable ones include: <http://www.rba.co.uk/sources/registers.htm>, <http://www.companieshouse.gov.uk/links/introduction.shtml#reg>, <http://www.commercial-register.sg.ch/home/worldwide.html> and http://en.wikipedia.org/wiki/List_of_company_registers

¹³ <http://opendefinition.org/okd/>

¹⁴ http://bit.ly/ogp_open_company_data

We therefore scored on the following basis (with a total of a possible 100 points):

- **Basic data online and available to search without charge or registration.** This is the base minimum threshold for an open company register. In short, if you can't search for a company without charge, restriction or registration (which implies restriction) they are essentially closed to the public. In the case where there are central registers that aggregate local regional registers (as in the case of Brazil, or Spain), we score on this central register rather than the regional registers.

20 points

- **Licensing** Is there an explicit open licence (e.g. CC-0, Open Government Licence)?

30 points for a licence that conforms to the Open Knowledge definition; **5 points** for no explicit licence;

0 points for a licence that explicitly prevents reuse or otherwise fails to conform to the Open Knowledge definition, including catch-all closed licences (e.g. All Rights Reserved).

- **Availability as data** Is the information freely available as data, either as an openly licensed data dump or via an openly licensed API?

20 points

- **Depth of data** How detailed is the data that is published on the register (specifically published for access to all without additional charges or restrictions).

10 points each for directors/officers, statutory filings (i.e. can the filings be downloaded or available as data), significant shareholdings, bringing a maximum total of 30 points.

Licence/data/errors/omissions

This document is openly licensed

This document and the information contained in it is published under the Creative Commons Share-Alike Attribution Licence (<http://creativecommons.org/licenses/by-sa/3.0/>), allowing it to be freely reused. Attribution should be to OpenCorporates with a hyperlink to the OpenCorporates website (<http://OpenCorporates.com>), where such a link is possible (e.g. web pages, PDFs, Word documents).

In addition, the underlying data is licensed under the Share-Alike Attribution Open Database Licence (<http://opendatacommons.org/licenses/odbl/>). Attribution should similarly be to OpenCorporates with a hyperlink to the OpenCorporates website (<http://OpenCorporates.com>), where such a link is possible (e.g. web pages, PDFs, Word documents).

Data

The results of this survey are available at http://bit.ly/ogp_open_company_data and can be downloaded as data from this same URL.

Errors/omissions

Any survey of this breadth may contain errors, and it is OpenCorporates' policy to correct errors when they are brought to our attention, updating this document and explicitly stating where changes have been made. In any survey of this kind there are also going to judgement calls – for example, whether a jurisdiction's officer and directors information is complete enough to qualify, or whether a licence is open or closed. In general we have erred towards being generous in the scoring, but there may also be cases where the judgement on reflection may be wrong, or where we have made the judgement was made in the absence of explicit information (e.g. an obvious licence page, or underlying law).

There are also omissions, some known, others in error. In particular, we failed to find the company registries for a handful of the Open Government Partnership countries, but given the difficulty in finding them it seems likely that inclusion of these would if anything lower the overall score. We would, however, be pleased to accept information on these.

Thanks

Thanks to the many people in the open data community who helped compile this data, or gave feedback on the methodology or the report.

Errata/Updates to this report

- 18 April 2012: Corrected typo on scoring. Basic online information scores 20 points, not 30 points. Individual scores/totals were all correct, just the words were incorrect.

- 3 May 2012: Page 6/7. Clarified section on role of company registers. The Netherlands Chambers of Commerce, despite the name, is a non-departmental public body, not a private company, added info about Companies House being a Trading Fund.
- 14 June 2012: Page 1. Italy was wrongly credited with having a zero score in the summary; the tables and charts correctly scored them as having 20/100.
Page 4: Typo. 'The next group of companies...' should have read 'The next group of countries...'