

Data requests from Open Data User Group request process

Paper for Tax Transparency Sector Board, 25th February

Summary

1. This paper sets out the requests for HMRC to release open data received through the Open Data User Group request process on Data.gov.uk.

Background

2. The Open Data User Group began a call for open data users to request data from Government in October 2012. This is hosted as a request form on Data.gov.uk. Any member of the public is able to put a request in for a dataset they would like to see released as Open Data. Requesters are asked to complete as much information as they have available on the data they are requesting. The request is then published on Data.gov.uk although they also have the ability to make a request confidentially.
3. These requests are initially analysed by the Cabinet Office Transparency Team and those which are identified as being of most value or which are likely to involve negotiations relating to building a more detailed case for release are referred on to the Open Data User Group for consideration.
4. Where cases are seen as being relatively straightforward these are referred by the Transparency Team departmental account manager straight to working level contacts in Departments to investigate whether the data can be released.

Requests for HMRC data

5. There have so far been eight requests received through this process that have been identified as relating to HMRC held data. Those six requests which have been submitted publicly are set out in Annex A which are for:
 - a. CHIEF database
 - b. Anonymised dataset of UK taxpayers
 - c. VOA business listings
 - d. Small area wealth or income data
 - e. Companies Tax Records
 - f. The London Rent Map Raw Data
6. A number of these are being pursued at a working level and one has been raised to the Data Strategy Board (VAT Register – see Annex B).

Considerations for the Board

7. The board may wish
 - a. to give their thoughts on the content of these requests to inform the prioritisation of the work that HMRC is undertaking to understand whether the data can be released; and
 - b. to consider how they would like to remain updated on an ongoing basis in terms of new requests that come in via this process.

Ed Parkes, Cabinet Office, Open Data and Transparency Team

Annex A: Public data requests for HMRC held data

| Data Name | | CHIEF database | |
|-------------------|---|----------------|--------------------|
| Post date | 26/09/2012 - 11:05 | Date Reformat | 26/09/2012 - 11:05 |
| Status | ODUG developing business case for data release | | |
| Action | 6 | | |
| DSB | No | | |
| Comment | | | |
| Data Description | <p>Import/export database, including detailed information about individual imports/exports.</p> <p>See http://customs.hmrc.gov.uk/channelsPortalWebApp/channelsPortalWebApp.portal?_nfpb=true&_pageLabel=pageOnlineServices_ShowContent&id=HMCE_CL_001614&propertyType=document</p> | | |
| Data Holder | HMRC | | |
| Submitter Name | Chris Taggart | | |
| Organisation Type | Small to Medium Business | | |
| Organisation Name | OpenCorporates | | |
| Data Themes | Environment, Finance, Administration | | |
| Reason | There are financial charges for the data | | |
| Further Detail | Given that the data is sold, and that the ICO has said that there are no Data Protection issues that are engaged for public publishing that are not engaged in sold data, there would not seem to be any issue about publishing this as open data | | |
| Suggested Use | Business UseCommunity WorkResearch | | |
| Further Detail | This is a critical database in understanding the flows of goods in and out of the UK, and | | |
| Benefits Overview | Allows wider societal understanding of which companies are sourcing in the UK and Europe, versus those importing from further afield. This has implications and benefits for local sourcing, sustainability, carbon, air miles and many more, giving consumers a much greater understanding, and greater opportunity to influence these issues than they currently have. In short, this is an important piece of the puzzle in understanding and positively influencing corporate behaviour | | |

| | | | |
|------------------|---|--|--|
| Data Name | Anonymised dataset of UK taxpayers | | |
|------------------|---|--|--|

| | | | |
|------------------|--------------------|----------------------|--------------------|
| Post date | 14/10/2012 - 21:31 | Date Reformat | 14/10/2012 - 21:31 |
|------------------|--------------------|----------------------|--------------------|

| | |
|----------------|---|
| Status | Being considered |
| Action | 10 |
| DSB | No |
| Comment | COTTAM HMRC to identify privacy issues; Need to develop stronger benefits case with requester |

| | |
|-------------------------|---|
| Data Description | Age, gender, geographical location (postal area?) amount of income tax paid, amount of... |
| Data Holder | HMRC |

| | |
|--------------------------|--------------------|
| Submitter Name | James Macfarlane |
| Organisation Type | Private Individual |
| Organisation Name | |

| | |
|--------------------------|--|
| Data Themes | Finance |
| Reason | |
| Further Detail | |
| Suggested Use | Personal UseResearch |
| Further Detail | We can find out interesting statistics on the distribution of wealth, etc. |
| Benefits Overview | |

| | |
|------------------|------------------------------|
| Data Name | VOA business listings |
|------------------|------------------------------|

| | | | |
|------------------|--------------------|----------------------|--------------------|
| Post date | 16/10/2012 - 09:21 | Date Reformat | 16/10/2012 - 09:21 |
|------------------|--------------------|----------------------|--------------------|

| | |
|----------------|---|
| Status | ODUG developing business case for data release |
| Action | 6 |
| DSB | No |
| Comment | May form part of a wider approach towards VOA data |

| | |
|-------------------------|---|
| Data Description | Full UK Business Ratings full listings including business classifications, rateable values, buildings area, business names and addresses. |
| Data Holder | VOA in England and Wales, the LPSI in Northern Ireland and the SAA in Scotland |

| | |
|--------------------------|-----------------|
| Submitter Name | Blair Freebairn |
| Organisation Type | Start up |
| Organisation Name | GeoLytix |

| | |
|--------------------------|---|
| Data Themes | Health, Education, Finance, Society, Transportation, Location, Policy |
| Reason | There are financial charges for the data |
| Further Detail | The financial charges just edges out licensing, publishing format and lack of co-ordination between the countries as reasons. |
| Suggested Use | Business UseResearch |
| Further Detail | The main use is to build accurate sets of POI for use in site location research and hyper-local marketing. |
| Benefits Overview | The data is richly attributed with size (essential for site research), names, business type etc. This data is simply not generally available anywhere else. For the marketing use case the comprehensiveness and accuracy of addressing is vital. |

| | |
|------------------|---|
| Data Name | Small area wealth or income data |
|------------------|---|

| | | | |
|------------------|--------------------|----------------------|--------------------|
| Post date | 02/11/2012 - 12:14 | Date Reformat | 02/11/2012 - 12:14 |
|------------------|--------------------|----------------------|--------------------|

| | |
|----------------|--------------------------------------|
| Status | Chair will write to publisher |
| Action | 11 |
| DSB | No |
| Comment | |

| | |
|-------------------------|---|
| Data Description | Some sort of analysis of wealth (not deprivation) at small area (OA level for preference). This could be something like average gross taxable income, average untaxed state benefits and any other useful additions |
| Data Holder | HMRC should be able to get to the income data anonymised via post codes from tax returns, DWP should have benefits data on a similar basis from awards. |

| | |
|--------------------------|---|
| Submitter Name | Martin How |
| Organisation Type | Voluntary sector or not-for-profit organisation |
| Organisation Name | Diocese of London (Church of England) |

| | |
|--------------------------|---|
| Data Themes | Finance, Society |
| Reason | Other |
| Further Detail | This data does not seem to be available at small area. |
| Suggested Use | Community Work |
| Further Detail | Its an important part of understanding the detail of complexity and diversity of the area that we serve. Some of this data is available at Local Authority level, but that is too crude and limited to be useful in communities that comprise 1/20th of the LA, and have high levels of diversity in adjoining communities. The LSOA deprivation data hints at this, but Deprivation data can not be reverse engineered to give wealth. Table 3.14a Income data by Local Authority provides indicative income per head for those in the tax system, but no information on household size, numbers of people paying tax in the cohort etc. |
| Benefits Overview | We use deprivation data to plan and target mission resources (spending), however the absence of wealth data makes it difficult to have a proper understanding of the financial resources available within various communities to support work in deprived areas and to achieve a fair basis for cost sharing (balancing income) across London. |

| | |
|------------------|------------------------------|
| Data Name | Companies Tax Records |
|------------------|------------------------------|

| | | | |
|------------------|--------------------|----------------------|--------------------|
| Post date | 07/11/2012 - 11:56 | Date Reformat | 07/11/2012 - 11:56 |
|------------------|--------------------|----------------------|--------------------|

| | |
|----------------|-------------------------|
| Status | Being considered |
| Action | 2b |
| DSB | No |
| Comment | |

| | |
|-------------------------|--|
| Data Description | I'd like to see all data released in an easily accessible way for each company registered in the UK and how much tax they pay. |
| Data Holder | HMRC and Companies House. |

| | |
|--------------------------|--------------------|
| Submitter Name | Jake Lea-Wilson |
| Organisation Type | Private Individual |
| Organisation Name | |

| | |
|--------------------------|---|
| Data Themes | Finance, Society, Administration |
| Reason | The data is published but not in a format I can download and use (e.g. only displayed onscreen or only downloadable as a PDF rather than CVS) |
| Further Detail | |
| Suggested Use | Business UsePersonal UseCommunity WorkResearchOther |
| Further Detail | A searchable database that lets anyone access the information about how much tax each company pays. |
| Benefits Overview | This would benefit society because it would allow people to make informed decisions about each company they deal with. |

| | |
|------------------|-------------------------------------|
| Data Name | The London Rent Map Raw Data |
|------------------|-------------------------------------|

| | | | |
|------------------|--------------------|----------------------|--------------------|
| Post date | 30/11/2012 - 17:22 | Date Reformat | 30/11/2012 - 17:22 |
|------------------|--------------------|----------------------|--------------------|

| | |
|----------------|-------------------------|
| Status | Being considered |
| Action | 2b |
| DSB | No |
| Comment | |

| | |
|-------------------------|---|
| Data Description | The Valuation Office Agency (VOA) data used to produce the London Rents Map is based on market rents from private rented sector lettings. Landlords, agents and tenants across England contribute this information. It is mainly used to help Rent Officers provide valuations for Housing Benefit purposes and Rent Act 1977 Fair Rent registrations. Rent Officers continuously evaluate and refresh lettings data to provide a representative sample of the private rented sector, and track the market. |
| Data Holder | The Valuation Office Agency, Mayor of London's Office |

| | |
|--------------------------|--------------------|
| Submitter Name | Racheet Dave |
| Organisation Type | Private Individual |
| Organisation Name | |

| | |
|--------------------------|---|
| Data Themes | Finance, Location |
| Reason | The data is published but not in a format I can download and use (e.g. only displayed onscreen or only downloadable as a PDF rather than CVS) |
| Further Detail | The data is provided as some sort of scrollable and searchable map. Unfortunately, I'm looking specifically for the rents in places along a specific tube line. I'd really like the raw data available from a JSON API so that I can query it and mash it up with other APIs that would tell me how desirable a given area was to live in, it would really make my house move a lot easier to plan if I had access to the data I needed to plan it. |
| Suggested Use | Personal Use |
| Further Detail | I'm looking to make a web-app that would let me actually plan my house move by mashing up information on the average rents in an area with information about the crime rate, police spending, quality of schools etc. This is so that I can plan my upcoming house move to London. I'd be happy to make the web-app available to the public for free, but it would initially just exist to serve my own need. |
| Benefits Overview | Currently having to work out what different areas in London are like to live in, and how much they cost is dependent on knowing people already in london who can tell you. All you can get is anecdotal information about each area. Being able to get a stable api |

UNCLASSIFIED

| | |
|--|--|
| | <p>that allows you to programatically query a regularly updated database that shows what the median and upper quartile rents in arbitrary areas of london are would be incredibly useful for planning moves within London. Despite being a phenomenally useful dataset, it appears to only be available on the internet through the London Rent Maps website: http://www.london.gov.uk/rents/about/ That website has a map that you can search, but it isn't possible to take the data out and mash it up with other sources, which is a terrible shame.</p> |
|--|--|

Annex B – ODUK Benefits case for releasing and Open VAT Register

Benefits of releasing an Open VAT Register

Description & Request Overview

HMRC hold the dataset known as the VAT Register which contains information on businesses such as; VAT number, Business status (i.e. Sole proprietor, partnership, corporate, etc), Name and business address, Contact information – telephone, email, website, Business activities, Other business associations in the past 2 years, New registration or ceasing registration, Turnover, Date of Establishment, Applicants Name & Address.

Releasing some or all of this data will enable banks to make better lending decisions to increase the availability of trade credit.

Data Release Rationale

BIPA (Business Information Providers Association) have estimated that the VAT register could benefit UK PLC to the tune of around £50 billion. This is based upon an increase in credit requests and associated credit approvals powered by the positive data from the VAT Register.

The VAT Register will be positioned as an additional source of information to those provided by Companies House and member of BIPA. It will also be used by organisations such as the members of BIPA to add value to existing data products and services, further increasing the positive economic benefits for their customers.

This enhanced data could be returned to HMRC to improve the efficiency of tax collections to provide greater revenue and / or cost savings.

Businesses (including SMEs and lenders) will be able to comply more easily with anti-money laundering (AML) requirements. This creates positive business efficiencies and cost savings.

Commercial fraud may be reduced through the enhanced identification services enabled by the VAT Register. This has obvious economic & social benefits. In 2011, CIFAS members reported 250,000 instances of commercial fraud (a 9% year on year increase).

While the data should be released openly, some elements could be offered as an enhanced, paid for product by HMRC (similar to the Companies House model).

There are European Countries that already have a full business register including both incorporated and unincorporated businesses e.g. France and some that in addition use the VAT number as the basis of business identification e.g. Belgium.

Benefit description

BIPA (Business Information Providers Association) have estimated that the VAT register could benefit UK PLC to the tune of around £50 billion through improved access to funding for small businesses.

This is based upon an increase in credit requests and associated credit approvals powered by the positive data from the VAT Register.

The VAT Register will be positioned as an additional source of information to those provided by Companies House and members of BIPA.

It will also be used by organisations such as the members of BIPA to add value to existing data products and services, further increasing the positive economic benefits for their customers.

This enhanced data could be returned to HMRC to improve the efficiency of tax collections to provide greater revenue and / or cost savings.

UNCLASSIFIED

Businesses (including SMEs and lenders) will be able to comply more easily with anti-money laundering (AML) requirements. This creates positive business efficiencies and cost savings.

Please identify where further case studies or quantifiable evidence to support the release of this dataset?

Benefits analysis

Release of the VAT register as Open Data will:

Enable the release of business credit leading to growth of the UK Economy by up to £50 billion (BIPA estimate).

Improve access to business funding for 750,000 businesses.

Reduce the instances of commercial fraud (250,000 cases in 2011).

Enable the creation of new services such as 'trusted business' apps with the VAT Register acting as the unique data identifier.

Create internal government efficiencies through better data sharing and tax gathering.

Can you identify further areas where this dataset release will create opportunities for innovation and new business?

Can you help identify sectors, businesses and organisations that will benefit from the release of these data?

Barriers and Requirements for Release

It is assumed that the VAT Register data contains no information of a personal or sensitive nature.

Any sensitive data would need to be removed before release based upon advice from the ICO.

HMRC would need to make the data available on a regular basis (weekly or more regular is recommended).

HMRC would need to make the data available as a download or API.

While one of the main aims of the request is to reduce commercial fraud, risks posed by the data (in terms of enabling fraud) will need to be examined by Government with partners such as BIPA.

An 'Opt Out' scheme for Unincorporated business owners (similar to Edited Electoral Roll system) could prevent any negative perception of the data.

Consideration should be given to how the data will be used (e.g. marketing, credit or other).

Can you identify other barriers to this data release, or solution to those listed?

Recommendation

The business case for releasing the VAT Register is clear. Today, small business credit scores are approximately 40% lower than is expected due to a lack of positive data available to credit reference agencies (and in turn, lenders). This makes access to trade credit lower than it should feasibly be.

By providing the VAT number and additional details; business information providers will be able to create linked, up to date data assets to improve the chances of successful, well managed small businesses having access to the kinds of funding required to help them grow.

Without access to the VAT Register, business lending will remain sluggish and the chances of success for many small enterprises will be endangered.

The Open Data User Group therefore call on HMRC to release the VAT Register as an Open Database under the Open Government License.

[Login to post comments](#)



Forum of Private Business supports data release with safeguards

Posted by Alex Jackman FPB on 21/01/13

Increased availability of information on unincorporated businesses can be used to support lending to those businesses. By making the information accessible it will ensure that Credit Reference Agencies are able to better categorise the business and improve the opportunities for that business to receive funds, if desired.

The Forum of Private Business understands there is currently no central register of unincorporated businesses at present. The release of this data will require some safeguards to be in place and HMRC must be vigilant to ensure that only accurate information is released, albeit on a regular basis. HMRC should also ensure a monitoring relationship is in place to ensure the proper use of this information by Credit Reference Agencies, though this role could be given to the Trade Association.

Government must ensure opportunities for fraud are identified before data is released and prevented to the greatest ability. In addition, an opt-out option for businesses from this data release should be made clear.

Lending in the economy to small businesses remains slow. Opening up this information can help to tackle some of the credit shortage that exists, though it should be in parallel with continued efforts to improve bank lending together with incentivising alternative forms of finance.

The FPB supports the opening up of this data.

[Login to post comments](#)

VAT Register data disclosure is supported by Graydon UK

Posted by Chris Oatts on 17/01/13

Graydon UK wishes to voice its support for requests to release data from the VAT register.

Increased levels of verification granted from VAT data will allow UK Credit Reference Agencies to give even more timely and up to date information to their clients with an improved level of customer confidence. With no UK central register of unincorporated businesses, organisations can spend large amounts of time and money in verifying that a prospective client is, in fact, who they say they are. Access to identification data on the VAT register will add an official source of current data which will reduce this workload for decision makers. This will be true of products supporting unsecured lending (Trade Credit), Commercial insurers, Anti money laundering/Know-Your-Customer products and services and secured lenders such as banks.

In times of recession where the appetite for lending is at its lowest ebb, any activity to increase the confidence of lending and maintain low levels of risk is essential to recovery plans, and the additional confidence that government verification can bring could make all the difference to a start-up business or a struggling SME. The impact will be felt worldwide as our ability, as suppliers to the international markets, to champion UK SMEs better in their search for international partners will also benefit from the addition of this data.

Graydon UK, through our representation within BIPA, is pleased to be included in discussions in relation to the potential release of data from the VAT register. We believe that even the smallest steps toward a data sharing solution involving VAT data will benefit all lenders, SMEs immediately and, by extension, the whole economy.

[Login to post comments](#)

A view from Wonga

Posted by OpenWonga on 16/01/13

Data is a key ingredient to any algorithm-based approach to assessing risk. The more business data that is


available, the better the basis for business risk decision-making.

An Open VAT register, such as the one proposed by gov.uk, would significantly increase the volume AND the quality of data available to credit reference agencies and platforms like Wonga for Business. It would lead to an increase in positive lending decisions, allowing more businesses to grow. We support the Open Data User Group's call, and look forward to working with all parties who wish to see it implemented.

• [Login to post comments](#)

Experian Support the Release of the VAT Register by HMRC

Posted by Mark Hargreaves on 15/01/13



Experian supports and recommends the release of the VAT Register, as it will allow CRAs to be able to identify and verify SME businesses more easily, and will allow for improved accuracy of credit scores. In turn this will allow Banks/Lenders better visibility of the risk associated with these businesses. This will have a knock on effect to the economy and support the Government in its aims to stimulate economic growth, whilst aiding the banks to hit their lending targets. Also, with the improved accuracy of credit scoring, there will be an improvement in trade credit between businesses, as there would be an increase in credit offered to SME businesses that fall within the VAT Register, that without the release of this data, would otherwise have had a lower credit score and therefore may restrict the level of credit available to them. Also, through the improved access to verification and identification data for these businesses, it will help support AML/Anti-fraud activity, in being able to get clear transparency as to whether a business exists or not. Experian through our representation within BIPA is keen to work closely with HMRC along with ODUG, the Data Strategy Board and the wider lending community to help make the release of the VAT Register possible for these purposes, within the current legislative framework.

• [Login to post comments](#)