Crowdsourcing in energy research

Introduction

In his original definition, (Howe, 2006) describes crowdsourcing as the activity happening when an organisation decides to outsource a specific task to a crowd. The main elements of the term are the “open call format”, which refers to the task being open for anyone to participate, and also the crowd element, which refers to the task being outsourced to a large undefined number of workers. The term has since been expanded by (Brabham, 2008; Brabham, 2012) to refer to a task issued by an organisation to an open online community which then contributes in performing that task for the organisation’s interest. Crowdsourcing work has been used from organisations in many applications ranging from product design and urban planning (Brabham, 2008; Brabham, 2012) to crowdsourcing for participation in policy-making and constitutional reform (Aitamurto, 2012).

Broadly, crowdsourcing can be split into voluntary and paid crowdsourcing (Aitamurto, 2012). Voluntary crowdsourcing resembles a modern form of citizen science where volunteers, usually non-experts community members, contribute to a data collection task (Catlin-Groves, 2012). Paid crowdsourcing refers to microtasks issued through an online platform such as Amazon Mechanical Turk. Microtasks in this case can be very simple and quick tasks such as identifying objects in images and transcribing audio (Kittur, et al., 2008).

As (See, et al., 2015) also argue, crowdsourcing is only just starting to be used in the field of energy research for data collection and analysis tasks. In contrast, other research fields, such as ecology, conservation or mapping, have traditionally involved citizens in scientific studies and primarily for data collection tasks. An example of a successful citizen science project is OpenStreetMap[1], a free and editable map developed and maintained by volunteers since 2004 which now has almost 6.5 million registered users (OpenStreetMap, n.d.).

In the energy sector, the availability of accurate and realistic data is crucial for the development of suitable energy modelling tools. In this context, crowdsourcing can be an effective method of collecting energy data that would otherwise be difficult to collect (Bazilian, 2012). A common application is the collection of geographic-specific data such as the location of small renewable energy installations (e.g. rooftop PV or solar-thermal systems) (Sagl, et al., 2014). Data collected through crowdsourcing can then be used in the development and validation of energy models, which usually form the basis of decision making in governments and local organisations. Better

[1] https://www.openstreetmap.org/

validated models can therefore lead to more informed decisions by leaders locally and nationally. Crowdsourcing can also help reduce costs and time in data collection, by having community members contribute with observations to systems that they already have access to (Bazilian, 2012). Finally, in the renewable energy sector, crowdsourcing has been found to be an effective way to engage with citizens and increase awareness of renewable energy and its benefits (Sagl, et al., 2014).

Three examples of crowdsourcing in energy research

Three case studies of crowdsourcing for collection of energy-related data are presented. For each case a brief description of the project is given, the need and motivation for crowdsourcing are discussed and finally each project's success is evaluated.

Repowermap.org

Repowermap.org[2] is a non-profit initiative that originated within the framework of the Intelligent Energy Europe programme and supported by the European Union (repowermap.eu, 2015). The scope of the initiative is to promote renewable energy and energy efficiency by making examples of such projects visible on an online map and by facilitating information exchange (repowermap.eu, 2015; See, et al., 2015). Any citizen can contribute by providing information on their local renewable energy installation or building energy efficiency measure which then becomes visible on the map. Organisations, businesses and local authorities can also contribute their own projects and they are encouraged to integrate the map to their website to promote public participation. The initiative originally focused on ten European countries, which were part of the Intelligent Energy Europe programme, but has since been expanded to other countries. The main partners to this project include renewable energy associations, NGOs, local authorities research institutes and businesses across Europe.

The main driver behind this initiative has been the lack of visibility on local examples of renewable energy projects which impacts people's confidence in technology and holds up progress in installations (repowermap.eu, 2015). By providing a map which showcases examples of installations at the neighbourhood level, people can become more widely aware of projects, gain confidence in promoting their own installations as well as become motivated to use a technology or adopt a specific energy measure that they have seen in a neighbour. The initiative also aims to facilitate information exchange between all stakeholders using the map including local organisations and the public. Information exchange can take place in the form of sharing best practices, learning about specific local conditions or discovering companies that are active in a particular area as technology installers or suppliers (repowermap.eu, 2015). The map is also designed to show events held by local organisations as site visits to specific installations.

[2] http://www.repowermap.org/

The initiative began with the goal of collecting information on 40,000 individual renewable energy or energy efficiency projects. It currently has information on almost 74,000 projects and although the majority of them are in the initial 10 countries, the initiative has gained support and is being promoted by organisations globally thus extending its geographic coverage. The map has also been integrated in the websites of high-profile organisations such as WWF Europe and the Climate Alliance. However, at the moment the information gathered can't be used outside the mapping tool and the website provides no means to download or access the data. (See, et al., 2015) also point out that the information repowermap.org gathers on each project has not been determined by energy modellers and is therefore quite basic. In fact, information is limited to the companies involved in the project and the installation's power capacity. Together with the inability to access the data, this makes the overall map unsuitable for use in any type of detailed energy modelling. Finally, (See, et al., 2015) argue that the project doesn't give any financial incentives to people to add their projects to the map which could hinder its expansion. The project's success so far has only relied on support from local organisations. Although this might have been sufficient so far, having a financial incentive and potentially some gamification element could make people contribute more.

OpenGridMap

The OpenGridMap[3] project started from researchers at the Technical University of Munich with the goal of creating an open platform of power grid data. The researchers, inspired by OpenStreetMap, aim to develop a map of electric power grids worldwide using crowdsourcing for geographic-specific power grid infrastructure data (Rivera, et al., 2015). Apart from crowdsourcing the data, the project also aims to provide it freely in the required format for power grid simulations, thus supporting other researchers in simulation studies (Rivera, et al., 2017). To support their crowdsourcing approach, the researchers have developed a mobile app, where participants can submit images of grid components with explanations on the type of grid component and the geographic location. The crowdsourced data are verified by an expert before being uploaded on the OpenGridMap platform. The verified data are then structured using a Common Information Model (CIM) to make it suitable for power simulation studies.

The team behind the OpenGridMap project was driven by the lack of available power grid data to researchers particularly at the distribution network level (Rivera, et al., 2015). Knowledge of the location and state of the grid infrastructure is highly important for analysing power flows and assessing the resilience of the grid to changes. This is particularly important when looking at the distribution level where most changes take place (more distributed power generation being connected, distribution loads rising due to electrification of heat and transport, etc.). The distribution grid is also where the lack of data is most apparent. Even the power grid data owned by utilities and network operators, are not often publicly available for use by researchers, or worse they are available at high cost (power-technology.com, 2017). This means that in most cases, research in this field takes place using design

[3] http://opengridmap.com/

data and standardised test cases. The motivation, therefore, behind the OpenGridMap project is to create the framework where actual power grid infrastructure data are gathered and used for realistic grid simulations.

The project has so far been highly successful in creating a near-complete map in Germany, with almost 99% of the country's power grid mapped (power-technology.com, 2017). Other European country maps are also getting developed with the UK's grid being mapped at almost 82%, but other non-European countries lagging behind. It is understandable that the project has been most successful in its home country, where the team can more effectively promote their project. However, this is also what could restrict the project's global success as completion of maps is dependent on the level of engagement from communities in each country (power-technology.com, 2017). To increase contribution the team has suggested developing gamification schemes to incentivise people (Rivera, et al., 2015). Another way that the research team could promote the project further globally is by collaborating with local researchers, institutes and organisations in each country. Also, supporting more international languages on the mobile app could help expansion. Finally, providing a platform or map where participants can visually see their contributions may support engagement and further contribution from the community.

Surveying household energy behaviour using paid crowdsourcing

In this study, (Pritoni, et al., 2015) explored how people in the US use programmable thermostats to heat their homes. (Pritoni, et al., 2015) used online paid crowdsourcing through Amazon Mechanical Turk to collect data. By taking this approach, the research team wanted to explore the effectiveness of crowdsourcing when compared to traditional phone or written surveys. To collect data, the team released a survey with questions on thermostat usage and also gave the option to participants (Turk workers) to upload photos of their thermostat. The researchers performed two iterations of the survey, using the first one as trial and the second to collect the actual data. The actual survey had 200 participants who uploaded 31 photos and was completed within a week. The researchers specified taking a number of measures for ensuring reliability of responses such as qualifying workers based on their quality rating and filtering based on three “nonsense” questions and completion time.

(Pritoni, et al., 2015) were driven by the need for more current and reliable information on how households use thermostats in order to improve understanding of energy savings. Other studies that explored indoor home temperatures have also identified the need for more reliable data on household thermal comfort and energy behaviour (Vadodaria, et al., 2014). In the US, such information is usually available from national large-scale surveys on residential energy consumption such as (EIA, 2015). However, surveys like the (EIA, 2015) are very expensive to run and only take place every 3-5 years. They also rely on participants self-reporting information on temperature settings which could be unreliable (Kempton, 1987; Pritoni, et al., 2015). (Pritoni, et al., 2015) were therefore motivated to explore an alternative method for collecting energy behaviour data that makes use of online paid crowdsourcing services as a way to reduce cost and time and also have the ability to include photos which can be used to validate survey responses.

As (Pritoni, et al., 2015) describe, survey responses from this study regarding household energy usage were in line with responses in national energy surveys however some key methodology considerations were made apparent. The most significant one was that self-reported data can be inaccurate. By using a crowdsourcing service where respondents can upload photos, the researchers were able to collect both quantitative and qualitative data and cross-validate them to identify inaccuracies in survey responses. Other advantages include this method being inexpensive and quick to do. (Pritoni, et al., 2015) were able to do a trial survey first to test the credibility criteria for uploading photos and then run the actual survey. It also means studies like this can be repeated many times during the year to capture seasonal behaviour changes. Nonetheless, online paid crowdsourcing is not perfect and still has many limitations such as non-representativeness of the sample and reliability concerns due to the intrinsic motivation of paid work (Paolacci & Chandler, 2014).

Challenges and the value of crowdsourcing in energy research

The challenges of crowdsourcing as a method for data collection are not unique to energy research and they are well known and documented in literature. The three examples presented earlier reveal three key challenges: maintaining motivation of participants, ensuring data quality and overcoming sampling bias in paid crowdsourcing.

Firstly, to maintain participants engaged, the role of communities is very important and projects like repowermap.org rely on local organisations to drive crowdsourcing activities and promote participation (repowermap.eu, 2015). Another way to address this challenge is by introducing gamification techniques to incentivise participants (Rivera, et al., 2015). In terms of data quality, citizen science projects have often been challenged in terms of the reliability of the data collected (Catlin-Groves, 2012). Quality assurance must definitely be a part of crowdsourcing projects and methods to ensure this include having experts to verify the crowdsourced information (Rivera, et al., 2015) or introducing ‘quality standards’ especially when it comes to submitting photos (Sagl, et al., 2014). Finally, it is known that paid crowdsourcing can lead to sampling issues with underrepresentation of certain population groups and unreliability of workers which could have different motives for participation. These issues highlight the importance of methodology and the need for recording sampling characteristics in paid crowdsourcing studies (Paolacci & Chandler, 2014).

Crowdsourcing in energy research is still in early stages and its value and benefits are just beginning to be realised. Crowdsourcing can be an incredible tool for filling in gaps in our knowledge of the energy system, especially when it comes to data that are not readily available for use in academic research. Examples might be power grid infrastructure data, which are usually owned by regulators and only available partially or at a cost, and data on consumer energy behaviour, which come either at the cost of a paid survey or researchers end up using national energy survey data which are out of date and can miss on behavioural trends. In addition, crowdsourcing can be valuable for bridging the gap between real and modelled energy analysis. Quantitative analysis using energy models is what forms the primary basis of energy research. Part of the validity of energy models comes from having realistic input data and crowdsourcing can support collection of such data. Finally, crowdsourcing can be a powerful tool in educating the public and improving awareness of energy issues. Even though crowdsourcing is still considered a niche in energy research, its potential and value are evident especially in data collection and citizen engagement.

Bibliography

Aitamurto, T., 2012. *Crowdsourcing for Democracy: New Era in Policy-Making.* s.l.:Publication of the Committee for the Future.

Bazilian, M. e. a., 2012. Open source software and crowdsourcing for energy analysis. *Energy Policy,* Volume 49(C), p. 149–153.

Brabham, D., 2008. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases.. *Convergence, 14(1),* pp. 75-90.

Brabham, D., 2012. THE MYTH OF AMATEUR CROWDS: A critical discourse analysis of crowdsourcing coverage.. *Information, Communication & Society: AoIR Special Issue, 15(3),* pp. 394-410.

Catlin-Groves, C., 2012. The Citizen Science Landscape: From Volunteers to Citizen Sensors and Beyond.. *International Journal of Zoology,* p. 14.

EIA, 2015. *Residential Energy Consumption Survey (RECS).* [Online]
Available at: https://www.eia.gov/consumption/residential/
[Accessed 19 July 2020].

Fattori, F., Albini, D. & Anglani, N., 2016. Proposing an open-source model for unconventional participation to energy planning.. *Energy Research & Social Science,* pp. 12-33.

Gardumi, F. et al., 2018. From the development of an open-source energy modelling tool to its application and the creation of communities of practice: The example of OSeMOSYS.. *Energy Strategy Reviews,* pp. 209-228.

Howe, J., 2006. *Crowdsourcing: A Definition.* [Online]
Available at: https://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html
[Accessed 19 July 2020].

Hyysalo, S., Juntunen, J. & Freeman, S., 2013. User innovation in sustainable home energy technologies.. *Energy Policy, 55,* pp. 490-500.

Intelligent Energy Europe, n.d. *repowermap.org.* [Online]
Available at: http://www.repowermap.org/
[Accessed 13 July 2020].

Jacobsen, P., n.d. *OpenGridMap.* [Online]
Available at: http://opengridmap.com/
[Accessed 19 July 2020].

Karnouskos, S., 2011. *Crowdsourcing information via mobile devices as a migration enabler towards the SmartGrid..* s.l., s.n., pp. 67-72.

Kempton, W., 1987. Thermostat management: Intensive interviewing used to interpret instrumentation data. *Energy Efficiency: Perspectives on Individual Behavior,* pp. 245-262.

Kittur, A., Chi, E. H. & Suh., B., 2008. *Crowdsourcing user studies with Mechanical Turk..* New York, s.n., pp. 453-456.

OpenStreetMap, n.d. *OpenStreetMap.* [Online]
Available at: https://www.openstreetmap.org/
[Accessed 13 July 2020].

Ornetzeder, M. & Rohracher, H., 2006. User-led innovations and participation processes: Lessons from sustainable energy technologies.. *Energy Policy, 34(2),* pp. 138-150.

Paolacci, G. & Chandler, J., 2014. Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science,* Volume 23(3), pp. 184-188.

Pfenninger, S. et al., 2017. The importance of open data and software: Is energy research lagging behind?. *Energy Policy, 101(C),* pp. 211-215.

power-technology.com, 2017. *Where does it all go? Mapping power worldwide using the OpenGridMap.* [Online]
Available at: https://www.power-technology.com/features/featurewhere-does-it-all-go-mapping-power-worldwide-using-the-opengridmap-5752438/
[Accessed 13 July 2020].

Pritoni, M. et al., 2015. Energy efficiency and the misuse of programmable thermostats: The effectiveness of crowdsourcing for understanding household behavior.. *Energy Research & Social Science, 8(C),* pp. 190-197.

repowermap.eu, A., 2015. *Achievements in building up the repowermap.org initiative within the framework of the Intelligent Energy Europe Programme,* s.l.: Intelligent Energy Europe Programme.

Rivera, J. A., Goebel, C., Sardari, D. & Jacobsen, H., 2015. *Opengridmap: An open platform for inferring power grids with crowdsourced data..* s.l., s.n., pp. 179-191.

Rivera, J., Leimhofer, J. & Jacobsen, H., 2017. OpenGridMap: Towards automatic power grid simulation model generation from crowdsourced data.. *Computer Science - Research and Development, 32(1),* pp. 13-23.

Sagl, G. et al., 2014. *Crowdsourcing Energy Data for Participatory Renewable Energy Planning and Modeling..* Vienna, s.n.

See, L. et al., 2015. The Potential of Crowdsourcing for the Renewable Energy Sector. *Handbook of Clean Energy Systems*, July.

Vadodaria, K., Loveday, D. & Haines, V., 2014. Measured winter and spring-time indoor temperatures in UK homes over the period 1969–2010: A review and synthesis. *Energy Policy,* Volume 64, pp. 252-262.