JCOM

BIG DATA AND DIGITAL METHODS IN SCIENCE COMMUNICATION RESEARCH: OPPORTUNITIES, CHALLENGES AND LIMITS

Making citizen science newsworthy in the era of big data

Stuart Allan and Joanna Redden

Abstract

This article examines certain guiding tenets of science journalism in the era of big data by focusing on its engagement with citizen science. Having placed citizen science in historical context, it highlights early interventions intended to help establish the basis for an alternative epistemological ethos recognising the scientist as citizen and the citizen as scientist. Next, the article assesses further implications for science journalism by examining the challenges posed by big data in the realm of citizen science. Pertinent issues include potential risks associated with data quality, access dynamics, the difficulty investigating algorithms, and concerns about certain constraints impacting on transparency and accountability.

Keywords

Citizen science; History of public communication of science; Science and media

Introduction

Just as one might expect, it is possible to trace the emergence of numerous examples of 'citizen science' long before the actual concept itself enters science journalism's lexicon. Let us briefly introduce one such example to serve as our point of departure here, namely because it also helps us to begin illuminating several pertinent issues for current debates regarding big data gathering and interpretation.

At the turn of the twentieth century, a fledgling conservation movement in North America inspired a concerted effort amongst bird watchers to provide an alternative to the traditional hunting event called the Christmas Side Hunt of wild birds and animals. Dismayed by the scale of the carnage, ornithologist Frank M. Chapman, editor of *Bird-Lore* magazine, announced on its pages the first 'Christmas Bird-Census' to be launched on Christmas Day, 1900:

> It is not many years ago that sportsmen were accustomed to meet on Christmas Day, 'choose sides' and then, as representatives of the two bands resulting, hie them to the fields and woods on the cheerful mission of killing practically everything in fur or feathers that crossed their path — if they could.
>
> These exceptional opportunities for winning the laurels of the chase were termed 'side hunts,' and reports of the hundreds of non-game birds which

> were sometimes slaughtered during a single hunt were often published in our leading sportsmen's journals, with perhaps a word of editorial commendation for the winning side. [. . .]
>
> Now, BIRD-LORE proposes a new kind of Christmas side hunt, in the form of a Christmas bird-census, and we hope that all our readers who have the opportunity will aid us in making it a success by spending a portion of Christmas Day with the birds and sending a report of their 'hunt' to BIRD-LORE before they retire that night. [. . .] [Chapman, 1900].

Chapman worked with fellow volunteers to count the numbers of birds, and to collect details about species, across 25 sites (each covering a 15-mile circle) in the Northeast of the United States. In the years to follow, the annual bird census quickly spread throughout the country, as well as into Canada, parts of Central and South America, the Caribbean and several Pacific islands. *Bird-Lore*, the 'official organ of the Audubon Societies' at the outset, was re-titled *Audubon Magazine* in 1941, before becoming simply *Audubon* in 1966. The non-profit National Audubon Society behind it, initially incorporated in 1905, has consistently ensured the necessary funding to underwrite the cost of publishing the census results in an annual report. In the most recent Christmas Bird Count (2015–16) tally, almost 59 million birds (2,607 species) were counted by 76,669 observers across 2,505 count circles, each being held on one day over the holiday (mid-December to early January) period.[1]

Today the Audubon Society's Christmas Bird Count is widely considered to be the most celebrated wildlife census in the world, one increasingly described in news reports as a remarkably successful example of 'citizen science' marshalling big data. Count data about population fluctuations, movements, and the status of species have consistently proven to be invaluable, not least for helping to set research and conservation priorities — particularly where birds designated as threatened are concerned. 'For over one hundred years, the desire to both make a difference and to experience the beauty of nature has driven dedicated people to leave the comfort of a warm house during the Holiday season,' the Michigan Audubon [2016] website dedicated to the bird count explains. 'Each of the citizen scientists who annually braves snow, wind, or rain, to take part in the Christmas Bird Count makes an enormous contribution to conservation.' This description of the volunteers as 'citizen scientists' is revealing, implicitly underscoring as it does the convergence of their personal interests and expertise with those of Audubon's scientific advisors. Not surprisingly, closer examination of how this grassroots network of concerned individuals operates reveals a multiplicity of motivations inspiring their involvement. Equally telling, however, is the extent opportunities for this type of collaborative engagement have been facilitated by further points of convergence between citizen science and science journalism. For many of the volunteer birders participating in the annual count, there is a news story at stake here that recurrently eludes the attention of the mainstream media — namely, in

[1] These figures are drawn from LeBaron's [2016] summary on the Audubon website. Errors can creep into the data, of course, although steps are taken to lessen their impact. 'The counting parties are organized to minimize the chances of misidentification and limit the possibility that any bird will be counted more than once,' Jane E. Brody [1989] explains. 'Unusual sightings or inordinately large numbers of birds are checked by regional compilers before the data are sent to the National Audubon Society,' she continues. 'When scientists use the Christmas count data to calculate the size of bird populations, they try to standardize the figures by taking into account the number of observers in each party and how much time was spent in the field.'

their view, the evidence gathered regarding the birds' activities reveals important insights into ecosystems in a serious state of crisis.

This article aims to explore several pressing issues for efforts to rethink science journalism in the era of big data by focusing on its engagement with citizen science. 'The organized pursuit of science by amateurs is a little known but spectacularly successful phenomenon,' a recent *Los Angeles Times* news account explains. 'Over time it has brought together people of all ages, creeds, backgrounds and political persuasions and enabled discoveries that were only made possible through their collective volunteer work' [16 March 2017; see also Lewenstein, 2016; Weitkamp, 2016]. In the next section, we briefly highlight various facets of this engagement, including one of the early interventions — Lancelot Hogben's [1938] *Science for the Citizen* — intended to help establish the basis for an alternative epistemological ethos recognising the scientist as citizen and the citizen as scientist. From there, the article proceeds to assess further implications for science journalism by examining the challenges posed by big data, particularly in the realm of citizen science. Pertinent issues include potential risks associated with data quality, access issues, the difficulty investigating algorithms, and concerns about certain constraints impacting on transparency and accountability. Accordingly, for science journalists striving to interrogate the uses of big data, we argue, it is crucial for them to better understand how best to frame new lines of investigation which avoid reaffirming as natural, legitimate or inevitable the stark polarities between 'professional' and 'amateur' so familiar to us today.

Citizen science and the 'democratization of positive knowledge'

Readers of major news sites will regularly encounter news items revolving around citizen science and data. Recent examples include:

> 'Astronomers need help to find planet' is the headline of a *Sydney Morning Herald* news account informing readers that astronomers 'want you to trawl through thousands of images taken by NASA's Wide-field Infrared Survey Explorer to see if you can spot Planet Nine.' The 'Backyard Worlds' project 'needs human input because algorithms cannot reliably search some star-rich regions of the sky' [17 February 2017].
>
> 'Telling Mosquitoes Apart With a Cellphone,' a *New York Times* news account, describes efforts to develop a crowdsourcing initiative to produce a 'worldwide mosquito distribution map' using cellphones to 'record mosquito wing beats accurately enough to distinguish, for example, Culex mosquitoes, which spread West Nile virus, from Aedes mosquitoes, which spread Zika' [21 November 2016].
>
> 'Mapping shoreline changes with simple scientific tools' appeared in the Indian broadsheet newspaper *Daily News & Analysis*, detailing how 180 community volunteers have been 'trained to observe, record, and document data on shoreline changes in chosen stretches of beaches.' The citizen science project involves beach profiling, together with sand grain size analysis, to plot monthly readings 'on a simple graph to map the sand erosion or sand accumulation' [19 February 2017].
>
> 'Record low number of British butterflies a "shock and a mystery",' is a *Guardian* news account describing the 2016 Big Butterfly Count, evidently 'the biggest annual citizen science insect survey in the world' involving more than 36,000 people who spotted 390,000 butterflies during the three-week recording period. 'The count is the first indication that the summer of 2016 may be the worst year on record for butterflies' [10 October 2016].

Citizen science initiatives prove newsworthy, these examples suggest, when they can be reported in a way that reaffirms the news values, priorities and protocols consistent with science journalism more generally [see also Allan, 2011; Allan and Ewart, 2015; Cooper, 2016; Dickinson and Bonney, 2012]. At the same time, similarly apparent in these examples is the extent to which familiar boundaries demarcating what counts as scientific inquiry are being redrawn, not least by rendering typically tacit, normative judgements regarding who is qualified to participate in related activities newsworthy in their own right. 'While the ivory towers of academia may have closed in around the concepts of professionalization and expertise over the last century and a half,' Geoffrey Belknap [2016] writes in *The Guardian*, 'the potential of the new digital communities allows us to start questioning what it means to participate in knowledge production.'

Scholarly inquiries into the boundary-work of scientists have shown how certain idealised visions have served as a means to protect professional autonomy over the years, including by normalising their proclaimed authority, reinforcing claims to expertise, even excluding 'rivals by defining them as outsiders,' not least with labels such as 'amateur' [Gieryn, 1983, p. 792; see also Lewis, 2012]. In so doing, the fluid, unevenly evolving boundaries of scientific communities have been defined contingently, in part, at least as much by rules of exclusion as those of inclusion. Early public science campaigns, typically espousing ecologies of 'vernacular' epistemology, tapped into a myriad of educational strategies intended to expand and enrich what was an elitist, exclusive scientific sphere in the interests of the 'common' or 'ordinary' person turned citizen of science [Irwin, 1995; Pandora, 2016; Turner, 1980].

First published in 1938, Lancelot Hogben's weighty tome *Science for the Citizen* is often credited with helping to inspire the general science movement calling for scientific education to be much more inclusive. Hogben, a British experimental zoologist and medical statistician, intended his book to be an intervention aimed at engaging the non-scientist in scientific enquiry. 'Natural science is an essential part of the education of a citizen,' he maintained, 'because scientific discoveries affect the lives of everyone' [1938, p. 9]. The history of science, he wrote, 'is co-extensive with that of civilized living.' He continued:

> It emerges so soon as the secret lore of the craftsman overflows the dam of oral tradition, demanding a permanent record of its own. It expands as the record becomes accessible to a widening personnel, gathering into itself and coordinating the fruits of new crafts. It languishes when the social incentive to new productive accomplishment is lacking, and when its custodians lose the will to share it with others. Its history, which is the history of the constructive achievements of mankind, is also the history of the democratization of positive knowledge [1938, p. 17].

Hogben was convinced the story of science amounted to more than a record of human achievement, it was 'a story of the satisfaction of the common needs of mankind, disclosing as it unfolds new horizons of human wellbeing which lie before us, if we plan our new resources intelligently' [1938, p. 17]. In this way, then, science was effectively politicised. Just as scientists were obliged to exercise their social responsibilities as citizens, thoughtful citizens 'realize that no society is safe in the hands of a few clever people' [1938, p. 1075]. The future of scientific enquiry,

Hogben argued, was dependent on co-operation between scientists and their fellow citizens in order to align the application of scientific knowledge with these common needs. Modern science, it followed, offered 'a new social contract,' one demanding 'a new orientation of educational values and new qualifications for civic responsibility' [1938, p. 1089].

Appearing at a time when the prospect of another European war loomed, *Science for the Citizen*'s declaration of scientific humanism took on further resonance. Borrowing the phraseology of the poet Lucretius, Hogben reminded the reader how 'science liberates us from the terror of the Gods' before extending this critique into a challenge against the capitalist system. 'Scientific planning gives us the means of planning for plenty,' he wrote, 'and also helps to free us from habits which prevent us from doing so' [1938, p. 1079]. While conceding the dividing line between progress in science and morals was not clear cut, he underscored the folly of reducing the social value of knowledge to perceptions of its material benefits or rewards. In essence, then, the social contract of scientific humanism entailed:

> . . . the recognition that the sufficient basis for rational co-operation between citizens is scientific investigation of the common needs of mankind, a scientific inventory of resources available for satisfying them, and a realistic survey of how modern social institutions contribute to or militate against the use of such resources for the satisfaction of fundamental human needs [1938, p. 1089].

Where Britain was concerned, Hogben contended such analyses of human needs would not be emerging from its universities, 'where the teaching on current social problems is dominated by the dreary futilities of deductive economics' [1938, p. 1090]. Rather, it was much more likely the men and women 'who bring the live curiosity and painstaking industry of the naturalist to bear on problems of contemporary society' will be 'the symbol of a popular movement' [1938, p. 1090]. He passionately believed the makers of the 'New Social Contract' would be, necessarily, the founders of a new social culture.

With the benefit of hindsight, *Science for the Citizen* can be read as signalling a formative moment in a history of ideas, one that would gradually coalesce into the wider impetus for the citizen science movement as we recognise it today. Definitions of 'citizen science' vary, of course, depending on who is doing the defining in question. Many scientists devote considerable time striving to create effective ways to engage ordinary members of the public (that is to say, 'laypersons' or 'non-scientists') in science, particularly where scientific uncertainty has become contentious. Such efforts have acquired even greater impetus with the advent of digital technologies in recent years — ranging from the personal computer to the smartphone or tablet of mobile participatory cultures — leading some to herald a new age of 'citizen science' dawning on the horizon.

Fostering synergies

Looking across a range of differing perspectives on what counts as citizen science in this regard, it soon becomes apparent that varied definitions tend to privilege diverse rationales for public participation. As the examples above selected from news coverage of citizen science make clear, however, scientists are embracing digital technologies to fashion new, collaborative forms of connectivity with

dispersed networks of practitioners with considerable success. Members of the scientific community typically welcome the enthusiasm of dedicated 'amateurs', even when some tend to be sceptical about whether the results being produced satisfy research-grade standards. 'Naysayers might chide, "The data are of poor quality, they cannot be trusted; they could be misleading or even dangerous; and they are certainly not admissible in court"', Schnoor [2007] points out. Certain scientific tasks are better than others for citizen science, he concedes, and findings always need to be interpreted with due attention to how they were achieved. Still, he adds, there 'is a sizable literature which attests that data collected by properly trained citizen volunteers are of as high a quality as those obtained by professionals with the same equipment' [Schnoor, 2007, p. 5923; see also Catlin-Groves, 2012].

The growing significance of online citizen science projects — striving to make the most of digital, web-based resources to handle big datasets — underscores how the boundaries of professional science are being redrawn. Lending shape to the ethos of these 'new wave' projects is their commitment to moving beyond more traditional, deficit-model (top-down, zero-sum) conceptions of the 'public understanding of science' in order to emphasise meaningful engagement in co-operative ventures. Some of these new projects serve not only social and scientific functions by contributing to discovery and information collection, but also simultaneously serve science and data literacy functions by illustrating the politics of data. In an age of big data, as pre-existing power dynamics are extended and new power dynamics are forged, efforts that increase awareness of the basic premise that data are not neutral become essential [Kitchin, 2014; Langlois, Redden and Elmer, 2015]. Often for the citizen scientists involved, Lewenstein [2016] observes, they 'are learning science at the same time they are challenging scientific orthodoxies and making claims on the governance of science' [Lewenstein, 2016, p. 2].

A recent illustrative example concerns the range of individuals and organizations who have come together to save climate data and act as watchdogs in response to fears about the threat the US President Donald Trump's administration poses to environmental and energy policy and infrastructure [EDGI, 2017]. The fear is that publicly accessible climate data will be erased. In response to this risk, data are being scraped and saved from online public sources through hackathons held in Canada and the United States and by a global network of volunteers organized by groups like Climate Mirror [Beeler, 2017]. Data archiving organizations like the Environmental Data and Governance Initiative and Data Refuge have been set up to collect and store data and to also perform a watchdog function by monitoring changes to data, regulation, enforcement, research, funding, websites, and agency management. Further, there has been a long-standing and growing movement among environmental justice activists, acting as citizen scientists, to address the lack of data in some areas. Citizen scientists globally are involved in environmental monitoring through sensors, making use of crowdsourcing to collect data, and social media data mining [Mah, 2015; Gabrys, 2016].

There is much to be gained by fostering productive synergies between citizen science and big data. Another less overtly political example of this is eBird, a global citizen science project to collect data about bird observations, which in some ways resembles a big data version of the Audubon Society's Christmas Bird Count, [Sullivan et al., 2009]. The project makes use of new digital technologies while also

responding directly to concerns about volunteer data by improving data quality through big data applications. As detailed by Kelling et al. [2015] the data submission process is designed to ensure data meets high quality standards and is complete and accurate. Further a 'sensor calibration' approach is employed to measure the variations in the ways citizens monitor birds, and species distribution models are used to both fill in data gaps and control for bias that may enter through data collection [Kelling et al., 2015, p. 602]. New big data processes make it easier to collect, combine, and analyse data generated by citizen volunteers. While the eBird project demonstrates the benefits that can be achieved when citizen science is coupled with big data tools, some of the epistemic and ethical issues raised by other uses of big data are hinted at in this project. One of the ways big data are used to evaluate data quality with this project is the use of various measures to rank the skill of those who participate [Kelling et al., 2015]. Ranking in this context is benign but becomes a more complex issue when considering the data generated and used by citizen scientists in areas like health and neuroscience.

Citizen scientists contributing data about themselves, and making use of open data in a wide range of areas, present promise and risk. Citizens are sharing their health and behavioural data through wearable technologies, the Internet of Things, online participation, and mobile tracking. Concerns are being raised about how data can be repurposed, potentials for re-identification, citizen access to their data profiles, who is using the data contributed and how, how data may be shared, and how it may be used in the future [Lewis and Westlund, 2015]. As an example, neuroscience researchers developing brain training programs are monitoring and collecting data through citizens who are willing to participate in them. The goal is to use the data collected to improve programs trying to enhance cognitive function. This kind of participation may lead to the contribution of highly personal details about behavioural and cognitive functions. Purcell and Rommelfanger [2015] argue the datasets collected about people through their use of online brain training programs will be worth more than the training programs being developed. These growing datasets contain information about people all over the world, and hold valuable cognitive performance data that certain advertisers and others will want to access [Purcell and Rommelfanger, 2015, p. 358]. What rights do people have with respect to the performance data held about them, how it is used, who has access to it, and who may delve into it in the future? Of course, these kinds of questions about data collection and use can be applied not just to citizen science participation, but also to the kinds of data being collected and combined more generally by advertisers, data brokers, and governments through citizens' online and mobile communications, transactions, and movements. These kinds of questions apply to journalists investigating big data uses and applications across public, corporate and government sectors. The challenge is that investigating big data practices is neither easy or straightforward.

A new social contract?

In thinking about how Hogben's [1938] notion of a 'new social contract' and its normative commitment to the 'democratization of positive knowledge' resonate in an era of big data, the importance of recasting longstanding binary oppositions between 'professional' and 'amateur' becomes a pressing concern. As we have seen, traditional assumptions regarding the ordinary citizen's relationship to the authorised, legitimate boundaries of the scientific community have lost much of their ideological purchase. Citizen science initiatives are opening-up alternative

epistemologies, particularly where knowledge claims aligned with big data are being generated and sustained across dispersed networks calling into question earlier (once again, typically top-down, zero-sum) conceptions of scientists' relationship with their publics.

In our discussion above, we have briefly surveyed a range of examples showing why it is so important for science journalism to interrogate uses of datasets in a manner alert to both new risks as well as opportunities. As governments increase the amount of data available to the public through open data programs citizen scientists are using this data to make scientific discoveries, develop new programs and tools, improve healthcare, lobby for policy changes, navigate systems, and enhance public education [Hoffman, 2015]. However, as health records become available there are a range of risks, including risks to privacy and anonymity given what we are learning about the ease of re-identification. There are also risks of discrimination, social sorting and targeting as discoveries may be re-purposed by others in ways not initially intended. There are risks of error due to data problems, assumptions based on false correlations, or poor project design [Hoffman, 2015, p. 1784]. As argued by Hoffman [2015], another risk in our internet age is how misinformation can be amplified through the sharing of content online. The efforts to discredit climate science and the false news widely shared during the last American election present two effective illustrations of this type of risk. The black boxed nature of big data processes, the dominant myths about big data as objective and neutral, as well as the inability of most to understand these processes would potentially make interrogating false conclusions difficult.

Ensuring adequate data quality is a significant challenge for those making use of big data. There can be 'signal problems' or dark areas in the data because particular groups, people, or communities are not represented, or conversely some demographic groups may be over-represented in the data [Barocas, 2014; Crawford, 2013]. Further, there can be errors due to data inputs. In the United States, government officials say problems with data regularly prevent them from working effectively [Barrett and Greene, 2015]. British and American companies who do social media data mining note they regularly have issues with quality and accuracy of data used [Kennedy, 2016]. For citizen scientists making use of government open data, assessing questions of data quality and adjusting for data problems are serious issues. Mah [2015, p. 7] contends the uncertainties introduced through uses of big data pose particular difficulties for citizen scientists who are working in contested areas and trying to raise awareness and change practices and policies.

A further range of challenges are connected to issues of access, transparency and accountability. The algorithms that facilitate the analysis of big data are often highly complex and can be inscrutable even to those who produce them. This is particularly the case with machine learning algorithms. As has been argued, algorithms are not neutral but the product of socio-technical systems and influenced by a range of historical, situated and contextual factors [Gitelman, 2014; Amoore and Piotukh, 2015]. They embed the assumptions of their architects [Gourarie, 2016]. Further uses of big data involve translation of complex and messy problems into algorithms, they also involve iterative processes of development, many people, as well as infrastructural considerations and limitations [Kitchin, 2017]. The point, Kitchin, Lauriault and McArdle [2015] argue, is that:

> Data do not exist independently of the ideas, instruments, practices, contexts, knowledges and systems used to generate, process and analyse them, regardless of them often being presented in this manner. Data are generated as the product of many minds working within diverse situations, framed and shaped within contexts and structures [Kitchin, Lauriault and McArdle, 2015, p. 16].

The 'black boxed' nature of many big data processes complicates the efforts of investigators, few of whom will likely have the access or knowledge to interrogate these systems [Pasquale, 2015; Diakopoulos, 2014]. Access can be an issue because often the algorithms in question can be proprietary [Angwin et al., 2016], developed in spaces and by people few have access to [Redden, 2015], or because they are part of automated machine learning systems constantly evolving and that even the original architects may not be able to understand. To cite Ruppert [2016], big data systems 'actualize and legitimize' worlds, ways of being and ways of knowing. These systems hold great potential, but for citizens and journalists alike also present real obstacles hindering participation, understanding and investigation.

To close, it is our contention that Hogben's [1938] aim of democratizing science needs to be extended to the realm of big data as a matter of urgency. Key questions going forward, it follows, include: how should science journalism evolve in order to further enhance citizens' abilities to interrogate big data practices while, at the same time, encouraging new, nuanced forms of engagement to increase wider public participation? Moreover, we ask, what sort of journalistic strategies would help to deconstruct familiar mythologies of big data, and thereby foster wider, critical debate about possible risks and opportunities? A first step in formulating answers, this article's critique suggests, is for science journalists to redouble their efforts to create spaces for active dialogue and debate about science's civic responsibilities.

References

Allan, S. (2011). 'Introduction: Science journalism in a digital age'. *Journalism* 12 (7), pp. 771–777. DOI: 10.1177/1464884911412688.

Allan, S. and Ewart, J. (2015). 'Citizen science, citizen journalism: New forms of environmental reporting'. In: The Routledge Handbook of Environment and Communication. Ed. by A. Hansen and R. Cox. London and New York: Routledge, pp. 186–196.

Amoore, L. and Piotukh, V. (2015). 'Life beyond big data: governing with little analytics'. *Economy and Society* 44 (3), pp. 341–366. DOI: 10.1080/03085147.2015.1043793.

Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016). *ProPublica*. URL: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (visited on 2nd September 2016).

Barocas, S. (2014). 'Data mining and the discourse on discrimination'. In: *Data Ethics Workshop, Conference on Knowledge Discovery and Data Mining*. URL: https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf.

Barrett, K. and Greene, R. (2015). 'How big data in government falls dangerously short of the mark'. *Government Technology*. URL: http://www.govtech.com/opinion/How-Big-Data-in-Government-Falls-Dangerously-Short-of-the-Mark.html.

Beeler, C. (2017). 'People around the world are helping the US save its climate data'. *PRI's The World*. URL: https://www.pri.org/stories/2017-01-26/people-around-world-are-helping-us-save-its-climate-data.
Belknap, G. (2016). 'People power: how citizen science could change historical research'. *The Guardian*.
URL: https://www.theguardian.com/science/the-h-word/2016/apr/26/how-citizen-science-could-change-historical-research-crowdsourcing.
Brody, J. E. (1989). 'Annual Bird Counts Are Starting to Yield Scientific Insights'. *The New York Times*. URL: http://www.nytimes.com/1989/12/19/science/annual-bird-counts-are-starting-to-yield-scientific-insights.html.
Catlin-Groves, C. L. (2012). 'The Citizen Science Landscape: From Volunteers to Citizen Sensors and Beyond'. *International Journal of Zoology* 2012, pp. 1–14.
DOI: 10.1155/2012/349630.
Chapman, G. (1900). 'A Christmas bird-census'. *Bird-Lore* 2 (6), p. 192.
Cooper, C. (2016). Citizen science. New York, U.S.A.: Overlook.
Crawford, K. (2013). 'The hidden biases in big data'. *Harvard Business Review*.
URL: https://hbr.org/2013/04/the-hidden-biases-in-big-data.
Diakopoulos, N. (2014). 'Algorithmic accountability: On the investigation of black Boxes'. *Tow Center for Digital Journalism*. URL: http://towcenter.org/research/algorithmic-accountability-on-the-investigation-of-black-boxes-2/.
Dickinson, J. L. and Bonney, R., eds. (2012). Citizen Science: Public Participation in Environmental Research. London, U.K.: Comstock.
URL: http://www.cornellpress.cornell.edu/book/?GCOI=80140100107290.
Environmental Data & Governance Initiative (2017). *Overview'*.
URL: https://envirodatagov.org/.
Gabrys, J. (2016). Program earth: Environmental sensing technology and the making of a computational planet. Minneapolis, U.S.A.: University of Minnesota Press.
Gieryn, T. F. (1983). 'Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists'. *American Sociological Review* 48 (6), pp. 781–795. DOI: 10.2307/2095325.
Gitelman, L. (2014). Raw Data is an Oxymoron. Cambridge, MA, U.S.A.: MIT Press.
Gourarie, C. (2016). 'Investigating the algorithms that govern our Lives'. *Columbia Journalism Review*.
URL: http://www.cjr.org/innovations/investigating_algorithms.php.
Hoffman, S. (2015). 'Citizen Science: The Law and Ethics of Public Access to Medical Big Data'. *Berkeley Technology Law Journal* 30 (2), pp. 1741–1805.
DOI: 10.15779/Z385Z78.
Hogben, L. (1938). Science for the citizen. London, U.K.: George Allen & Unwin.
Irwin, A. (1995). Citizen Science: a Study of People, Expertise and Sustainable Development. Oxon, U.K.: Routledge.
Kelling, S., Fink, D., La Sorte, F. A., Johnston, A., Bruns, N. E. and Hochachka, W. M. (2015). 'Taking a 'Big Data' approach to data quality in a citizen science project'. *Ambio* 44 (S4), pp. 601–611. DOI: 10.1007/s13280-015-0710-4.
Kennedy, H. (2016). Post, mine, repeat: social media data mining becomes ordinary. Basingstoke, U.K.: Palgrave Macmillan.
Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. London, U.K.: Sage.
— (2017). 'Thinking critically about and researching algorithms'. *Information, Communication & Society* 20 (1), pp. 14–29.
DOI: 10.1080/1369118X.2016.1154087.

Kitchin, R., Lauriault, T. P. and McArdle, G. (2015). 'Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards'. *Regional Studies, Regional Science* 2 (1), pp. 6–28. DOI: 10.1080/21681376.2014.983149.

Langlois, G., Redden, J. and Elmer, G. (2015). Compromised data: from social media to big data. New York, U.K.: Bloomsbury.

LeBaron, G. S. (2016). 'The 116th Christmas bird count summary'. *Audubon*. URL: http://www.audubon.org/news/the-116th-christmas-bird-count-summary.

Lewenstein, B. V. (2016). 'Can we understand citizen science?' *JCOM* 15 (1), E, pp. 1–5.

Lewis, S. C. (2012). 'The tension between professional control and open participation'. *Information, Communication & Society* 15 (6), pp. 836–866. DOI: 10.1080/1369118X.2012.674150.

Lewis, S. C. and Westlund, O. (2015). 'Big Data and Journalism'. *Digital Journalism* 3 (3), pp. 447–466. DOI: 10.1080/21670811.2014.976418.

Mah, A. (2015). 'Environmental justice in the age of big data: challenging toxic blind spots of voice, speed, and expertise'. *Environmental Sociology* 3 (2), pp. 122–133. DOI: 10.1080/23251042.2016.1220849.

Michigan Audubon (2016). *Christmas bird count*. URL: http://www.michiganaudubon.org/get-involved/citizen-science/.

Pandora, K. (2016). 'Amateurs'. In: *A Companion to the History of Science*. Ed. by B. Lightman. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 139–152. DOI: 10.1002/9781118620762.ch10.

Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Boston, U.S.A.: Harvard University Press.

Purcell, R. H. and Rommelfanger, K. S. (2015). 'Internet-Based Brain Training Games, Citizen Scientists, and Big Data: Ethical Issues in Unprecedented Virtual Territories'. *Neuron* 86 (2), pp. 356–359. DOI: 10.1016/j.neuron.2015.03.044.

Redden, J. (2015). 'Big data as system of knowledge: Investigating Canadian governance'. In: Compromised data: from social media to big data. Ed. by G. Elmer, G. Langlois and J. Redden. New York, U.S.A.: Bloomsbury.

Ruppert, E. (2016). 'Big data economies and ecologies'. In: An end to the crisis of empirical sociology? Trends and challenges in social science research. Ed. by L. Ryan and L. McKie. London, U.K.: Routledge, pp. 12–26.

Schnoor, J. L. (2007). 'Citizen science'. *Environmental Science & Technology* 41 (17), pp. 5923–5923. DOI: 10.1021/es072599+.

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D. and Kelling, S. (2009). 'eBird: A citizen-based bird observation network in the biological sciences'. *Biological Conservation* 142 (10), pp. 2282–2292. DOI: 10.1016/j.biocon.2009.05.006.

Turner, F. M. (1980). 'Public Science in Britain, 1880–1919'. *Isis* 71 (4), pp. 589–608.

Weitkamp, E. (2016). 'From planning to motivations: citizen science comes to life'. *JCOM* 15 (03), E, pp. 1–5.

Authors

Stuart Allan is Professor and Head of the School of Journalism, Media and Cultural Studies at Cardiff University, UK. E-mail: AllanS@cardiff.ac.uk.

Joanna Redden is Lecturer in the the School of Journalism, Media and Cultural Studies at Cardiff University, UK. E-mail: ReddenJ@cardiff.ac.uk.

How to cite

Allan, S. and Redden, J. (2017). 'Making citizen science newsworthy in the era of big data'. *JCOM* 16 (02), C05.

ISSN 1824-2049. Published by SISSA Medialab. jcom.sissa.it