CONFERENCE OF EUROPEAN STATISTICIANS

Workshop on Statistical Data Collection

10-12 October 2017, Ottawa, Canada

WP.3-6

16 August 2017

A citizen data app: reflections on a collaborative practice

Evelyn Ruppert and Francisca Grommé (Goldsmiths, University of London)

e.ruppert@gold.ac.uk

Abstract

The paper discusses an experiment in collaboration between national statisticians and university researchers to establish design principles for a 'citizen data app' for generating official statistics. The researchers are part of an ERC funded project, ARITHMUS (Peopling Europe: How data make a people), which is concerned with how new digital technologies and new data sources are remaking and challenging official statistics. The term 'citizen data' is used to refer to data co-produced with citizens and which engages them in all stages of production, from the design of an app or platform to the interpretation and analysis of data. We attend to various issues that co-production raises such as data quality, reliability and security. At the same time, we address how co-production can potentially mitigate problems associated with the re-use of Big Data and issues of trust in relation to the collection of government data. Through a focus on four key design principles - experimentalism, citizen data, privacy by design, and smart systems – we argue that the future of official statistics not only depends on inventing new data sources and methods but also new relations to citizens.

A citizen data app: reflections on a collaborative practice

Evelyn Ruppert and Francisca Grommé, Goldsmiths, University of London
e.ruppert@gold.ac.uk
8 July 2017

For the UNECE Workshop on Statistical Data Collection
10–12 October 2017, Ottawa, CA

Abstract
The paper discusses an experiment in collaboration between national statisticians and university researchers to establish design principles for a 'citizen data app' for generating official statistics. The researchers are part of an ERC funded project, ARITHMUS (Peopling Europe: How data make a people), which is concerned with how new digital technologies and new data sources are remaking and challenging official statistics. The term 'citizen data' is used to refer to data co-produced with citizens and which engages them in all stages of production, from the design of an app or platform to the interpretation and analysis of data. We attend to various issues that co-production raises such as data quality, reliability and security. At the same time, we address how co-production can potentially mitigate problems associated with the re-use of Big Data and issues of trust in relation to the collection of government data. Through a focus on four key design principles - experimentalism, citizen data, privacy by design, and smart systems – we argue that the future of official statistics not only depends on inventing new data sources and methods but also new relations to citizens.

Introduction

NSI experiments on the potential of Big Data as a new source of data for the generation of official statistics have now been underway for many years. The ARITHMUS project has documented several issues identified by these experiments such data access, data ownership, privacy and ethics, data representativeness, data quality and so on.[1] Much attention has been given to technical, legal and organisational solutions to these issues or rejecting the use of Big Data as a source altogether.

The possible solution that we are investigating concerns how NSIs might generate statistics by co-producing new forms of data with citizens, instead of predominantly relying on data generated by the private sector with all the associated drawbacks. We came to this after two years of studying how NSIs are experimenting with new digital technologies and new data sources in the remaking of official statistics. Through discussions and meetings with the project's Advisory Group (consisting of representatives of NSIs and international statistical organisations), we agreed to engage in an experiment in collaboration to establish design principles for a 'citizen data app' to generate data for governing and research. Whereas we have engaged with national statisticians through conversations, meetings and

[1] ARITHMUS (Peopling Europe: How data make a people) is a sociological study of the making of official statistics. Since 2014, a team of 6 researchers1 has been following the working practices in five EU national statistical institutes (NSIs) and two international organisations (Eurostat and UNECE). The ARITHMUS researchers are: Evelyn Ruppert (Principal Investigator), Baki Cakici, Francisca Grommé, Stephan Scheel, Ville Takala and Funda Ustek-Spilda. This presentation builds on and summarises key points in a recently published ARITHMUS working paper: Grommé, F., Ustek-Spilda, F., Ruppert, E., Cakici, B. 2017. *Citizen Data and Official Statistics: Background Document to a Collaborative Workshop*. ARITHMUS Working Paper Series, Paper No. 2. Available here.

conferences, in this workshop we aim to engage through the practice of developing a method and technology together.

The collaboration involves many stages and one involves the organisation of a workshop in September 2017 to identify possible design elements of a 'citizen data app' or web platform. The workshop participants will include the ARITHMUS researchers; statisticians from national and international organisations; researchers from different universities in the EU; representatives of research organisations engaged in related issues; and information designers and programmers. At this stage, the app will not be materialized and most of its technical specifications will remain undefined. Instead, the outcome of the workshop will be a description of design principles about how the app could be configured in terms of its aims, issue areas and the topic of statistics addressed. The results of the workshop will then feed into the writing of a funding application to develop a Proof of Concept (PoC) application to the European Research Council (ERC).

In my presentation, I will report on the outcomes of that workshop. I will also discuss what is outlined in the remainder of this paper, the key starting concepts that underpin our proposition for a citizen data app that we developed and discussed with the ARITHMUS Advisory Group: experimentalism, citizen science, smart statistics and privacy-by-design. These concepts are based on a set of concerns we have identified in relation to the use of existing Big Data sources for official statistics: the distance between citizens and the practices involved in the generation of data. Unlike long established methods of data collection such as surveys and questionnaires, methods of repurposing data generated by social media, mobile phones and browsers constitute various forms of detachment: between citizens and NSIs; and between citizens' actions, identifications and experiences and how they are categorised, included and excluded. These concerns, we argue, not only introduce problems of interpretation but potentially undermine trust in and the legitimacy of data and in turn official statistics. We recognise that these forms of detachment and consequences are also relevant to academic researchers who repurpose the same data. Our objective then is to explore the causes and possible solutions to detachment through the design of a citizen data app co-produced between national statisticians, academic researchers and citizens.

Experimentalism

In areas as diverse as wheelchair design, Big Data and synthetic biology,[2] social scientists have adopted experimentalism to open up spaces for problem formulations, engage with different actors and consider alternative futures. Academia is not unique in these efforts, as government agencies and corporations have engaged in similar experiments. Official statistics is a good example of a government domain openly embracing experimentalism, as attested by innovation laboratories, sandboxes, hackathons and exploratory research projects.

The adoption of experimentation as a method is most commonly undertaken to achieve social benefits of various sorts. Broadly speaking, we can distinguish two formats through which collaborative experiments seek to achieve this. The first is through various forms of participation intended to achieve a degree of democratisation by opening up scientific and technical debates and processes to publics (Marres, 2012). The second, which is the format we engage with in the workshop, is to experiment collaboratively with stakeholders to develop and explore new problem formulations, transcend ingrained styles of reasoning, disrupt existing hierarchies and critically examine how the data we study are made (Rabinow and Bennett 2012; Ruppert et al. 2015). This is the model of a 'collaboratory' (or, co-laboratory) where participants engage in the common exploration of a topic. There are also

[2] For these three examples, see: https://entornoalasilla.wordpress.com/english/; Ruppert et al. (2015); and http://www.anthropos-lab.net/about.

different models of doing this; one engages participants in the definition and development of shared concepts for understanding a topic or issue. This was the model experimented in the fore-runner of the ARITHMUS project, which involved a collaboratory to 'socialise' Big Data through discussion and provocations (Ruppert et al. 2015).[3] The model of the citizen data workshop follows a different model. Through the design of a 'thing'- a tangible end-product – we seek to *practically* explore and develop shared concepts and issues. By working on a common product (a proposal for an app/platform, in our case) is to make "issues experimentally available to such an extent that 'the possible' becomes tangible, formable, and within reach".[4] Working on a common object thus forces participants to make future modes of working explicit (Muniesa and Linhardt 2011),[5] in this case to consider the future of data collection and analysis in official statistics through the design of an app. Generally, from the social studies of science we learn that experiments are not impartial and objective processes of discovery. Instead, they require a reshaping of relations between participants, objects of knowledge and things. They bring into being new entities, agencies and problematisations (Haraway 1988; Latour 1993) that may pose unexpected risks and problems to their environment, as has been shown in experiments in finance and biotechnology (Millo and Lezaun 2006).

What we suggest, therefore, is a 'care-full' approach to collaborative experimentation (Grommé 2015). Elements of such an approach are to monitor and document who and what are (unavoidably) in – and excluded; avoid ambiguity about terms of evaluation (when do we think something is 'good enough'?); avoid attributing failure solely to perceived local circumstances; avoid separating normative elements from scientific fact (Latour 2006); and to ensure documentation. The elements of a care-full experimentalism include the key starting concepts we have identified such citizen privacy, anonymity and consent, which we treat in a separate section.

Citizen Science

A second concept informing the workshop is that of citizen science, which is generating new relations between citizens, statisticians and users. Different models of citizen science make it possible to think of citizens as not mere research subjects, but as actively involved in the production of data as opposed to traditional methods where they have been otherwise 'passive' subjects. There are many definitions and interpretations of citizen science in the literature and the terms of their engagement in the making of data. Goodchild uses the term to describe communities or networks of citizens who act as observers in some domain of science (2007, 218). This is the most commonly accepted definition especially evident in the significant momentum citizen science has gained in the natural sciences in recent years (Kullenberg and Kasperowski, 2016, 2). However, the practice of engaging people in collecting and submitting data for scientific purposes goes back at least to the 1960s, though the term itself was not used until the 1990s; for some, it includes the National Audubon Society's Annual Christmas Bird Count in early 1900s, where citizens were asked to participate in the observation and enumeration of bird species (Ibid.).

A second version involves citizens not as only observers but co-designers of scientific studies to reflect their own concerns, needs and questions. Other versions include local and activist-oriented approaches such as 'community based auditing', 'civic science', community

[3]Another inspiration for this project is the model of the ethnographic 'para-site', see http://www.socsci.uci.edu/~ethnog/theme3.htm.

[4] Also see Callon et al.'s notion of 'co-researchers' (2011).

[5] A related research approach is practice-based research, in which participants from different disciplines develop a common project together. This would not necessarily result in a material object, but could also be an event or participation in a common activity.

environmental policing', 'street science', 'popular epidemiology', 'crowd science', and 'Do It Yourself Science' (Kullenberg and Kasperowski 2016, 2). These versions range from citizens seeking close alliances with scientific and knowledge institutions to citizens seeking to produce independent knowledge together with scientists.

The reasons citizens engage are multiple, from documenting concerns over environmental issues to creating an online archival map of local historical sites. Goodchild notes two reasons why people might be motivated to participate: self-promotion and personal interest (2007, 219–220). He argues that self-promotion is clearly an important motivator for internet activity, as increasingly citizens spend their lives online and engage in online social networks. Personal interest can also be a good motivator, if citizens believe that the information might be useful for them personally or their communities at some point in future. For instance, Web 2.0 sites have been a convenient way of making information available to users' friends and relations, even though by engaging in Web 2.0, users make the same information available to all (219).

Jasanoff notes that models of citizen science can facilitate meaningful interaction among policy-makers, scientific experts, corporate producers and the public (2003, 235–236). She argues that the pressure for accountability in expert decision-making is manifest in the demand for greater transparency and participation by stakeholders. However, participatory opportunities cannot alone ensure the representative and democratic governance of science. She underscores that the attention of modern states has focused on refining the 'technologies of hubris' which are designed, generally, to facilitate management and control, even in areas of high uncertainty (238). What is lacking is not just knowledge, but ways to bring uncertain, unknown processes and methods into the dynamics of democratic debate (239-240). What Jasanoff points to is that beyond personal and individual reasons, citizens, governments and scientists also engage in citizen science for a more inclusive generation of knowledge.

A contentious issue concerning the role of non-scientists in the production of science is the implications for established scientific principles.[6] However, as Goodchild demonstrates, although citizen science might not fulfil scientific criteria, it can potentially open up new ways of thinking and approaching data. He gives the example of traditional mapping agencies which have elaborate standards and specifications to govern the production of geographic information and who employ cartographers with documented qualifications (2007, 219). The quality of information produced by these cartographers stems from their scientific expertise. Google Maps, on the other hand, has no such reputation and its database is mostly generated by users. Yet, Google Maps is used every day by millions around the world with little or no scepticism about its performance. While Google's authority in mapping might be arising from its success in other areas of computation (and that the mapping platform itself is based on principles of computer science and a previous cartographic software called *Earthviewer*), the wide acceptance of Google's mapping reference systems demonstrates that quality in data can be ascertained in many ways, especially considering new technologies (Ibid, 219-220). Goodchild calls this "the democratisation of GIS" because the platform enables citizens to access and use data but also influence it, through marking newly developed areas or identifying wrong addresses. Providing an easy to use platform that does not require coding knowledge or a scientific understanding of geo-mapping further facilitates users' engagement with the platform.

Traditional methods in social sciences and official statistics have a long tradition of engaging closely with citizens as subjects of study, especially in survey-oriented research and the census. Engaging in citizen science requires a shift to active participation or contribution

[6] Also see Gabrys et al. (2016) and Freitag et al. (2016) for discussions about data quality and credibility.

from citizens as research citizens. This is more closely aligned with a model of that conceives of citizens as co-producers. Given the requirements of scientific principles, together with the call for more inclusive production of science, we suggest that a specific form of citizen science is required for citizen involvement in the production of official statistics. This form could attempt to combine statistical science and citizen science. It could entail the production of data that is more representative and inclusive of citizens' concerns, needs and experiences, as well as their own identifications.

Smart Statistics

Active participation and contribution from citizens for public goods has also been extensively conceptualised and practiced by civil society organisations and foundations. De Waag Society in the Netherlands, Nesta in the UK, and the FabLab-Medialab Prado in Spain are some examples. What these organisations have in common is their aim to use digital technologies for social, bottom-up innovation. In this line of thinking, innovation occurs in an 'innovation ecosystem' consisting of various communities, such as innovation labs, open hardware and software communities and open data communities (Bria 2015).

With regard to the production of data by citizens, some of these initiatives build on the concept of smart cities, understood here as the use of Big Data, urban sensors, Internet of Things and other forms of data collection and data integration to streamline municipal governance and transportation infrastructures, rejuvenate local economies, transform the urban environment to make it more sustainable, liveable, and socially inclusive (see, for instance, Henriquez 2016). While smart cities have been defined in various ways, the concept generally refers to 'how cities are increasingly composed of and monitored by pervasive and ubiquitous computing and, on the other, whose economy and governance is being driven by innovation, creativity and entrepreneurship, enacted by smart people' (Kitchin 2014). In this view, Big Data enables real-time analysis of city life, new modes of urban governance, and provides for envisioning and enacting more efficient, sustainable, competitive, productive, open and transparent cities. When citizens actively participate, monitor and use data in relation to smart cities they are often referred to as 'smart citizens'.

What then is the potential role of citizens in the development of 'smart' official statistics? We understand smart statistics and its conceptualisation are still in the making, and perhaps permanently ambiguous. In its current formulation, smart statistics draws on a genealogy of 'smart systems,' such as that on smart cities just mentioned, but also smart energy, smart meters, smart transport, and so on. As a way of thinking, it builds on the massive proliferation of electronic devices and sensors connected to the internet that generate and communicate large volumes of data.[7] How this data might be embedded in statistical production systems such that statistics could be generated in real time and automatically is what smart statistics could be. In this view, data capturing, analysis and processing are envisioned as embedded in activities that generate and simultaneously analyse Big Data. The adoption of such smart systems would dramatically transform the production system for official statistics. They would not only transform technical infrastructures, but also facilitate rethinking business processes and architectures, laws and regulations, ethics, methodologies, and so on.

How then might pervasive and ubiquitous devices and computing be used to compose and monitor populations, businesses and economic phenomena in new ways? Two models have been suggested: using third party systems that exist for other purposes than statistics but from which statistical information can be extracted or developing entirely new data gathering approaches such as sensors and devices exclusively for generating statistical information.[8] In

[7] Eurostat Big Data Task Force, 'Smart Statistics'. Draft document, Oct 2016.
[8] Ibid.

either case, data could be joined up with that from a variety of other sources generated by statistical offices. The latter approach pertains to our objective: to develop a citizen data app that could incorporate elements of smart systems but additionally address the concepts of experimentalism, citizen science, and as developed in the next section, privacy by design. In sum, we would call this an approach to smart statistics that facilitates and fosters citizenship.

Privacy by Design

Privacy by design can be understood as the embedding of privacy protection at the software design stage of data collection platforms, devices or applications. It entails designing privacy protection with citizens in mind at the outset and the implementation of these designs in a transparent manner. As such, privacy by design is a response to the problem of privacy, consent, and confidentiality through software. It can be used in tandem with other tools, such as privacy impact assessments. By employing privacy by design, software designers tackle privacy issues at the beginning of the process, in contrast to other approaches that aim at solving privacy issues after software development is complete or leave privacy considerations to legal or regulatory frameworks.

Cavoukian, Taylor, and Abrams (2010) define privacy by design through seven foundational principles: proactive not reactive and preventative not reactive; privacy as the default; privacy embedded into design; full functionality that leads to positive sum, not zero-sum outcomes; end-to-end lifecycle protection; visibility and transparency; and respect for user privacy. These principles require designs to be committed to privacy from the beginning and to limit data collection to ways that are respectful of citizen expectations. The principles also require that data collection software address the likelihood that data may exist after the software stops functioning. The authors also emphasise that the lifecycle of the software must be considered when deciding on how to best protect privacy, including making plans for retiring data once the software reaches the end of its lifecycle. Finally, the principles compel organisations dealing with personal data to be transparent in their goals and to remain accountable to citizens.

The collection and processing of personal data present many other challenges for privacy in addition to individual privacy. Nissenbaum (2004) argues that privacy norms need to be tied to specific contexts. She describes three principles that have dominated debates around privacy throughout the 20th century, namely, limiting surveillance of citizens by governments, restricting access to private information, and curtailing intrusions into private places. She suggests a new term, ‘contextual integrity’, to deal with the new challenges introduced by information technologies. Contextual integrity demands that information gathering is kept appropriate to the context and obeys the governing norms of distribution within it. The key insight is that norms of distribution vary across cultures, historical periods, locales, and other factors. Additionally, contextual integrity requires awareness of not only the specific site of data collection but also the relevance of related social institutions (Nissenbaum 2009).

Approaches that aim to protect individual privacy may still lead to undesired outcomes in large-scale data collection efforts. When individually anonymised data are joined to create profiles, individuals who fit the profile could still experience effects of profiling even when they are not identified individually. For example, Graham (2005) discusses how software can be used to assign different categories to different parts of the city based on school performance, house prices, crime rates, etc., which might potentially orchestrate inequalities and discriminate inhabitants, even when they are not personally identified. Similarly, Zwitter (2014) has also identified and problematised the potential discriminatory ‘group effects’ of anonymised data.

The use of Big Data also introduces new privacy challenges. Barocas and Nissenbaum (2014) argue that anonymity and consent are often fundamentally undermined in Big Data applications, and that we need other approaches to protect integrity, such as policies based on moral and political principles that serve specific contextual goals and values. Instead of focusing on anonymity in Big Data applications, they instead emphasise securing informed consent, not only as a choice for subjects to waive consent or not, but a requirement that data collectors justify their actions in relation to norms, standards, and expectations. In sum, privacy is not a single thing but depends on the context of production, accountability for group effects and mechanisms of informed consent.

Recently, scholars working to address the technical challenges of privacy in relation to Big Data have proposed a method of privacy protection by taking advantage of blockchain technology (de Montjoye et al. 2014; Guy Zyskind, Nathan, and Pentland 2015; G. Zyskind, Nathan, and Pentland 2015). Blockchain is a distributed computing method where many devices communicate with one another over a shared network, without requiring a central server to authorise the participation of each member or to keep a list of currently connected members. By applying the blockchain technology to privacy, it becomes possible to encrypt and distribute private data over a large network without requiring a trusted central server.

Blockchain privacy methods are intended to solve underlying privacy challenges using a technical framework during software development. They do not stand on their own as the sole solution to preserving privacy, but rather supplement the legal policy-oriented considerations such as contextual integrity, and software design methodologies such as privacy by design.

Conclusion

The foregoing has summarised a set of propositions and starting concepts for a collaborative workshop to be conducted in September 2017 in which the aim is to specify design principles for an app/platform co-produced with citizens that generates citizen data for official statistics. We propose that a citizen data approach has the potential to produce new statistical variables desired and identified by users and citizens, and increase identification with official statistics which might facilitate trust in official statistics. However, these approaches also challenge the principles of data quality, reliability, privacy, security and anonymity. Moreover, they raise important privacy and ethical questions. In this regard, the issues that the workshop will address are not dissimilar to those that arise in relation to Big Data sources. However, designing a bespoke data generation app/platform has the advantage of potentially measuring, documenting and mitigating data principles and issues. It could moreover incorporate citizens' demographic data, their subjective evaluations and feedback. Furthermore, like Big Data sources, a citizen app/platform does not need to replace existing sources. It can also be auxiliary and supplementary to more traditional and longstanding statistical sources.

Considering the novelty of this approach and its (unknown) challenges, the workshop may or may not lead to applicable outcomes. At the very least, the practical work of conceptualising a citizen data app/platform may contribute to specifying the role(s) of citizens in a smart statistics regime.

References

Barocas, Solon and Helen Nissenbaum. 2014. 'Big Data's End Run Around Anonymity and Consent'. in *Privacy, Big Data, and the Public Good*, edited by J. Lane, V. Stodden, S. Bender, and H. Nissenbaum. Cambridge University Press.

Bria, Francesca. 2015. *Growing a Digital Social Innovation Ecosystem for Europe: DSI Final Report*. European Commission. Retrieved 12 January 2016 (https://waag.org/sites/waag/files/public/media/publicaties/dsireport.pdf).

Callon, Michel, Graham Burchell, Pierre Lascoumes, and Yannick Barthe. 2011. *Acting in an Uncertain World: An Essay on Technical Democracy*. Cambridge, Mass.: MIT Press.

Cavoukian, Ann, Scott Taylor, and Martin E. Abrams. 2010. 'Privacy by Design: Essential for Organizational Accountability and Strong Business Practices'. *Identity in the Information Society* 3(2):405–13.

Freitag, Amy, Ryan Meyer, and Liz Whiteman. 2016. 'Strategies Employed by Citizen Science Programs to Increase the Credibility of Their Data'. *Citizen Science: Theory and Practice* 1(1). Retrieved 9 January 2017 (http://theoryandpractice.citizenscienceassociation.org/article/10.5334/cstp.6/).

Goodchild, Michael F. 2007. 'Citizens as Sensors: The World of Volunteered Geography'. *GeoJournal* 69(4):211–21.

Haraway, Donna J. 1988. 'Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective'. *Feminist Studies* 14(3):575–99.

Henriquez, Laurence. n.d. *Amsterdam Smart Citizens Lab: Towards Community Driven Data Collection*. Amsterdam: De Waag Society and AMS Institute. Retrieved 22 November 2016 (https://www.waag.org/sites/waag/files/public/media/publicaties/amsterdam-smart-citizen-lab-publicatie.pdf).

Jasanoff, Sheila. 2003. 'Technologies of Humility: Citizen Participation in Governing Science'. *Minerva* 41(3):223–44.

Kitchin, Rob. 2014. 'The Real-Time City? Big Data and Smart Urbanism'. *GeoJournal* 79(1):1–14.

Kullenberg, Christopher and Dick Kasperowski. 2016. 'What Is Citizen Science? – A Scientometric Meta-Analysis'. *PLOS ONE* 11(1):e0147152.

Latour, Bruno. 1993. *The Pasteurization of France*. Harvard University Press.

Latour, Bruno. 2006. 'Which Protocol for the New Collective Experiments?' *Boletín CF+S* (32/33). Retrieved 9 November 2016 (http://habitat.aq.upm.es/boletin/n32/ablat.en.html).

Marres, Noortje. 2012. *Material Participation: Technology, the Environment and Everyday Publics*. Basingstoke: Palgrave Macmillan.

Millo, Yuval and Javier Lezaun. 2006. 'Regulatory Experiments: Genetically Modified Crops and Financial Derivatives on Trial'. *Science and Public Policy* 33(3):179–90.

de Montjoye, Yves-Alexandre, Erez Shmueli, Samuel S. Wang, and Alex Sandy Pentland. 2014. 'openPDS: Protecting the Privacy of Metadata through SafeAnswers'. *PLOS ONE* 9(7).

Muniesa, Fabian and Dominique Linhardt. 2011. 'Trials of Explicitness in the Implementation of Public Management Reform'. *Critical Perspectives on Accounting* 22(6):550–66.

Nissenbaum, Helen. 2004. 'Privacy as Contextual Integrity'. *Washington Law Review* 79(1).

Nissenbaum, Helen. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

Rabinow, Paul and Gaymon Bennett. 2012. *Designing Human Practices: An Experiment with Synthetic Biology*. University of Chicago Press.

Ruppert, Evelyn et al. 2015. ‘Socialising Big Data: From Concept to Practice’. Retrieved 3 November 2015 (http://www2.warwick.ac.uk/fac/cross_fac/cim/research/socialising-big-data/sbd_wp_2015.pdf).

Zwitter, Andrej. 2014. ‘Big Data Ethics’. *Big Data & Society*.

Zyskind, G., O. Nathan, and A. Pentland. 2015. ‘Decentralizing Privacy: Using Blockchain to Protect Personal Data’. Pp. 180–84 in *2015 IEEE Security and Privacy Workshops (SPW)*.

Zyskind, Guy, Oz Nathan, and Alex Pentland. 2015. ‘Enigma: Decentralized Computation Platform with Guaranteed Privacy’. *arXiv Preprint arXiv:1506.03471*. Retrieved 19 October 2016 (http://arxiv.org/abs/1506.03471).