



# Smart Meter Energy Data: Public Interest Advisory Group

A policy dialogue and work programme  
led by  
Centre for Sustainable Energy & Sustainability First

Stimulus paper 7

## Possible routes to the data for a public interest purpose

Author : Maxine Frerk. Sustainability First

with support from Judith Ward. Sustainability First  
& Simon Roberts and Nicky Hodges. CSE

### Status of this Document

This paper was prepared as an input to the work programme of the  
Public Interest Advisory Group on access to smart meter energy data.

An earlier version was discussed with PIAG members at their workshop on 20 November 2018.

Last Revise – 23 January 2019

## PIAG Stimulus Paper 7

### Possible Routes to the Data for a Public Interest Purpose

#### Introduction

The aim of this paper is to explore possible eventual routes to smart-meter data for a ‘public interest’ purpose.

This discussion is not taking place in isolation. There are other debates about access to customer smart meter data in relation to the BEIS Data Access and Privacy Framework, settlement reform, the future of supply (Energy Midata), the data privacy plans of the electricity networks and UCL’s smart meter research portal.

In [Stimulus Paper 5](#) we examined key ‘public interest’ use-cases which stood to benefit from access to more granular smart-meter consumption data. These include: improved national energy statistics; better data to support regional and local-level planning; better data to allow policy-makers to model the impacts of new policies; and, potentially, access to better sample data to enable better targeting and development of new services. Our conclusion was that significant benefits could be delivered from the use of smart-meter consumption-data for such public interest purposes, provided customer privacy concerns can be satisfactorily addressed.

The focus in all of the PIAG use-cases was on energy consumption data although for certain uses other energy data, such as export, would be relevant. As noted previously, smart meters provide a wide range of data which could be of public interest. However our focus at this point has been on energy consumption data which is also where the strictest privacy rules apply.

Stimulus Paper 5 also made a helpful distinction between (1) the **input data** required for each public interest use-case – of necessity at the individual customer-level and (2) the resulting **output datasets** which would only become available for use by policy-makers and others once suitably aggregated or anonymised. **As such we do not foresee fundamental privacy questions arising on the use of anonymised output data for a policy-making purpose.** But for **input data**, some real practical challenges remain within the current privacy frameworks and statute as to how, potentially, this could be obtained and then aggregated / anonymised.

Whatever route is chosen, decision-makers and consumer bodies need to feel confident that customer sentiment is largely amenable to smart meter data being used in this way for a public interest purpose and that overall an appropriate balance has been struck between respecting individual privacy and the wider public interest. These issues are considered at length in PIAG [Stimulus Paper 6](#) and the Ipsos MORI paper (January 2019), which we will return to at the workshop on 4 April 2019.

In practice, a move to make better use of smart meter data aligns well with the current direction of travel across all of government to make better use of administrative data for public interest purposes (as reflected in the Digital Economy Act 2017). Annex 1 provides more context.

This paper therefore puts forward a **suggested framework** to explore options for accessing customer smart meter data for a public interest purpose. In particular the paper outlines:

- A **strawman process to access smart meter data for a public interest purpose** - in terms of data-inputs, data-outputs, the role of a trusted processor and potential users of the data;
- A summary of **potential sources for the input data**. This is where the UK faces a particular challenge as there is no single data-base of smart meter data that could be used as a source (unlike in the US);
- A **set of proposed criteria** by which to judge between alternative potential routes to obtain customer smart meter input data;
- A **basic and initial strawman assessment** of different options against these criteria;
- Reflections on **other elements of the process**;
- **Early concluding thoughts – for discussion**.

**The purpose of this note was to frame the discussion at the PIAG workshop on 20 November 2018 and has been updated in the light of comments made.** A number of industry actors were invited to set out how they potentially foresee their own roles evolving to support such a process.

## **Overview – a strawman process to access smart meter data for a public interest purpose**

As noted, any public interest use-case would need granular customer-level input data.<sup>1</sup> Given that there are no plans for a comprehensive GB database of the consumption data from smart meters we have considered the options that exist in terms of how this **input data** could be obtained and then converted, by what we are calling a **‘trusted processor’** into the **output data** that is needed to support each use-case.

Our first step was to review what meter data is currently held – or is planned to be held – by different industry players. We then explored how these data-sources might (in principle at least) provide a route to input-data for a trusted processor. We also considered possible options for a range of future routes to input-data longer term.

In every case, and in line with design of the GB smart meter system, the consumption data sits on the individual smart meter (for a maximum of 13 months). For SMETS2 meters (and in due course, SMETS1 meters), the data is retrieved from the meter by the DCC on instruction from organisations such as energy suppliers in line with the rights and consents those organisations have to access the data as

---

<sup>1</sup> Reasons for this are explained in our use-case discussion in [Stimulus Paper 5](#)

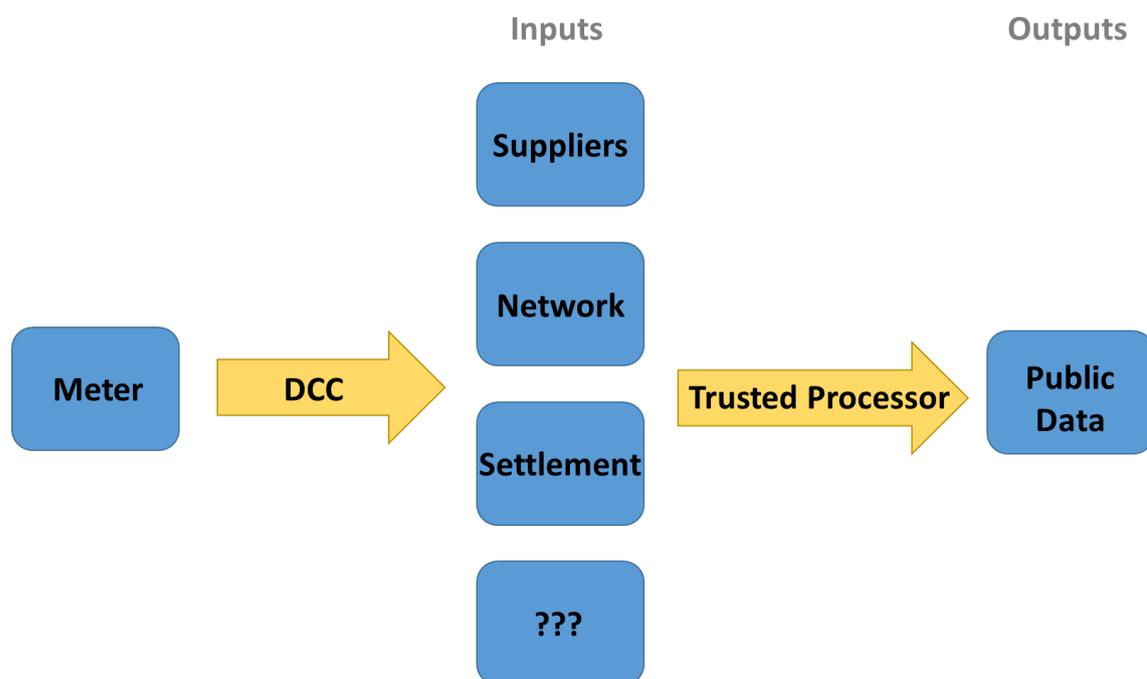
registered DCC-users. In accordance with its licence, the DCC does not itself hold the data; rather it routes the encrypted data to the relevant party without access to it.

In terms of potential sources for input data, discussed at more length in the next section, we have identified:

- **suppliers** who collect and store smart meter (and ‘dumb’ meter) data at individual customer-record level for both gas and electricity. For smart meters, this will be at least monthly consumption data (which they have rights to access). The Data Access and Privacy Framework requires that suppliers obtain customer consent to access data at finer granularity (such as half-hourly).
- **distribution networks** who have access to half-hourly data for regulated purposes without customer consent subject to approval of a privacy plan which shows how the data will be aggregated and anonymised. To date only Western Power Distribution (WPD) have a plan approved.
- in future, **the settlement system** might arguably become an alternative source of granular input-data (assuming the settlement process eventually has comprehensive access to such data).
- and ultimately – if the case were considered strong enough - provision could be made for a **new party** to collect smart meter data specifically for a ‘public interest’ purpose.

A **trusted processor** role is essentially about obtaining MPAN/MPRN-level input data (ie consumption data), processing it securely and providing the outputs needed. To do this may require the input data to be held securely for a certain period. For some public-interest use-cases, processing would essentially then be a relatively simple task of aggregation to a suitable level to ensure individual consumer data cannot be identified. Other use-cases, such as the provision of input-data for modelling, would involve linking the smart meter data with other data at an individual record level (which means that the trusted processor would also have to have access to this other household level data) and then making that de-identified linked-data available on a suitably secure basis. (See [Stimulus Paper 5](#) for more detail on the relationship between Input and Output data for different use-cases and the role of a trusted processor).

This strawman process is illustrated in Figure 1 below.



**Figure 1 – A strawman process to access smart-meter data for a public-interest purpose**

**Source : CSE & Sustainability First**

Two established entities already play the role of a trusted processor within the UK’s legal and institutional arrangements:

- The UK Data Archive (funded by ESRC and primarily used by academics)
- The ONS (for government statistics).

Practical considerations suggest, for the short-term at least, that rather than establish a new trusted processor it would make sense to rely on existing arrangements as provided for in the Digital Economy Act - where ONS would appear to be more relevant for wider public interest issues (but noting that UCL are expecting to use the UK Data Archive for SMRP given their focus is on the academic community).

Under the Digital Economy Act the ONS also already has the powers to request data from private sector companies for public interest purposes. Use of these powers would avoid the need for an explicit individual consent but obviously – and crucially in this context - can only be used where the data already exists. Further detail is provided in Annex 2.

In line with the provisions of the Digital Economy Act, and should it take on such a role, the ONS (or any other trusted processor) would need to consider carefully how it made any outputs available to ensure privacy was protected. If data was suitably aggregated so there was no risk of re-identification (eg for national or local area statistics) then the data outputs could be published. In other cases where

smaller datasets are required or pseudonymisation is used to give access to individual records linked to other socio-demographic data for research purposes then the ONS would be required to follow its own published guidance which is based around the '5 safes' (described more fully in the [Data Ethics PIAG Stimulus Paper and summarised in Annex 2](#)). In particular, this includes arrangements for accrediting researchers and projects to access the linked data.

Longer term the Digital Economy Act does provide for other bodies to be accredited by the National Statistics Authority as data processors, responsible for linking, de-identifying, storing and making data securely available. However they would not have the same data gathering powers as the ONS. Details of the accreditation process have not yet been published but are expected shortly.

There could also be an option, if a single regulated body already held all the smart meter input data – be it for a particular geographic area or nationally - that they could be required through their own licence to publish aggregated data for public interest purposes. This could allow certain use cases to be met without a third party (such as the ONS) acting as a trusted processor.

PIAG members were asked to consider whether this strawman process for taking input data and providing the outputs needed under the various use-cases made sense or, whether there were other models that should be considered. No alternative models were identified.

## Potential sources for the input data

### *Current sources of input data*

As noted above, a number of parties already have access to certain smart meter data (via DCC as registered DCC-users), or will do so in future. These parties are listed below with a summary of what data they can access. In the next section we explore the strengths and weaknesses of each as a route by which a trusted processor could potentially access smart-meter input-data.

**Energy Suppliers** – currently have access to: monthly gas and electricity data without customer consent (for billing and for fulfilling statutory and licence requirements); daily data on a customer 'opt-out' basis (for any purpose except marketing<sup>2</sup>); half-hourly data on a customer 'opt-in' basis. In its consultation on access to data for half hourly settlement<sup>3</sup> Ofgem noted that not all suppliers currently seek to collect half-hourly data and opt-in rates vary widely between suppliers reflecting the different approaches taken. Consumer attitudes can also be expected to change over time. It is understood that where suppliers seek consent to collect half-hourly data they would typically do so for both electricity and gas.

---

<sup>2</sup> Customers must explicitly opt-in to the use of their data for marketing

<sup>3</sup> [https://www.ofgem.gov.uk/system/files/docs/2018/07/access\\_to\\_data\\_for\\_settlement\\_consultation\\_5.pdf](https://www.ofgem.gov.uk/system/files/docs/2018/07/access_to_data_for_settlement_consultation_5.pdf)

**Electricity Distribution Networks** – are entitled to collect half-hourly *consumption* data from every customer subject to having their privacy plan approved by Ofgem setting out how they will aggregate or anonymise the data. To date, WPD is the only network to have a privacy plan approved. Other DNOs are working on their plans and are expected to submit them in the first half of next year. Assuming the broad approach agreed by Ofgem for WPD is followed for other energy networks, distribution networks will have access (via the DCC) to smart meter data at individual record level for system planning and other agreed regulated purposes. WPD's approved privacy plan only allows for data aggregated to feeder-level to be stored (which may include some individual households). Our understanding is that at least one of the other DNO plans to seek approval in their privacy plan submission to store all data at individual meter level, suitably protected.). One challenge is that it would seem all DNOs are approaching this in different ways and there are several different underlying systems for collecting and aggregating the data.

As well as the consumption data authorised by their privacy plans, networks can also be expected in the future to access other meter data intended to help them manage the network such as voltage alerts, maximum demand, reactive power, export etc. Currently networks can access data only from SMETS2 meters but should be able to access data from SMETS1 meters when these are adopted by DCC in due course.

**Gas Distribution Networks** – are entitled to collect data on the same basis as DNOs but so far are not doing so and, to our knowledge, are not actively preparing proposals to Ofgem to do so. Indeed no GDN is yet registered as a DCC user.

#### **Possible future sources of input data**

Other bodies also have access to certain customer smart-meter data – for I&C customers as well as for small customers and households. This includes the bodies responsible for system settlement (Elexon, Xoserve) plus some industry delivery bodies (e.g. DCC, ElectraLink). Looking to the future, and subject to any necessary rule-changes, these or other bodies may also offer potential routes to consumption data and other input-data for a public-interest use.

**Electricity settlement (Elexon)** - current reform of electricity settlement will lead to new data-flows and processes to feed into future settlement arrangements. These are currently out for consultation. The settlement system is likely to receive some customer half-hourly data, but not from every customer as a matter of course<sup>4</sup>. Ofgem is also looking at how data submitted to settlement will be collected and thereafter aggregated. Of particular relevance to the PIAG discussion is the question - still out for Ofgem consultation - on whether suppliers (or their agents as data-collectors) should each separately aggregate customer meter-data prior to submission into settlement. Or, whether individual customer- data should be submitted direct by suppliers (or their agents) to the settlement body without prior aggregation. The latter arrangement would mean that in the future the settlement body would hold a large set of half-hourly electricity meter-level data for the purpose of settlement. While

---

4 A current Ofgem consultation proposes that customer half-hourly data should be submitted into settlement on an 'opt-out' basis (or opt-in where a meter has already been installed). By the time Ofgem's decision takes effect, the majority of meters could be 'opt-in'.

no final decision has yet been taken on this, Ofgem has published an open letter stating that the working assumption for system design should be that individual records will be submitted into settlement.<sup>5</sup>

**Gas settlement** (Xoserve) - is done on a daily basis but domestic customers are typically settled based on an annual profile. Project Nexus, implemented recently, allows for daily data (or other longer periods) to be used for settling domestic customers. Changes to arrangements and supplier practices could lead to this becoming an additional source of smart meter data on gas – but at most at a daily level.

**Smart DCC** - DCC securely transmits consumption and other data in line with its licence. However, DCC is currently not able to access that data because it is encrypted. For DCC to access that data would require legislative or regulatory change and a fundamental change to DCC's systems. DCC does hold some other smart-meter data such as records of service requests and alerts and has a role also in procuring the central switching service as part of Ofgem's faster and more reliable switching project.

**Electralink** – Electralink provides data communication services for the electricity sector to support switching and settlement. They have been exploring the opportunity to make wider use of this data under the “legitimate purpose” gateway in GDPR and in particular are now providing I&C half-hourly data to National Grid.

**ONS direct as a DCC user** – theoretically and subject to appropriate statutory arrangements, the ONS (or another trusted processor) could become a registered DCC-user and so access individual-level meter-data records for each MPAN and MPRN. However to access consumption data in this way would require legislative change to provide a basis for the ONS obtaining that data without customer consent.

**Other entity** - more generally, if, and counter to current policy, a decision was made for public policy reasons to have a single comprehensive data-base of granular electricity and gas consumption data and/or other meter data - for regulated use within the energy system - then responsibility for collecting that input-data could be given to a single body (for example, Smart DCC, the settlement system, the system operator etc). Input data for any such data-base would still be accessed from the smart-meter via the DCC and specific legal provisions would be needed to allow this. A trusted processor would then have access to the data-base as an input to its work to produce outputs which meet various public-interest use-cases. However, this would be a major departure from current arrangements which were originally designed to avoid having a central database.

PIAG members were asked to comment on the potential sources of smart-meter input-data listed here – including on possible over-sights or omissions. No other potential sources were identified.

---

<sup>5</sup> [https://www.ofgem.gov.uk/system/files/docs/2018/11/dwg13\\_ofgem\\_policy\\_steer\\_v1.0.pdf](https://www.ofgem.gov.uk/system/files/docs/2018/11/dwg13_ofgem_policy_steer_v1.0.pdf)

## **Proposed criteria by which to assess routes to smart-meter input-data for a public-interest purpose**

In exploring possible alternative routes for accessing smart meter data for a public interest purpose we have assessed these possible options against a number of basic high-level criteria:

**Complexity** – number of parties (and / or different systems) involved.

**Current data availability** - how much smart meter data is accessible through that route currently

**Comprehensive coverage longer-term** – how much smart meter data would be accessible through that route longer term (eg in 5 years plus)

**Cost** – eg if new arrangements would need to be put in place to collect or access the data (rather than relying on data already collected for another purpose). This aligns with a principle of ‘collect once, use many times’.

**Legal basis** – ie is there a current legal basis for that body to collect the input data (and what prospects are there for securing changes to legislation, taking account of government commitment to certain aspects of the overall design).

**Capability** – ie does the body have the required skills around secure data handling and access to supplementary datasets.

**Consumer confidence** – ie is the body likely to be trusted reputationally by government, regulators and customers . Ideally the body would be a respected neutral party without a commercial interest in the data.

PIAG members were asked to consider whether these were the ‘right’ criteria by which to assess possible routes to accessing smart-meter input-data. Some of the comments have been incorporated into the explanatory wording of the criteria above. Additional criteria suggested by PIAG members that are worth bearing in mind are:

**Coherence / interoperability** – there are benefits in a uniform approach that would facilitate a national, whole system view, providing a solution that also serves other system data needs and / or can borrow from data capture being done by other system processes;

**Future proofing / future enabling** – for example noting the potentially radical changes being contemplated around the role of suppliers and programmes such as faster switching.

## Strawman assessment: possible sources of smart-meter input-data against key criteria

In Figure 2 we have summarised our assessment of the different routes to input data identified so far against our suggested criteria. The assessment will vary for the short and long term as shown in the table and discussed further below.

Route to access input data	Energy Supplier	Network (DNO)	Network (GDN)	Settlement (Elexon / Xoserve)	Smart DCC	ONS - as a DCC user
<b>Complexity</b>	60+ energy suppliers	6 DNOs – different approaches	4 GDNs	Single body	Single body	
<b>Current data availability</b>	Gas+ Elec SMETS1+2 Varying granularity (monthly / daily / half-hourly)	Elec SMETS2 Half-hourly Aggregated? Subject to privacy plan (WPD only)	Gas SMETS2 Not collected	None	No right to access to encrypted data (currently SMETS2 only)	No basis under DAPF for access without consent
<b>Long term Comprehensive coverage</b>	Gas+ Elec SMETS1+2 Only monthly for some	Elec SMETS1+2 Half-hourly Aggregated (or pre destruction?)	Gas SMETS1+2 May be collected	Elec - Subject to Ofgem decision on HHS Gas - Daily or monthly data	Gas+Elec All - subject to legislation	Gas + Elec All - subject to legislation
<b>Additional cost</b> (i.e. of being 'route to data' for trusted processor - on top of core purpose)	Depends what each supplier already collects & processes for own purpose	Depends on what each network plans to collect & process for own purpose			Would involve capturing and unencrypting data - so new systems needed.	New system and DCC costs per meter read
<b>Legal Basis</b>	Y	Y	Y	Being considered for elec	Not at present	None at present
<b>Capability</b>	Reliability of all suppliers as providers?				Highly secure approved systems.	Handle other sensitive data
<b>Consumer trust</b>	Mixed	Low awareness	Low awareness	No awareness	No awareness	Trusted
<b>Overall – short term</b>	Best option?					
<b>Overall – long term</b>	Limited granularity (plus question re reliability of collection)	Need gas too	Unclear if will collect	Subject to HHS on elec. Level of gas data unclear.	Fundamental change to current design of smart meter systems, incl on DCC role.	Would need a legislative change. Costs of DCC use.

Figure 2 - Strawman assessment : possible sources of smart-meter input-data against key criteria  
CSE & Sustainability First

**Short-term options for access to input-data**

In practice, short-term options for access to input-data are determined by what smart-meter data is available now (or in say the next 12-18 months). On the basis of our strawman assessment, the most expedient and best short-term route to access input-data would be via suppliers. This would cover both gas and electricity and includes SMETS1 meters, albeit for many customers only monthly data would be available. Even this level of granularity would be a significant improvement on energy consumption data submitted into government statistics today.

**Longer-term options for access to input-data**

The strawman assessment suggests a number of options could perhaps be on the table longer-term as routes to access smart meter input data. For example :

**Suppliers** - remain a potential long-term route to input-data. Relative to other options in Figure 2 the value of this route will depend on what proportion of customers have consented to access and storage of their more granular data. Plus, to what extent every supplier could be relied upon routinely to provide input-data (due to the number and diversity of players and the implications of faster switching).

**Energy networks - DNOs** – In due course once SMETS1 meters have been adopted by DCC - and once DNO privacy plans are approved - DNOs could potentially provide a better source of electricity half-hourly consumption data than energy suppliers (i.e. if the raw meter-level data could be passed-on by DNOs to a trusted processor rather than deleted immediately following aggregation). However complexity will remain if each DNO takes a different approach.

**GDNs** – by contrast, unless and until GDNs change their stance, there is no clear equivalent route to obtaining more granular gas input-data (e.g. daily) from a GDN. More granular gas consumption-data would have particular value in supporting future policy development around heat-decarbonisation, and more thought is therefore needed around responsibility for collection of more granular gas data.

**Settlement** - reformed electricity settlement arrangements could potentially provide an alternative long-term route to more comprehensive half-hourly electricity consumption and other input-data, subject to Ofgem's decision. The gas settlement system may also be an eventual source of at least some daily gas input-data but at present there are no plans for mandating this. More granular data is arguably less important on gas but could still be of value for some of the use cases.

**Smart DCC** – Hypothetically, DCC could collect a 'copy' of the gas and electricity meter-data it transmits to suppliers and network companies. This might have the potential to be a comprehensive and cost-effective route to granular input-data for the future. However, this would entail a significant change in DCC's current role, requiring legislative and licence changes.

**ONS** – the ONS, if it became a registered DCC-user in its own right, could also potentially access comprehensive gas and electricity input-data in the long-term. However, legislation would be needed for ONS to request smart meter-data from the DCC without customer consent. The DEA does not give the ONS the right to that data because, under the DEA, administrative data can only be supplied to ONS by a party who *already* anyway collects it. Moreover, the cost for ONS of accessing this data via DCC could be material (and would duplicate others already accessing the same data).

## Other elements of the strawman process

As discussed, the role of trusted processor in handling customer-level input data is central to the thinking around our strawman process. One party to play the role could be the ONS given their powers under the DEA and their access to wider sources of data. They already have the necessary governance and technical skills to deal with publication of data and for making sensitive data available for research purposes; and consumer research suggests they would be viewed as a trusted body. However clearly there will be a question for ONS about their own priorities and resources.

In terms of the **outputs**, as noted above, there is a distinction to be drawn between:

- Outputs where the risk of re-identification has been removed (eg through high levels of aggregation) and which can simply be published, as statistical data;
- Outputs where the risk of re-identification has been reduced but some risk might remain and where the governance arrangements need to reinforce the level of protection (by limiting who can access the data for what purposes, under what conditions etc).

Over-arching governance arrangements already exist for access either to ONS or to UK Data Archive data (some of which were described in PIAG Stimulus Paper 3 on Data Ethics). However, there may also be specific questions around who might be considered public interest actors in relation to smart meter data and levels of aggregation likely to provide suitable anonymity (where international experience as well as thinking on DNO privacy plans may be relevant).

## Summary of early thoughts on possible routes to smart meter data for a ‘public interest’ purpose

Many queries and unknowns remain, particularly around the longer-term options. In practice taking forward any of the options to access smart meter data as we describe in this note would require detailed work, including on licence or other implementation approaches. That said, our very initial thinking points towards the following.

**For the short term:** energy suppliers seem to offer the most straightforward initial route to more granular smart meter data for a ‘public interest’ purpose. This would be of value even though suppliers may only hold monthly data for some customers. For the purpose of creating more detailed and accurate national and local energy statistics this would still be a marked improvement on current

practice – both in terms of greater granularity and also speed. It would also build on existing arrangements whereby individual-level annual energy consumption data is obtained by BEIS from suppliers (under the Statistics of Trade Act 1947<sup>6</sup>) and hence should not raise any material or new issues. One would expect that suppliers' privacy plans provide for them to pass on data where they are legally obliged to do so<sup>7</sup>. In terms of improving both national and subnational energy consumption statistics, this could prove an early 'win'.

From a BEIS standpoint some additional effort and resource may be needed to manage monthly data inputs. Significant effort could well be needed to develop new systems to handle half-hourly input-data where available. Arguably, one potential option might be for ONS as an established data processor and using its powers under the Digital Economy Act 2017- to take-on this role of data collection for statistical purposes from the energy suppliers on behalf of BEIS. This would allow ONS to start from scratch in collecting the data at whatever granularity suppliers had available. In so doing, it could also allow ONS to create the kind of data capture and storage capability needed longer-term to hold the 'input data' necessary to serve the full range of public interest use-cases.

Similarly for suppliers there would be cost and effort in developing their systems to pass on more granular data securely. There would be little point doing this in advance of SMETS1 meters being integrated in the same systems. To provide a comprehensive picture the requirement to provide data would need to apply to all suppliers.

***In the slightly longer term:*** energy suppliers would still provide a straightforward and low-cost route to reasonably comprehensive gas and electricity data. If half-hourly electricity data is routinely collected either by **DNOs** or the **electricity settlement body** then either might eventually provide a better data-set for the ONS or other trusted processor to call on. But, this would also still have potential limitations; and in particular it would not provide any route for obtaining gas consumption data. XOServe could potentially be a route for accessing gas data but is dependent on the level of daily data that suppliers collect and provide into the system.

***For the long-term:*** delivery of a more comprehensive solution to obtaining high quality, fine grain customer smart-meter input-data for a public-interest purpose – and / or to deliver modelling-data based on individual linked records – is likely to require **a new bespoke route** (or routes) to collect energy meter data. The optimal route for doing this is not yet clear, and is dependent on wider market developments, but there are a number of potential options. For example to :

- Place a new obligation on DNOs and GDNs to provide smart-meter consumption and other data in standard formats to the ONS (in effect giving the networks a new regulated duty). This might well not require them to collect additional information, but instead to change the way they currently process it i.e. to pass the raw customer data to the ONS or other trusted processor instead of deleting it as now (and as DNO privacy plans require).

---

<sup>6</sup> See Annex 3 for further detail

<sup>7</sup> For example, this would also include providing data to enforcement agencies. Interestingly this is the angle that Privacy International flagged as having raised most concerns in the US context.

- Or, put new obligations on a body such as the settlement system, the system operator, the DCC or the ONS to collect the fine grain input-data specifically for these purposes. However this would be a significant departure from current smart meter data privacy arrangements, albeit potentially in line with government's wider data-strategy to improve and make better use of data<sup>8</sup>.

There would be cost implications if such a body was collecting the data purely for these purposes. There would also need to be some legislative or regulatory changes to establish the rights of such a body to access half-hourly individual consumption data via the DCC without a customer consent.

In terms of the trusted processor role, in the longer-term an obvious party remains the ONS. However, other options should also be considered, taking into account how the roles of other parties may well evolve. In particular in thinking about different use-cases it might be practical for other parties to publish aggregated data (and indeed this is arguably a part of the future DSO role to support local planning). For modelling - where it is necessary to link data with other socio-demographic data - then responsibility would need to sit with a body that also has access to that other data such as the ONS, subject to how the wider data landscape develops.

In considering these options thought should also be given, as an alternative, to how the SMRP data could be made available to a wider set of researchers to support the modelling use case.

Clearly any change to the arrangements for collecting and sharing data would need to be considered in the context of consumers concerns about privacy and the implications this could have for the rollout itself.

**We will return to this central question of how to achieve a balance between safeguards for consumer privacy and the benefits of access to data for a wider societal purpose in our final PIAG paper and workshop on 4 April 2019.**

**It is also clear that gas data presents some particular challenges and we will look to explore further with XOServe the level of data to which they expect to have access. We will also explore with BEIS and Ofgem their ambitions for GDNs to make use of smart meter data as part of delivering full benefit from the rollout of gas meters.**

## **Sustainability First & CSE**

**23 January 2019**

---

<sup>8</sup> Including by reducing the need for / expense of less than satisfactory alternatives – such as collecting energy consumption data via surveys or via sales data submitted by retailers of oil, gas and power.

## Annex 1: Wider Government Ambition

The Statistics Authority recently produced a report<sup>9</sup> on “Joining up data for better statistics” which signalled their ambition to see more done to build on the DEA powers and link different administrative data sources in the expectation that this would yield new insights and opportunities.

The term ‘**administrative data**’ refers to information obtained by a public or private sector organisation in the course of undertaking its normal operations, rather than with a view to its use for statistical purposes.

While this more ambitious use case has not been considered to date in the PIAG work it is worth noting this broader direction of travel across government.

The Bean Review<sup>10</sup> in 2016 highlighted the scope for more use to be made of administrative “microdata” and highlighted some of the barriers to doing so which included legislative barriers (as Information Sharing Orders or specific provision in legislation was required to give permission for the ONS to access each dataset and the specific uses had to be set out in the order which precluded exploratory uses of the data). The Bean Review also highlighted cultural barriers (risk averseness and concerns about the possibility of accidental data release) and a lack of ambition.

The Digital Economy Act (see Annex 2) was a response to these legal challenges and there is now no need for secondary legislation for ONS to access data.

However the Statistics Authority report mentioned above does highlight that there is still a reluctance to share data between departments and a lack of ambition. It notes the benefits that can be derived from linking data but also that this takes time and effort.

---

<sup>9</sup> <https://www.statisticsauthority.gov.uk/wp-content/uploads/2018/09/Data-Linkage-Joining-Up-Data.pdf>

<sup>10</sup>

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/507081/2904936\\_Bean\\_Review\\_Web\\_Accessible.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/507081/2904936_Bean_Review_Web_Accessible.pdf)

## **Annex 2: Digital Economy Act 2017**

### ***ONS Access to Information from Public and Private Sectors***

The Digital Economy Act 2017 amends the Statistics and Registration Service Act 2007 to provide the UK Statistics Authority and its executive office, the Office for National Statistics (ONS), with greater and easier access to a range of data sources held within the public and private sectors.

The ONS can only seek access to data for the purposes of fulfilling one or more of its statutory functions, including to produce official statistics and undertake statistical research that meets identifiable user needs for the public good.

The Act requires the ONS to set out principles that it will adhere to in exercising its new powers. It has done this and the principles cover confidentiality of data, transparency, ethics and the law, public interest, proportionality and collaboration.

### ***Disclosure of information for research purposes***

The Digital Economy Act 2017 also facilitates the linking and sharing of datasets held by public authorities for research purposes. To ensure data are processed and made available in a safe and secure way the legislation sets out six conditions under which this can take place:

- Data should be de-identified before disclosure;
- The risk of accidental disclosure should be minimised,
- Disclosure should be made by the relevant public body (or someone appointed by it)
- The research must be accredited
- The individuals involved (in disclosing or using the data) must be accredited
- The individuals must have regard to the statutory Code of Practice...

The DEA thus creates a gateway to enable public authorities to make data available to researchers for research that is in the public interest using a trusted third-party model. Under this model, a data-holding public authority discloses identifiable data to an accredited third party processor (or the public authority itself acting in this capacity), who is then responsible for processing the data (that is, linking, de-identifying, storing and making data securely available) before the de-identified data are made available to an accredited researcher.

Researchers must be accredited (but do not have to come from academic institutions) and projects must be accredited as being in the public interest.

The ONS is required to produce a Code of Practice which it has done and which includes examples of research in the public interest as being to:

- provide or improve evidence bases that support the formulation, development or evaluation of public policy or public service delivery;
- guide critical decision-making with anticipated impacts on the UK economy, society or quality of life of people in the UK;
- significantly extend existing understandings or social or economic trends or events, either by improving knowledge or challenging accepted analyses; or
- replicate, validate or critically analyse existing research (including official statistics) in a way that leads to improvements in the quality, coverage or presentation of existing research.

The National Statistics Authority will accredit third party processors. The ONS is considered an accredited processor but there may be others in future.

### **Annex 3: Current arrangements for data collection by BEIS**

#### ***Sub-National Energy Consumption Statistics***

In December each year, BEIS publish sub-national energy consumption statistics for the previous year. (BEIS has recently received 2017 data, due for publication in December 2018). Statistics are published down to Lower Layer Super Output area (minimum 400 households)<sup>11</sup>.

In June each year, a second document is published which links consumption data and NEED data (National Energy Efficiency Data Framework)<sup>12</sup> covering energy efficiency and socio-demographic data to provide for example breakdowns of consumption by income level.

In 2014 an anonymised dataset of NEED data was made publicly available. This included linked data for energy consumption, energy efficiency measures installed and certain socio-demographic data. The public use file comprises around 50K records. There is also an End User Licenced File available to researchers via the UK Data Archive which contains around 4 million records.

To produce these outputs, estimated gas and electricity annual consumption data are sourced separately by the sub-national energy statistics team in BEIS. Their vires is the Statistics of Trade Act 1947 (which predates DEA & ONS powers). The methodology used (for both the sub-national statistics and for NEED) is published but in summary:

**Gas** – Xoserve on behalf of suppliers compile and provide weather corrected annual consumption gas data for 23 m meters (with their MPRNs) directly to BEIS. They also provide a dataset mapping MPRN numbers to address. So, BEIS receive address-level annual consumption data and assign to geographic areas.

**Electricity** – On behalf of energy retailers, each data collector, as an agent of the supplier, collects and submits annual consumption meter data (by MPAN) into BEIS, who create a single dataset for all 27 m meters from the data submitted. BEIS also obtain a dataset matching MPANs to address from Gemserve (the ECOES database). BEIS then assign the data to geographic areas.

**Smart Meter Data** - Data from properties with a smart meter comes with a 'flag'. There are no plans currently to change the frequency of data collected for smart meters.

**Data-linking** – The energy consumption data submitted to BEIS is also joined at address level by BEIS with NEED data using a Unique Property Reference Number. The NEED data includes property attributes from the Valuation Office (eg floorspace) together with data on energy efficiency measures / microgeneration installed under government schemes and Experian data on household characteristics. Obviously, address-level data is not published and there are restrictions on the individual researchers in BEIS who can access the address level data. To help data confidentiality and

---

<sup>11</sup>

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/678653/Sub-national\\_electricity\\_and\\_gas\\_consumption\\_summary\\_report\\_2016.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/678653/Sub-national_electricity_and_gas_consumption_summary_report_2016.pdf)

<sup>12</sup> <https://www.gov.uk/government/statistics/national-energy-efficiency-data-framework-need-report-summary-of-analysis-2018>

reduce cost the analysis for NEED is now carried out on a sample of 4 million households. Where data is published no sub-category is allowed to fall below 30 observations to prevent re-identification.

### ***DUKES***

UK energy statistics are top-down, not based on granular data. DUKES is compiled from monthly survey data – both on the supply-side and on the demand-side. Every big energy company is surveyed monthly on how much oil, gas etc they have sold. Among energy suppliers the survey extends well beyond the big 6 but is still not comprehensive.

### ***Energy Consumption in the UK***

Energy Consumption in the UK basically uses existing available evidence to produce consumption data. e.g re-presents DUKES data.