Workshopping a Data Equity Manifesto

Organizers
Catherine D'Ignazio, Emerson College, Boston, MA, USA
Ari Happonen, Lappeenranta University of Technology, Lappeenranta, Finland
Firaz Peer, Georgia Institute of Technology, Atlanta, GA, USA
Maria Palacin-Silva, Lappeenranta University of Technology, Lappeenranta, Finland
Annika Wolff, Lappeenranta University of Technology, Lappeenranta, Finland

Abstract. The goal of this workshop is to generate a publicly shareable manifesto around data equity. Data equity refers to the degree of fairness in responsibilities and benefits, opportunities trade-offs that all members of society experience as a result of civic datasets. As a range of civic datasets about the government, environment, education and others become increasingly available, it is important to understand how current technologically-mediated practices can be improved to achieve better data equity and accountability for all, irrespective of their data literacy skills. We reflect on these datasets and technical practices through hands-on activities that have been specifically designed to expose the barriers that prevent individuals, communities, businesses, nonprofits and governments from engaging with data. We pay particular focus to the differentials based on sexism, racism and other forms of structural oppression that tend to go underexamined within such settings.

Keywords: Data equity, civic data, manifesto

1   Relevance to iConference

Data driven decision-making is increasingly valued within business, nonprofits and government. Recent advancements in free online data analysis and visualization tools, the publicizing of previously confidential data sets have added to the mythology and hype around Big Data [2]. Citizen sensing (also known as ICT enabled citizen science or crowdsensing) is also a popular means for empowering citizen participation in the management and governance of cities [7, 8]. Data is also collected by other actors such as IoT devices, cameras, satellites, and traces of online activity. Many cities, organizations and even private individuals are opening up their data with the aim of making governments more accountable and supporting the innovation of new data-driven products and services. Engaging individual organizations and the wider public in telling their own data stories has been as aspiration of the Open Data Movement [4]. But open access does not always translate into equity in consumption [4]. This so called data revolution has resulted in a data divide, where those with privileged access and knowledge about such data are given a seat at the bargaining table, while the voices of those who lack such skills, continue to be ignored [3].

Janssen et al. [5] have identified a number of technical difficulties that act as barriers to open data use, including 'a lack of ability to discover appropriate data', 'no explanation of the meaning of data', 'difficulty in searching and browsing due to no index or other means to ensure easy search for finding the right data', 'even if data can be found, users might not be aware of its potential uses' and 'focus is on making use of single datasets, whereas the real value might come from combining various datasets'. Essentially, open data interfaces are not user friendly, there is a lack of interoperability between platforms [6] and open data is often released in a raw form (i.e. data filtering and cleaning is left to the end user) that is hard for users, especially non data experts, to understand [9]. This can be especially problematic for very large datasets where even a single attribute might have millions of data points. Even for expert users, some data processing and visualization is often needed to find patterns and to begin to understand what information the data can convey and how it might be used.

2   Purpose and Intended Audience

Our workshop will introduce participants to computer, paper and craft-based activities that have been designed to make civic data more accessible and overcome barriers to using it. The data interfaces and scenarios we are proposing for this workshop have been designed specifically to remove barriers that prevent individuals and communities with limited data literacy from engaging with data. These new interfaces will be used to represent carefully selected, real-world datasets and to provoke discussions on different aspects of equitability. They will be used alongside more traditional methods for representing either raw data, or curated data summaries. In addition, we will use metrics such as evidence of creativity, questions asked, identifying equitability or ethical concerns and combining data to assess which interfaces and tools better support our goal of identifying and overcoming barriers to a civic data culture. Reflection on these tasks will provide feedback on the data interfaces and provide clear goals for improvement. More importantly, the reflection will broaden our understanding of data equitability issues for civic data.

Our data scenarios will consider the whole life cycle of data, from conception of an idea that frames data collection, setting up the collection, storage of data, licensing, curation, combination with other datasets, re-use and sharing data with others. Within each scenario, we will reflect on the socio-political nature of these datasets, the biases they come laden with, the values and ethical issues they harbor and the practices they enable or preclude through the multiple actors involved.

3   Proposed Format

The proposed one-day workshop will be structured as a series of interactive activities and feedback sessions. We would ideally hope to have between 20-25 participants for the workshop. The morning session will focus on introducing and discussing the concept of data equity through hands on activities, where attendees will work within small

groups (4-5 each) to work through and deliberate on ideas. The afternoon session will also be broken up into activities but will focus on collaboratively writing different aspects of the manifesto.

We are interested in learning how our attendees define the term 'data equity' within their own research and practice, and the tools, methods and techniques they use to achieve it. In addition to the demonstration of our methods, our proposed schedule also sets aside time for discussions and to brainstorm additional techniques that can be used within this realm. We would like to get a sense of what empowerment through data means to our participants and the organizations they collaborate with. How can designers negotiate issues of racism, sexism, power and privilege differentials in relation to access and skills of working with data?

4 Goals or Outcomes

Our goals in proposing this interactive workshop to the iConference audience are twofold. Firstly, we offer the activities as methods that they can use to engage with individuals and communities who are new to working with data. We hope to create connections between the iConference audience and practitioners within the civic data space to learn from each other and inform this burgeoning field. Secondly, we hope to catalogue the conversations and reflections resulting from this workshop into a publicly shareable data equity manifesto, not unlike the Karlskrona Manifesto for Sustainable Design [1]. Participants will be invited to collaboratively author the manifesto, which will address issues around data literacy, engagement, empowerment, access, power and privilege that are typically associated with civic data and its associated technologies.

We have reserved web space under https://datacultureproject.org/dataequitymanifesto and www.civics.ninja/dataequitymanifesto to live blog the workshop and document the day's proceedings. A Twitter hashtag #dataequitymanifesto will be created for the event.

References

1. Becker, C. et al.: The Karlskrona manifesto for sustainability design. (2014).
2. Boyd, D., Crawford, K.: Critical Questions for Big Data. Information, Commun. Soc. 15, 5, 662–679 (2012).
3. Boyd, D., Crawford, K.: Six provocations for big data. A Decad. internet time Symp. Dyn. internet Soc. (2011).
4. Gurstein, M.B. (2011): Open data: Empowering the empowered or effective data use for everyone? First Monday. 16, 2, 1–8 (2011).
5. Janssen, M. et al.: Benefits, Adoption Barriers and Myths of Open Data and Open Government. Inf. Syst. Manag. (2012).
6. Martin, C.: Barriers to the open government data agenda: Taking a multi-level perspective, (2014).

7. Palacin-Silva, M., Porras, J.: Shut up and take my environmental data! A study on ICT enabled citizen science practices, participation approaches and challenges. Presented at the .

8. Rotman, D. et al.: Motivations affecting initial and long-term participation in citizen science projects in three countries. In: iConference 2014 Proceedings. (2014).

9. Weerakkody, V. et al.: Open data and its usability: an empirical view from the Citizen's perspective. Inf. Syst. Front. (2017).