

Open Data User Group (ODUG)
Driver and Vehicle Licensing Agency (DVLA) Bulk Data
August 2013

1. Summary:

- The DVLA holds a highly valuable bulk data set covering all vehicles in the UK. It currently is sold at £96,000 plus VAT per annum. Licensing via the companies currently supplied with the data is limited and subject to strict terms & conditions.
- The data request is for an open release of all 47 fields under an Open Government Licence.
- **Economic benefits** include increased economic activity and innovation. There is currently no alternative to this data therefore making it open creates a gateway to new markets. For example, the potential to deliver competitive car check services which would generate estimated saving for consumers of up to £126 million per year.
- The **social benefits** are vast: the value of having safer roads, less pollution and fraud prevention is hard to quantify. Cost savings in petrol thefts are estimated at £2.3m p.a. Current estimates of 100,000 stolen vehicles in the UK every year imply annual costs to the UK economy of £several billion, so a reduction in the number of stolen vehicles would also generate significant savings.
- The risk of releasing DVLA bulk data as open data is **low** because the data does not contain personal information and is already available to paying customers.
- Our work has also highlighted the opportunity for DVLA to make substantial cost savings in its data infrastructure costs, which should be actively pursued to reduce the cost of releasing this dataset as open data.

ODUG recommends that:

- ***DVLA makes an open release of all 47 fields of the data under an Open Government Licence.***
- ***DVLA renegotiates its bulk data contracts (IBM and Fujitsu) where there is a great opportunity for cost and efficiency savings.***

2. Context

- The [DVLA](#) holds vehicle licensing data for all vehicles in the UK. The bulk dataset is a file of 47 fields of data such as registration plate number, date of registration and the VIN (Vehicle Identification Number).
- Data is currently available on a commercial re-sale **licence of £96,000** (plus VAT) p.a. and is only licensed by 8 organisations. The total revenue stream is therefore around £768k p.a.
- Anecdotal evidence suggests that the re-sale licence is not clear-cut, which leads to uncertainty and high administrative costs for users further down the supply chain and also burdens the DVLA with the cost of policing and litigation.
- Data use is currently controlled by strict compliance measures to prevent an individual being identified and the risk of fraud (including log book & car cloning).
- A request has been made for the release of the bulk data and further validated to include only the fields that would prevent the risks highlighted above. Previous consultation by DVLA has suggested that they are open to releasing more data.
- The DVLA are not authorised to make a profit from selling the data. We have an estimate of DVLA's cost (see below) and expect that a renegotiation of their contracts would result in substantially lower costs.

3. Benefits from the data and an open release

- There is currently **no alternative** to this data therefore making it open creates a gateway to new markets. In terms of some examples, access to this data would allow entry into the market place for supplying or reducing cost to the data for:
 - a) Insurance firms (who need the basic vehicle data to create vehicle insurance policies)
 - b) Car parking firms (who use the details to track the vehicles that are using their car parks)
 - c) Vehicle part replacement firms (who need the data to match parts to vehicle models/makes/derivative numbers)
 - d) Motor deals and car hire firms (who need the data for managing their business processes)
- Examples of open data show that easier and cheaper access can dramatically increase demand, which can result in a **positive spiral**, as more users accessing the data may result in better data quality, which then facilitates wider usage, which increases demand.
- The DVLA outlines several benefits of vehicle data to **society** on their [website](#). These include safer roads, e.g. by keeping illegal cars off the road;

less pollution, e.g. through improved traffic flow; and other benefits such as freeing up disabled parking spaces.

- Automobile Association estimates speak of 100,000 stolen vehicles in the UK every year at a **cost of several billion £ to the UK economy**. *Even a 0.1% reduction in vehicle theft would surpass the current DVLA revenue stream.*
- There is also a wide need for accurate information on a vehicle to **prevent fraud**. Case Study 3 shows how petrol stations can avoid the problem of ‘ bilking’ .
- An open release would stimulate **economic activity** and **innovation**. Case Study 2 on Total Car Check, a start-up which has developed a low-cost ‘ pay per click’ application, shows how consumers could benefit from **potential savings of up to £126 million per year**.

4. Effects of unclear re-use and licensing policies

The DVLA has a monopoly position in licensing this public data on to others. The licensing terms for the re-use of the data are unclear and it is difficult and expensive for third-parties (especially SMEs) to contest these terms, dissuading the widest possible use of this data for maximum economic benefit. This is exemplified by the recent events in Case Study 2, Total Car Check.

Total Car Check (TCC), a start-up, developed a low-cost pay-per-click application based on DVLA data, properly sourced from a third party. Following a successful launch of the service it was effectively closed down for two months because the DVLA challenged whether TCC was able to use DVLA data in their product as they are not a primary licensee. The argument employed by the DVLA followed other DVLA licensees with more expensive competitive services to TCC complaining to the DVLA. TCC has been forced to contest this with DVLA and, at the time of writing, the DVLA have accepted TCC’ s latest request to provide vehicle provenance checking services based on their data. These concessions do not allow TCC to use the data with the same freedom as the primary license holders, significantly affecting TCC’ s business opportunities.

This example emphasises the current barrier and administrative overhead the DVLA licensing regime presents for businesses wishing to use and re-using DVLA data, also the difficulty DVLA has in trying to oversee and police all existing and future services based on its data. Disputes of this nature over the re-use of public data serve only to benefit the lawyers involved in the debate, inflicting substantial, avoidable costs on both sides.

5. Cost structure

The DVLA sells its bulk data on a not for profit basis. We have been unable to obtain cost estimates from the DVLA as precise cost estimates are not available. DVLA’ s data sharing manager states "we do not hold the specific information because as this is a long standing product, it is not information held“ . However DVLA provided some **unofficial calculations about their annual running costs**:

104 weekly data scans at £3,000 = £312,000

12 monthly data scans at £8,000 = £96,000

2 full bulk data extracts at £50,000 = £100,000

Team support costs = £75,000 (very rough estimate)

The estimated total annual cost is therefore **£583,000**, which, at the current license fee, requires DVLA to have around six bulk licensees to break even. The DVLA currently have eight licensees meaning that *revenue exceeds estimated costs by approximately £185,000* at the current levels.

The cost breakdown indicates that *the majority of DVLA's costs are to fund data requests from 'their own' IT services*, run via a tripartite contract between themselves, Fujitsu and IBM.

Why the costs for the data scans are so high is unclear. Despite missing official data, *we have strong reasons to believe that the data scans can be provided more efficiently.*

6. Concerns and risks around an open release

- **Loss in revenue**

The open release would void the £96k * 8 companies = £768,000 p.a. in expected future revenue to the DVLA. This is not the only revenue stream for the DVLA and the figure dwindles in the context of the available economic, social and environmental benefits.

- **Increase in cost**

There is no evidence that more users will increase the cost of releasing the data because the data is already provided for current customers. In fact, re-negotiations of existing IT contracts should result in lower costs.

- **Disclosure of personal information**

The DVLA bulk data set does *not* contain any personal information. All the fields relate to characteristics and history of the car. Name, address and other information of the registered owner are not present in the dataset.

- **Privacy**

The DVLA data only affects privacy if the car is matched to the individual. This may happen via third parties, who collect various amounts of data as part of their customer relationship management. Ultimately, this practice is covered under general privacy and marketing laws and is remote from the open release of the DVLA bulk data.

- **'Cloning' of cars**

The most valuable information for cloning a car from the DVLA bulk data is the Vehicle Identification Number (VIN) field. Our recommendation is that the full VIN should be released because the VIN is visible from outside the vehicle on almost all modern vehicles manufactured since 2003, generally located just below the traditional tax disc position on the passenger side of the dashboard. However, an alternative solution might be to censor this field and only releasing the last two

digits.

- **Residual risk**

The residual risk is low because the data does not contain personal information and is already available for paying customers.

Annex A

A.1 Current licensees

There are currently eight companies which are supplied with the DVLA data (six of which are listed below, two chose to remain anonymous). CarweB are now fully owned by HPI and could merge licences reducing the number of customers to seven.

- CAP Automotive Ltd, Phone: 0113 222 2000
- CarweB Ltd, Phone: 01732 373002
- CDL Vehicle Information Services Ltd, Phone: 08444 821965
- Experian Ltd, Phone: 0870 0131 696
- HPI Ltd, Phone: 01722 422 422
- Vehicle Data Services Ltd, Phone: 01661 854388

The main benefits of opening up the data for the companies above are financial savings associated with licensing the DVLA data (£96,000 p.a.).

Ideally, these cost efficiencies have positive external effects:

1. They are **passed on to the end user**, so that consumers can directly experience lower rates for products based on the DVLA data.
2. This would encourage their role as **intermediaries**, i.e. they could release data or services on a 'per check' basis (as it is current practice) at reduced fees and hence stimulate business models based on the DVLA data.
3. Companies avoid the decision to justify the licence fee for upcoming years. Thus, they may be more open to use the data for a broader variety of applications and build services, which allow a **more widespread use of the DVLA bulk data set**.

A.2 Case Study 2 – [Total Car Check \(TCC\)](#)

The Total Car Check case study shows that there is **strong demand** for a more accessible release of the DVLA data. They access the data via one of the 6 direct customers. However, the business model is stifled by the DVLA licencing cost and terms & conditions. This is exemplified by this case study:

TCC “ helps people check the history of a vehicle” . They provide a low-cost ‘ pay £1.99 per click’ access to the DVLA data. Competitors like the market leader HPI charge £19.99 for each check. Individuals and companies have a strong incentive to access this information to prevent fraud if it is available at a low cost.

There are around 7 million used vehicles sold in the UK every year. Webuyanycar, for example, advertise the fact that they have bought 425,000 vehicles which all require a vehicle history check before purchase.

*This means an open release would, at current prices, have a **potential up to £126m of savings** for consumers.*

A.3 Case Study 3 – Petrol Stations

Statistics from the Automobile Association speak of 100,000 vehicles stolen per year of which only around half are recovered. Stolen vehicles pose a risk to society and businesses because their drivers cannot be easily prosecuted.

Petrol stations, in particular, are affected by the practice of ‘ bilking’ (the act of filling a vehicle with fuel and leaving without making payment). The British Oil Security Syndicate estimates the **cost of bilking at £23.4 million** for 2012.

An open release of the DVLA data would allow petrol stations to mitigate the problem of bilking. Some conglomerates already employ customer data identifying vehicles with number plate recognition systems. Individual petrol stations could install their own systems, link it to the data set and prevent bilking and related crimes. This would also have a positive effect on businesses, which sell cameras, develop pattern recognition software, and may spread the use of the DVLA data to other industries.

*A 10% reduction in the cost of bilking alone would amount to roughly **£2.3 million savings per year**. The benefits for society and further innovation are not included.*

Key Benefits

Data release goals rated in the scope of ODUG key benefit areas.

Key Benefits

	1	2	3	4	5
Efficiency		X			
Environmental			X		
Growth			X		
Social		X			
Transparency			X		

Data Theme Fit

Data Request rated in the scope of ODUG key data themes.

Data Theme fit

	1	2	3	4	5	6	7	8	9	10
Land & Property		X								
Environment					X					
Social						X				
Orgs. & Companies				X						
Education	X									
Transport									X	
Financial							X			
Health	X									

Not at all So me s Applie s Fully applies Single goal