Just Good Enough Data and Environmental Sensing: Moving Beyond Regulatory Benchmarks toward Citizen Action

Jennifer Gabrys[1] and Helen Pritchard[1]
[1]Citizen Sense, Goldsmiths, University of London
j.gabrys@gold.ac.uk

Forthcoming in

Environmental Infrastructures and Platforms 2015 - Infrastructures and Platforms for Environmental Crowd Sensing and Big Data
Proceedings of the Workshop

co-located with the European Citizen Science Association General Assembly 2015 (ECSA GA'2015)
Barcelona, Spain, October 28-30 October, 2015.

Edited by
Arne J. Berre * Sven Schade **
Jaume Piera ***
* SINTEF, Oslo, Norway
** JRC , Joint Research Centre, Ispra, Italy
*** ICM-CSIC , Institute of Marine Sciences, Barcelona, Spain

Draft (2015) published at
http://ecsa.citizen-science.net/sites/ecsa.citizen-science.net/files/ENVIP-2015-Draft-Binder.pdf

Just Good Enough Data and Environmental Sensing: Moving Beyond Regulatory Benchmarks toward Citizen Action

Jennifer Gabrys[1] and Helen Pritchard[1]

[1]Citizen Sense, Goldsmiths, University of London

j.gabrys@gold.ac.uk

Abstract

Discussing a research project that investigates citizen sensing practices in relation to monitoring air pollution from hydraulic fracturing sites, this paper investigates the types of data that citizen monitoring generates, and the uses to which it might be put. The discussion is located within the wider context of the rise of environmental sensing technologies and practices that are emerging and that seek to enable citizens to use DIY and low-tech monitoring tools to understand and act upon environmental problems such as air pollution. These "citizen sensing" projects intend to gather data, which can indicate environmental change and give rise to political action. However, regulators often contest citizen-gathered data as inaccurate, and as collected through sub-standard instruments and practices. Drawing on a report developed by the US EPA, we use the concept of "just good enough data" to demonstrate that citizen-gathered data can have multiple other uses beyond regulatory comparison and compliance. Describing the collaborative development of an environmental monitoring kit, as well as the deployment of this kit within a participatory research setting, we suggest that the relevance of citizen-collected air quality data should not be solely evaluated through absolute criteria such as alignment to state- or federally managed air quality data, but rather should be incorporated for the unique citizen-based insights and perspectives it provides.

Keywords: Citizen sensing, citizen science, just good enough data, air pollution, participatory research, environmental regulation

1. INTRODUCTION

A number of environmental sensing technologies and practices are emerging that seek to enable citizens to use DIY and low-tech monitoring tools to understand and act upon environmental problems. These "citizen sensing" projects intend to gather data, which can indicate environmental problems and give rise to political action (cf. Burke et al., 2006; Cuff, Hansen and Kang, 2008; Goodchild, 2007). One of the primary ways in which citizen sensing projects have sprung up is through engagement with environmental pollution. Noise, air, soil and water pollution are local if distributed environmental disturbances that can now be monitored using a range of digital sensing devices (Aoki et al., 2008; Maisonneuve et al., 2009; Paulos et al., 2009). These devices can be mobile, and provide new data in comparison to fixed monitoring stations, while allowing citizen sensors to understand personal exposure more fully. A key motivating aim of many citizen sensing projects is to attempt to make the data gathered about environmental pollution a catalyst for political action. Such an objective could be seen to be a continuation of earlier citizen science initiatives that focused on gathering data or otherwise intervening within scientific practice in order to provide alternative forms of evidence based on different and diverse citizen experiences (Irwin, 1995; Jasanoff, 2003).

In this way, citizen-generated data sets are often gathered with equipment that diverges from state and regulatory standards, and through practices that differ from standard scientific protocols. When monitoring air pollution or other environmental disturbances with low-cost technology, citizen-led initiatives are typically challenged about the validity or accuracy of their data. Environmental regulators at times dismiss citizen-collected data since it is perceived to be biased, lacking in standardised procedures for collection, and generated through imprecise instruments. Yet citizens now deploy environmental monitoring technologies in multiple contexts, and the amount and type of environmental data that they collect continues to grow. While citizen sensing technologies and practices might not typically involve consistently observing air pollution with sophisticated instruments to meet regulatory standards and to ensure compliance with air pollution policy, they do involve capturing fine-grained pollution data through DIY devices that are often located in spatially dense networks, and which further provide ongoing indications of changes in air quality, rather than absolute measurements.

As the US Environmental Protection Agency has expressed in its analysis of new modes of “next generation” environmental monitoring, “types of data” and “types of uses” are interlinked (US EPA, 2013). Data typically only become admissible for legal claims when gathered through specified scientific procedures by state certified users with approved (as well as expensive) instrumentation. However, there may also be situations in which data gathered through citizen sensing practices are “just good enough” for establishing that a pollution event is happening. It therefore remains a relatively open question as to what the uses and effects of data gathered through citizen sensing technologies might be, since these are data practices that are still emerging.

“Just good enough data” is a phrase and concept that we use here to address issues of accuracy in relation to citizen-collected data, and to expand the types of uses that might accompany these new types of data. “Just good enough data” draws attention to attempts to counter the reliance on high levels of measurement accuracy as the sole criterion by which data are evaluated. What different practices emerge when environmental data are engaged with in a more indicative register? What do these practices enable? And what other relations, connections and points of focus might "just good enough data" generate?

Examining these questions in the context of participatory environmental sensing conducted by the Citizen Sense research group in northeastern Pennsylvania, this paper considers how the use of air pollution monitors by residents living next to hydraulic fracturing (or fracking) infrastructure produces different registers and types of data. The paper outlines citizen sensing practices that monitor fracking-related pollution that are already underway, and it discusses our attempts to contribute to monitoring processes through further participatory and practice-based citizen sensing initiatives. The paper reviews the multiple forms of data generated through this participatory citizen sensing project that diverge from state and regulatory monitoring, including air quality data, data logs, citizen observations and stories, as well as a data analysis tool developed by Citizen Sense to facilitate citizen-led analysis of their data collected over 9 months. The paper further discusses how residents attempted to mobilise data and engage in discussions with regulators, and the ways in which citizen-gathered data could provide other insights beyond a regulatory-only focus on monitoring.

Citizen-sensed data is rich with trans-local experiences and collective insights, often bringing attention to environmental change from new perspectives. Rather than limit discussions of citizen sensing practices to accuracy and regulatory alignment, we investigate how to develop practices and infrastructures for "just good enough data" that enable citizen-sensed data to make expanded contributions to environmental sensing. We suggest that the relevance of citizen-collected air quality data is not solely determined through absolute criteria, or alignment to state- or federally managed air quality data, since these criteria can often shift depending upon modes of governance, location, and available resources. If data are understood instead as entities that transform depending upon the uses to which they will be put--and how "good enough" they might be to achieve these ends--it then becomes possible to attend to how data are differently mobilised through monitoring practices and political encounters (cf. Gabrys, 2016).

2. AIR QUALITY MONITORING AND NATURAL GAS EXTRACTION

Unconventional natural gas extraction in the form of hydraulic fracturing began in the Marcellus Shale region in Pennsylvania in 2003, however, by around 2006 the number of wells drilled in the state began to increase rapidly and communities began to notice more intensively the impact of the industry (State Impact, n.d.). At the time of writing this paper, permits have been given for almost

17,000 wells, and nearly 7,800 wells are in operation as sites of natural gas production, with more wells becoming active daily. As recorded by a local citizen-led website that collates and provides information and data on gas well production in Pennsylvania, on average one new well was opened every two days in the state during September 2015 (MarcellusGas.Org, n.d.).

Many of these wells and the related natural gas infrastructure of compressor stations, well pads, glycol dehydrators, water impoundment ponds and pipelines are densely located in northeastern Pennsylvania. Along with this infrastructure, inevitable concerns have arisen about environmental impacts, especially in relation to water and air pollution. While much attention has been given to water pollution through several high-profile cases of contaminated well water in areas of northeastern Pennsylvania, residents of this community have also had concerns about the relatively under-monitored effects of fracking on air quality.

In part, Pennsylvania residents' interest and sense of urgency in developing monitoring practices has also been in response to the lack of governmental monitoring in this rural area. Existing monitoring for the nationwide Air Quality Index (AQI), which is facilitated on a state level by the Pennsylvania Department of Environmental Protection (DEP), typically focuses more intensely on urban areas and roadside sites, and does not have a particular remit or attention to accounting for emissions from particular industries such as oil and gas. In this way, DEP stations for monitoring air quality and criteria pollutants such as particulate matter 2.5 ($PM_{2.5}$) are located in relatively distant urban centres such as Scranton, where monitors are often placed close to busy highways. Although the DEP also do undertake some mobile monitoring on a sporadic basis, due to economic and political constraints there is no consistent monitoring taking place by regulators that would fully account for local emissions from the natural gas industry in the northeast of the state. Within the context of a newly expanding industry that residents felt was not sufficiently monitored, an interest then emerged to develop techniques for documenting environmental pollution in this area.

In order to understand the air pollution arising from the processes of natural gas extraction and production, residents of Pennsylvania have undertaken many diverse practices of monitoring with differing aims and objectives to gain a more immediate sense of their environmental conditions. Attempting to capture their experiences of air pollution and associated health effects, residents in this area have used an extensive range of monitoring technologies either on their own, or collaborating with or assisting scientific studies. For instance, many monitoring practices have required that residents collect samples for lab analysis, which takes place in distant sites of data processing. Or they have required residents to use technologies that produce data in forms that are not immediately useable or comparable to other datasets. The promise of low-cost and "next-generation" environmental sensors is that data will be made available in real time, in a legible output, to the users of the technologies.

3. A PARTICIPATORY APPROACH TO CITIZEN SENSING

In a participatory approach (cf. DiSalvo et al., 2012) to developing citizen sensing practices and technologies further with residents of northeastern Pennsylvania, the Citizen Sense research project held a series of discussions and monitoring events with residents during 2013 to 2015. Through this collaborative process, we developed the Citizen Sense Kit for the citizen-based monitoring of air quality in the region. After a period of developing the Citizen Sense Kit with participants, the Citizen Sense research project deployed the kits in October 2014 with a training workshop and walk to field test the technologies. The Citizen Sense research team then undertook visits to participants' homes to help set up the technologies, and participants developed a number of situations and experiments to monitor areas of particular concern to them.

The Citizen Sense kit distributed during these events was developed in response to the concerns of community members, who provided information via Citizen Sense "logbooks" that asked for input on what the key concerns were for natural gas infrastructure in relation to air pollution. The Citizen Sense Kit, which was distributed to around 30 participants, contained a passive sampling badge for monitoring BTEX emissions (or benzene, toluene, ethylbenzene and xylene, which are substances associated with gas production and that are also hazardous to human health—see Moore, 2013); along with a "Speck" device from the Create Lab at Carnegie Mellon for monitoring $PM_{2.5}$. The kit also includes a custom-made prototype device, the Frackbox, which was installed at three compressor station sites. The Frackbox ran off a RaspberryPi and included sensors for monitoring nitrogen oxide (NO), nitrogen dioxide (NO2), ozone (O3), and volatile organic compounds (VOCs), as well as wind speed, humidity and temperature. Participants were able to upload the data they gathered to the Citizen Sense Kit platform, and to refer to a Citizen Sense logbook with instructions for use of the various part of the kit.

The Citizen Sense Kit was used by a wide range of residents living near infrastructure, and also taken up by a local group, Breathe Easy Susquehanna County (BESC), which was interested in maintaining constructive dialogue with industry about changes in the environment particularly in relation to air quality. The kit attempted to provide accessible and unobtrusive ways for participants to document pollution events and experiences, and to observe patterns and relations that emerged from collected data. This approach to environmental sensing was important for a number of reasons. Due to the sensitive context where air quality monitoring of fracking infrastructure was taking place, many participants needed to take part anonymously, as fears of reprisal from neighbours and industry were very present. The small size of the kit meant that it could be installed and used in an inconspicuous way, where citizens could install sensors on their porches, as well as in gardens, sheds and under overhanging eaves near homes for a duration from 3 to 6 months. As participating residents also lived distributed across the local region, there was a spatially dense concentration of over 20 individual monitoring locations, rather than the 2 to 3 monitoring points that might be found across rural areas of an entire state. This in turn gave rise to the possibility of identifying localised sources of emissions, which could be read together with state air quality data.

While data was collected and logged on the Citizen Sense platform, some participants began to notice patterns in their own data, particularly in the $PM_{2.5}$ data sets. Using the data together with additional sources of weather data, including wind speed and wind direction data from Weather Underground, participants were able to rule out spikes in their data that were most likely caused by regional sources and instead focus their energies on pollution events over more than 6 hours at time when the wind speed was lower, which would indicate a more local pollution source. The participants who knew each other also formed groups so they could compare their data with each other.

The Citizen Sense Kit for monitoring air quality did not just focus on the gathering of numeric data, however. Photographs, mobile phone videos, YouTube comments, FLIR camera footage, diaries, and multiple other forms of documentation that on one hand might have seemed like a disparate set of resources, all contributed to the making of a "just good enough" collective data-set for the region.

Participants were further able to use their local expertise about fracking processes and infrastructure, in particular in relation to compressor stations, to answer the questions of the regulatory bodies, which had little day-to-day experience of living so close to natural gas extraction infrastructure. Another participant set up two monitors at a site opposite to a location that was scheduled to be fracked. Due to the temporal and unpredictable nature of much fracking, it had been difficult for regulators to monitor a well pad from start to completion. Companies may have a permit to drill a well for five years, and often may start fracking without warning. In comparison, one participant who passed a potential well pad site daily was able to establish a period of monitoring data before the fracking took place, as well as during the fracking operation. This monitoring combined with the participant's daily YouTube videos documenting ongoing drilling and fracking underway, have in turn contributed to a unique set of evidence that can be read alongside more official regulatory monitoring data.

Although in the view of the regulators the data generated by the Citizen Sense research project was not comparable to AQI air quality data, it was however "just good enough" for the participants to read together with state-collected air quality data and locally collected wind data from Weather Underground. The distribution of devices also contributed to recognising a regional source of $PM_{2.5}$ in the area, which was good enough to form a pattern that could be excluded when looking at the local sources. One device on its own would probably not have been "just good enough," but the distribution of devices, maintained by participants on a day-to-day basis over 6 months, made the data useful for entering into discussion with regulators, since in some cases even regulators and industry are unsure what is being emitted from these sites of concern.

Data that emerged through these techniques further became a useful negotiation tool. It was used to arrange a number of conference calls with regulatory bodies such as the Center for Disease Control and Prevention (CDC), the Agency for Toxic Substances and Disease Registry (ATSDR), the Pennsylvania Department of Health (DOH), and nonprofit environmental organisations as well as local political representatives. Although responses to the citizen-collected data ran the spectrum from outright dismissal to interest, there was just enough evidence to lead to one environmental agency requesting that local monitoring be undertaken, something which BESC participants had been campaigning for since the inception of their organisation.

4. DISCUSSION AND CONCLUSIONS

Although some citizen sensing projects have worked closely with regulators and scientific disciplines, many others have departed from these practices, and instead have used devices in unconventional ways, creating infrastructures that might be very different both spatially and temporally from those of the regulators. As citizen science and citizen sensing stabilize, there is a call for practices to become more standardised to enhance the legitimacy of citizen-monitoring efforts. Indeed, the US EPA (2015: 28) cites the need to establish the standardisation of protocols and data sets, and both the North American Citizen Science Association (CSA, 2014), and the European Citizen Science Association (ECSA, 2015) cite the need to establish best practice guidelines as central to the aims of the organizations. Gatherings such as the ECSA assembly and Citizen Science Initiative Switzerland are specifically coming together to form working groups to address this problem of standardisation.

Much of the ongoing debate by practitioners and organisations that we have observed in citizen science meetings focuses on the importance of developing practices that can be directly comparable to existing regulatory practices. To some extent, there is a gap between the current citizen sensing infrastructure and this vision of comparability. This has led to a drive toward designing devices that create data in similar formats, and to the calibrating of devices in reference to regulatory monitoring equipment. In some cases, we have observed regulators recommending that citizens should only monitor in scenarios that are pre-approved by official bodies. Yet in this context, the inevitable question arises as to what new possibilities for environmental monitoring and citizen-gathered data might be missed by attending only to the ways in which citizen sensing practices might replicate monitoring practices focused on regulatory compliance.

Citizen-gathered data using next-generation environmental sensing could have multiple uses, and the trajectories of citizen sensing initiatives in making connections from environmental data to action do not need to exclude data that does not conform to regulatory practices, or which might have a more speculative starting point. Such an approach would imply that any production of data by citizens that does not aim toward regulatory targets and processes could not be useful. We recognise that for some contexts these new arrangements of infrastructures have proved challenging to both regulators and scientists, whose disciplines and professions have established practices of measuring, monitoring and accounting for environments. This, in turn, has often resulted in creating points of tension and disagreement between regulators and citizens.

But making citizen sensing practices and data conform only to regulatory standards would be to exclude the other creative and political possibilities of what we are calling “just good enough data.” And to align data practices exclusively with regulatory modes of monitoring might even exclude citizens from any participation in citizen sensing completely. For instance, to be comparable to the state DEP and US EPA air quality data in the context of air quality monitoring for $PM_{2.5}$, citizen data would need to be collected by officially trained personnel on approved equipment. Further to this, to be comparable to regulatory data, monitoring would have to take place at the very same location, height and position in which regulatory monitoring is already situated. In the context of the AQI $PM_{2.5}$ monitoring, citizen monitoring would then also have to be done over a timespan of 3 years. One could argue that as the data analysis process is also one of many decisions, data would have to be analysed (including averaged and smoothed) using the same software and algorithms as state and federal processes. In this scenario, citizen sensing as a practice would become completely redundant as the process would have to replicate the monitoring performed by governmental agencies and experts, rather than opening up opportunities for monitoring to be undertaken by a wider range of participants, in varied locations, over different timescales, and in response to distinct events.

Instead, we suggest that “just good enough data,” while not ignoring the important issues of accurate instrumentation, calibration, and measurement, along with robust monitoring practices, might also allow more expansive uses of citizen sensing technologies and data, while still opening up a dialogue on environmental change between citizens and regulators. With this proposal, we are not regressing to earlier conceptions of public science, where the collection of data is a cursory one oriented toward raising public awareness, but on the contrary we suggest that "just good enough data" is a practice that creates a shared space for discussion that can communicate community awareness of pollution events to regulators. Citizen-produced data sets are often declared to be inaccurate due to the devices used, illegitimate due to the protocols followed, and unscientific due to perceived community bias (such as citizens monitoring to create deliberate evidence for pollution events). However, we have shown that citizen sensing is also an entry point for testing the claims about the ease of participation that next-generation environmental sensors are meant to offer, as well as for developing expanded aspects of monitoring, data collection and environmental politics, which might allow communities to engage more readily with environmental problems.

5. ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313347, "Citizen Sensing and Environmental Practice: Assessing Participatory Engagements with Environments through Sensor Technologies." Thanks are due to participating residents in Pennsylvania, including Frank Finan, Rebecca Roter, Meryl Solar, Vera Scroggins, Chuck and Janis Winschuh, Paul Karpich, Barbara Clifford, John Hotvedt, Barbara Scott, Audrey Gozdiskowski, and Alex Lotorto, along with anonymous participants, as well as previous Citizen Sense researchers, including Nerea Calvillo, Tom Keene, and Nick Shapiro, and consultants including Kelly Finan (illustration) and Dr Benjamin Barratt (atmospheric science). Thanks are also due to the Create Lab for loaning Speck devices for use in this study.

6. REFERENCES

Aoki, P. M., R. J. Honicky, A. Mainwaring, C. Myers, E. Paulos, S. Subramanian, and A. Woodruff (2008). "Common Sense: Mobile Environmental Sensing Platforms to Support Community Action and Citizen Science", *Adjunct Proceedings UbiComp 2008*, pp. 59-60. Seoul: ACM.

Breathe Easy Susquehanna County, at https://www.facebook.com/BreatheEasySusq [accessed 7 January 2016].

Burke, J., D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and N.B. Srivanstava (2006). "Participatory Sensing", *Proceedings of the World Sensor Web Workshop*, Boulder, CO: ACM SENSYS, 2006.

Citizen Science Association (2014) "Vision, Mission, Goals", at http://citizenscienceassociation.org/overview/goals/ [accessed 8 January 2016].

Cuff, D., M. Hansen, and J. Kang (2008). “Urban Sensing: Out of the Woods”, *Communications of the Association for Computing Machinery* 51(3): 24-33.

DiSalvo, C., M. Louw, D. Holstius, I. Nourbakhsh, and A. Akin (2012). “Towards a Public Rhetoric through Participatory Design: Critical Engagements and Creative Expression in the Neighborhood Networks Project”, *Design Issues* 28(3): 48–61.

European Citizen Science Association (ESCA) (2015). “Ten Principles of Citizen Science”, at http://ecsa.citizen-science.net/sites/ecsa.citizen-science.net/files/ECSA_Ten_principles_of_citizen_science.pdf [accessed 8 January 2016].

Gabrys, J. (2016). *Program Earth: Environmental Sensing Technology and the Making of a Computational Planet,* Minneapolis: University of Minnesota Press.

Goodchild, M. F. (2007). “Citizens as Sensors: The World of Volunteered Geography”, *GeoJournal*, 69: 211-221.

Irwin, A. (1995). *Citizen Science: A Study of People, Expertise and Sustainable Development,* London: Routledge.

Jasanoff, S. (2003). “Technologies of Humility: Citizen Participation in Governing Science”, *Minerva*, 41: 223-244.

Maisonneuve N., M. Stevens, M. E. Niessen, P. Hanappe, and L. Steels (2009). “Citizen Noise Pollution Monitoring”, *The Proceedings of the 10th International Digital Government Research Conference,* 17-21 May 2009, Puebla, Mexico, pp. 96-103.

MarcellusGas.Org, at https://www.marcellusgas.org [accessed 7 January 2016].

Moore, C. W. (2013). “Air Impacts of Gas Shale Extraction and Distribution”, *Workshop on Risks of Unconventional Shale Gas Development*, 30-31 May, at http://sites.nationalacademies.org/DBASSE/BECS/DBASSE_083187 [accessed 7 January 2016].

Paulos P., R.J. Honicky, and B. Hooker (2009). “Citizen Science: Enabling Participatory Urbanism”, in Foth, M. (Ed.) *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*. Hershey, PA: Information Science Reference, IGI Global, pp. 414-436.

State Impact. “The Marcellus Shale, Explained.”

US Environmental Protection Agency (US EPA) (2013). “Draft Roadmap for Next Generation Air Monitoring”, draft report, 8 March, at http://www.eunetair.it/cost/newsroom/03-US-EPA_Roadmap_NGAM-March2013.pdf [accessed 7 January 2016].

US Environmental Protection Agency (US EPA) (2015). “Citizen Science and Crowdsourcing: Creative Approaches to Environmental Protection”, draft report, September, at http://www.epa.gov/sites/production/files/2015-09/documents/nacept_background_material_2.pdf [accessed 8 January 2016].