

# First research engagement workshop: Smarter Household Energy data: Opportunities and challenges for scaling up research

---

9 December 2015

With generous support from the recent ESRC-NRF centres funding initiative, the UK Data Service has embarked on a new project, *Smarter Household Energy Data: infrastructure for policy and planning*. This is a joint project across two established social science data archives: The [UK Data Service](#), University of Essex, and [DataFirst](#), based at the University of Cape Town in South Africa. The project is building on the current activities of the ESRC Big Data Network Support investment in the UK, by developing a demonstrator for a modern, fit-for-purpose research data infrastructure utilising current industry-standard [Open Data Platform](#) technology in an affordable and accessible way. The platform should enable ready discovery, linkage and visualisation of datasets of interest to the social scientist.

Household energy data has been selected as the use case for this UK-ZA demonstrator project as it is large in scale, and has the potential to be linked to other sources to answer important research questions. Energy usage is a priority issue for citizens and governments in the UK and around the world, with serious implications for policy as well as personal decisions. It is hoped that beyond this demonstrator, the platform can be scaled up to support social and economic research across the spectrum.

The project hosted its first engagement event in central London on 9 December 2015, bringing together UK-based energy specialists, statisticians and data scientists from academia, industry, the third sector, and government, including the Department of Energy and Climate Change (DECC) and Office for National Statistics (ONS). The aim was to scope data needs and challenges, and relevant use cases for the project.

After a series of presentations on energy epidemiology, smart metering and household energy consumption, and on current data management and data science techniques, including scaling up for big data, the experts formed into smaller groups to consider and discuss a series of pertinent questions. These included questions around identifying priority or policy areas where energy data can provide insight, assessing types and formats of useful

data and opportunities for linking across datasets, and considering how stakeholder perspectives converge and differ, for example between government departments and university-based research institutes.

Discussion and brainstorming on these questions produced a wealth of insights, which for convenience can be summarised under five broad headings: data acquisition; legal and ethical considerations; consumer behaviour and interventions; technology; and data utilisation.

- **Ethics and legality** – While the potential benefits of using household data are huge and in the future may well be indispensable, the very idea of monitoring behaviours within households raises serious legal and ethical questions around privacy and consent. At present, government use of household energy data is regulated under the 1947 Statistics and Trade Act, which prohibits analysis of ‘live’ data while still in use. Identification of individual households is also proscribed. Many energy experts feel that the very strict regulations make the most robust public/policy oriented research very difficult, but while reform might be welcomed, public concerns about smart meter data contributing to digital profiling and building a new form of ‘big brother’ must be dealt with responsively and transparently, as security and trust are essential. Existing access to energy data for researcher purposes is covered by the Data Protection Act (DPA), while access to smart meter data will in the future be additionally governed by the Data Access and Privacy Framework, incorporated into the Smart Meter Implementation Programme.
- **Data acquisition** – Given the challenges of securing and using robust energy data, effective acquisition of useful data, re-usable for research emerged as a priority. Standardised metadata would be useful and, where possible, consent for research use should be at point of collection, but this is not always the case with datasets collected for other purposes (such as billing). Seeking consent post-hoc can be so difficult and expensive as to be unfeasible, whilst opt-in consent for research use tends to provide small, biased samples. Triangulation/replication of findings across datasets can be useful, but is not a panacea for limited data, as different questions are asked of the same datasets and across datasets.

Data purchase from commercial organizations with pre-packaged consent is rarely an option and where it is available, is usually prohibitively expensive. A single UK-wide brokerage could be more cost-effective, creating a resource pool administered by the UK Data Service and accessible to the social and economic research community. Some good suggestions included prioritising high-value datasets for better access and buying small samples for exploratory/proof of concept work before committing to larger purchases, although questions around getting good information from 'data poor' areas remain. Backdated data could also be a good resource, freed of some of the concerns around using current information.

- **Data use** – The potential uses of household energy data, particularly in informing energy savings and efficiency, are not merely at the level of individual households, but could inform research and policy at local, national and international levels, as well as in industry. Household energy data can be enriched and made more useful by linking to other data sources, such as health, finance, meteorological, and social and economic datasets. All of these contribute to policy evaluation, assessing the key factors in developing and implementing effective and ineffective policies. Aggregated statistics, such as the average household consumption, often mask diversity in energy profiles which can be critical to energy demand research. Understanding individual buildings' energy performance, and their construction/maintenance are important.
- **Consumer behaviours and intervention** – Before considering how better data might contribute to reshaping consumer behaviours, better knowledge is needed on what consumers use energy for, when, and how, and what measures are most effective to persuade people to change their behaviours, both in the short-term and the longer-term. It was agreed that Intervention studies are particularly useful here, but are very expensive to conduct. For macro-level research, data from national and international smart grids to compare CO<sub>2</sub> in OECD/developing countries are useful. For micro-level investigation, one participant outlined his research on a small number of households who agreed to usage of a fine-grained level of data of appliances in their homes. The appliances were sub-metered to provide data about how much energy was used by a specific kettle or vacuum cleaner at which particular times. Similarly detailed data provided by household energy management

systems from the likes of [Hive](#) and [Nest](#) are much more intrusive and disclosive about household habits than smart meters, but some consumers seem to gladly accept this exchange of information in return for potential efficiency and savings benefits, provided they believe their data will be managed securely. Querying this data can inform not just better individual decisions, but could also contribute to Demand Side Response/Management, improved facility management, inform policy, and new research.

- **Technology** – many systems currently in use are stretched and will struggle to cope with new big data streams, but the range of platforms and tools which might be used to manage and analyse energy data have proliferated. Although the technical requirements of constructing big data systems are not always well understood in academia, the trend among the scholars present was a combination of older and newer software products, with Stata and R playing prominent roles. The UK Data Service Big Data team members took the opportunity to showcase the ODP model, and to demonstrate a small proof of concept solution visualising a 'heat map' of household energy usage across UK regions.

Overall, the organising team were pleased that the discussions covered much known territory already being debated in the energy research domain. They unearthed some pragmatic ideas about trying to secure backdated data, and to promote the idea of academic partnership and big data experiments to the energy companies. Given that the purpose of the exercise was engagement with the research community and scoping researchers' data service needs, this can be judged as a resounding success. The event generated a meeting of minds between leading UK experts, and has provided clear lines of work and development in scoping a robust data infrastructure enabling research supporting smarter household energy policy and planning. A similar engagement workshop will be held in Cape Town in January 2016.

If you are interested in contributing a research use case please let us know.

Project: <http://www.ukdataservice.ac.uk/about-us/our-rd/smarter-household-energy-data>

Blog: <http://ukdataservicesmartenergydata.wordpress.com>