Key points

- Work on gender equity is an important new frontier for open data communities, drawing on, and connecting to, many decades of wider work on feminist advocacy and action.
- Gender impacts issues of access and infrastructure, representation within datasets, and labour and leadership within open data communities.
- A commitment to gender equity must address both gender bias in data collection and publication, as well as in patterns of exclusion.
- Action is required to challenge embedded patriarchal attitudes and norms and practical steps must be taken to create better data, adopt clear policies, and promote opportunities and safe spaces for women.

Introduction

Gender is on the open data agenda and Sustainable Development Goal (SDG) 5 directly commits governments to “achieve gender equality and empower all women and girls”, specifically targeting the use of “communications technology to promote the empowerment of women”.

Groups like Data2X have expressed the need for more attention to be given to gender when data is collected, so that progress toward these goals can be tracked, and, since late 2017, the #MeToo movement has drawn global attention to the threats that women face in many professional settings. However, in the open data field, gender equity is still an emerging issue, often contextualised as a new frontier for progressive work. This chapter will explore the early movement to embed gender considerations in open data policy and practice and look at the challenges to be addressed in the years ahead.

Gender equity is defined as “fairness of treatment for women and men, according to their respective needs. This may include equal treatment or treatment that is different, but which is
considered equivalent in terms of rights, benefits, obligations, and opportunities." Although "gender equity" differs from "gender equality", the two concepts are inherently linked. Gender equity involves practices and processes that contribute to attaining gender equality as an end goal. This involves challenging and changing socioeconomic mechanisms around us on a day-to-day basis. On the other hand, gender equality as an end goal has been rooted in constitutional provisions, women's rights, and the labour movement ever since the principle was first recognised in international law in the 1948 Universal Declaration of Human Rights.

Given that the foundational principles of open data involve the promotion of accessibility and usability for all, gender equity should be an integral element in making this a reality. However, like gender-related work in many other spaces at present, gender data efforts are often siloed within the open data field. Although work to mainstream gender considerations within the open data space has already been a long process, encountering many institutional challenges and norms that stifle progress, it is still early days. The open data community has much to learn from pre-existing and long-standing movements for gender equality, as well as from the broader feminist literature, particularly with regard to its collective understanding of its approach to data, and how data collection and use relates to social justice and equity.

A note on limitations

The chapter does not explore issues beyond the binary. Our focus on the gender binary is not meant to be at the expense of, or to overlook, the non-binary. It is rather an effort to find initial avenues to open up political and institutional acknowledgement of the marginalised. We do, however, acknowledge this limitation of the chapter. Moreover, gender (even in binary terms) is viewed in this chapter as a separate issue from other factors that may marginalise the role of women, such as race.

From invisibility to awareness

At the start of this decade, it would not have been uncommon to attend major open data events and experience all-male panels or to observe that very few women were attending, let alone participating as speakers. Gender was rarely, if ever, discussed as a topic at these events. However, over the last few years, gender awareness and discussions of gender issues have become more central to the agenda of forums, conferences, workshops, and sessions around the world.

In 2016, the TechMousso (TechWoman) open gender data competition was held in Abidjan, Cote d’Ivoire, organised by the Millennium Challenge Corporation, Data2X, and the World Wide Web Foundation (Web Foundation) to highlight the difficulties of obtaining data on gender and on women’s organisations. During that same year, the first dedicated open data and gender session was held at the 4th International Open Data Conference (IODC) in Madrid, Spain. It was followed by sessions at the Open Government Partnership (OGP) Summit in Paris, France, where the Web Foundation held a workshop on why gender equality matters for open government.
and the Open Heroines network of women in open data, open government, and civic tech held their first Open Gender Monologues session.⁹

### Creating spaces for gender discussions

The growth of formal and informal networks working to advocate for, and support, women’s representation and participation in data and, at times, open data debates and discussions has been a positive sign over the past few years. While open data is the explicit focus of some of these networks, it is implicit or indirect in others.

**Open Heroines** ([https://openheroines.org/](https://openheroines.org/)) was launched in January 2016. Organised around a Slack team (instant messaging), the Open Heroines network was established to provide “a safe virtual space where women can meet like-minded women, share their experiences, and get advice and words of encouragement”.¹⁰ The network has hundreds of members around the world and was recently able to access funding to provide travel grants to IODC 2018.

**GeoChicas** ([https://geochicas.org/](https://geochicas.org/)) is an initiative that aims to close the gender gap in the OpenStreetMap community through collaborative and participatory projects in Latin America. Currently, women are substantially underrepresented among those who contribute to collaborative online projects like OpenStreetMap, and GeoChicas is one of a number of grassroots projects aiming to create balanced representation.

**Data2X** ([https://www.data2x.org/](https://www.data2x.org/)) is housed at the United Nations Foundation and is an alliance “dedicated to improving the quality, availability, and use of gender data in order to make a practical difference in the lives of women and girls worldwide”.¹¹ Created in 2012, Data2X has established a range of partnerships, focusing on issues related to big data, data on displaced populations, SDG monitoring, financial inclusion, and women’s work and employment. To date, they have not developed a specific open data partnership.

**EQUALS** ([https://www.equals.org/](https://www.equals.org/)) is a global partnership of corporate leaders, governments, non-profit organisations, communities, and individuals around the world working to bridge the gender divide in technology. Founded in 2016 by the International Telecommunications Union (ITU), UN Women, the International Trade Centre, Global System for Mobile Communications (GSMA), and the United Nations University, the partnership focuses on capacity building, political leadership, and action projects across four areas: access, skills, leadership, and research.

In 2017, the UN World Data Forum in Cape Town, South Africa, held important sessions on gender-disaggregated data. This was followed by the ATGENDER (European Institute for Gender Equality) Spring Conference on gender mainstreaming and gender budgeting¹² in Vilnius, Lithuania, and the “Open Data Brussels goes gender-smart” hackathon¹³ in Belgium. Also in 2017, at the Africa Open Data Conference in Ghana, a panel was held to discuss building an African open data gender agenda,¹⁴ and, in Wonolelo, Indonesia, women’s groups discussed the importance of having access to women’s development budgets as open data.¹⁵
In 2018, Article 19 Brazil hosted Dados y Feminicid (Data on femicides in Brazil) on Open Data Day, and the OGP launched the Feminist Open Government (FOGO) initiative to advance gender equity and hosted a high-level panel on the initiative at the OGP Summit in Tbilisi, Georgia. Additional events to incorporate gender and open data took place at IODC 2018 in Buenos Aires, Argentina, and at the UN World Data Forum in Dubai, UAE.

Taken together, these events indicate a gathering momentum with gender being increasingly addressed at key meetings of the open data field. However, it is notable that few other chapters in this volume are able to highlight sector-specific work on gender and open data, suggesting that global attention may not yet have filtered through to action in more specific sectors of open data work. This may reflect the reality that we are in the early stages of gender mainstreaming in open data, although it also highlights the risk that gender could become a sub-field separated from other major focus areas rather than an integrated strand within all open data action.

In preparing this chapter, the authors sought community input to identify stakeholders involved in supporting and championing efforts to bring a gender focus to work on big data, data for development, open data, and civic technology, mapping out the main stakeholders geographically, their primary focus, and the programmes they are working on. The two graphics below provide a summary, and a full table with details of each programme can be found in the online version of this chapter. While not exhaustive, this mapping gives an indication of the breadth, and limitations, of work on gender and data. In analysing the list of programmes and organisations, it is important to note, in particular, that many organisations are also working on advancing data literacy, gender equality, sustainable development, and women’s human rights. In addition, there are a range of funders supporting these activities, as well as a number of new organisations that have been established.
From awareness to action

There are three main challenges that organisations working on gender and data need to address: the capacity of women to contribute to and use data; the representation of women in data; and the representation of women in labour and leadership within open data organisations.

Access

Although open data advocacy often assumes that making data accessible equals putting data online, one cannot ignore the exclusions created by unequal access to technology. The digital world still has many gaps and limitations and is effectively inaccessible to the 50% of the world’s population who are still offline. Women are overrepresented among those without internet access. In 2015, women in poor urban communities of the Global South were 50% less likely to use the internet than men. In 2017, this finding was confirmed by ITU data that identified, on average, a 12% gap in internet access between men and women, particularly in Asia (17%) and
Africa (25%). This gender-based digital divide prevents women from making full use of the internet. Furthermore, research findings indicate that the dramatic spread of mobile phones is not enough to get women online, as well as that many potential applications of open data are not optimised for mobile phones.

Given the dire reality of the disparity in infrastructure and access, it is no surprise that translating the availability of open data into meaningful applications and inclusive use has been an ongoing challenge. The prerequisite to open data, digital connectivity, cannot be overlooked. Although, as Stephen Song suggests in Chapter 14: Telecommunications, open data on telecoms would help in advocating for better connectivity and addressing regional inequalities. Open data advocates also need to speak up more strongly on issues of internet access. Ultimately, many of the potential benefits of open data are available only to those who have access to internet-connected mobile and/or desktop technology. As a result, anyone developing an open data project needs to be able to answer questions related to who will be able to access its results, what consequences does unequal access have for the project, and whether they can be mitigated.

Even when access is present, alarming impediments to women's participation in digital spaces can be found in gender-based violence (GBV), personal data protection, and privacy concerns. Women worldwide experience higher levels of online harassment than men. They may be deterred from contributing to data sources that could otherwise be used to surface and address policy issues because of fears about how personal data may be abused. Although the open data movement has generally focused on non-personally identifying information, many important datasets are derived from personal data, and, as a Web Foundation study conducted in Nigeria has shown, without strong data protection legislation, women may have little or no recourse when personal data breaches occur. For gender data activities to be sustainable, women need to feel safe and secure online, as well as confident that their personal data will be handled safely. Companies, governments, and online communities must work together to stamp out hate speech and misogyny online, making the web a safer space for women.

A key consequence of present patterns of unequal access to internet and data technologies can be found in unequal representation of women within data itself. In the next section, we address the considerations that open data creators, intermediaries, and users need to take into account to work toward gender equity.

**Representation**

Data is never neutral: neither politically, technologically, or socio-culturally. Questions of what to collect, how to collect it, and which analytical framework to apply in order to present it, must acknowledge implicit and explicit gender dynamics. Opening up data for those with internet connectivity is a start and can challenge societal gender biases in terms of who controls access to data. However, opening up data should be about more than access. It needs to involve opening that data to scrutiny and bringing into focus questions of how it was created and shaped, and, in particular, how women and issues affecting women's lives are represented in the data upon which decisions may be based.

There are two main issues to be addressed here: data gaps and data distortions. Awareness of both is growing, although, to date, more has been done on addressing gaps in data than has been achieved in creating critical awareness of the way that datasets and their analysis, knowingly or
unknowingly, present only a partial and politically loaded view of the world, inhibiting efforts to ensure gender equity.

A gender data gap exists when a dataset describing people, or phenomena affecting people, does not include gender disaggregation, or omits women entirely. Such gaps are often a result of an inherent lack of understanding of how problems affect different genders in different ways. For example, there are problems that women face that cannot be fully understood without specific data in areas such as land inheritance, wage discrimination, and maternal mortality. The Data4SDGs network has argued that “without data equality, there is no gender equality”, because when women are invisible in statistics, or statistics cannot tell a story of the comparative position of women and men, it is impossible to meaningfully track progress on gender equity and equality targets.

Data gaps may occur at the time of data collection or when data is aggregated for presentation. Data2X has mapped 28 critical gender data gaps faced in monitoring SDG performance on health, education, economic opportunities, political participation, and human security. They found gender data lacked coverage in terms of countries and/or regular country production. Almost half of the topics reviewed lacked international standards and complexity (information that would allow patterns and determinants of variables to be understood), and more than half of them lacked granularity.

Data2X findings show gaps were particularly pronounced in data relating to employment and entrepreneurship, access to childcare, connectivity, and participation in peace and security processes. While some gaps may exist because survey tools lack fields for disaggregation, others occur due to gender-biased assumptions. One of the most prominent examples of the latter can be found in household surveys that reference the primary activities of the “head of the household”, who is often assumed to be male. This kind of bias is also illustrated by the lack of data on women’s non-market activities.

**Femicide data case study**

In Latin America, Article 19 has drawn on the open data agenda in their work to create the report “Data on femicide in Brazil – #InvisibilidadeMata”. The project, while it has produced a dossier on the problem, has called for better synchronisation between state datasets that address aspects of violence against women, demonstrating how when data is more open, citizens and NGOs can engage in advocacy for a better data infrastructure and for reliable sources of data on problems that are being ignored or side-lined by a patriarchal political system.

The Latin America Open Data Initiative (ILDA) has also developed a programme of work on data, gender, and security issues across the region with a goal of exploring the development of data standards that could be used to monitor femicides.

As organisations explore complementing conventional data collection with use of private sector big data and citizen-sourced open data to monitor the SDGs in relation to women, approaches to identify data gaps will become more and more important. Beyond the Global North, open data initiatives need to undergo major restructuring efforts in terms of participatory data collection,
access, and use. To work toward advancing gender equity, it is critical that marginalised groups are consulted and have a say when new data collection or data release efforts are being designed. Moreover, inclusive design processes can help tackle existing institutional norms of discrimination and exclusion. As long as gender data gaps persist, any open datasets created based on raw data that does not adequately represent women will have limited potential to support transformative action on gender equity.

It is not only statistical data where women may or may not be represented. Connecting sex-disaggregated data to gender-responsive budgeting is another way that the gender data gap can be addressed. Although there have been a number of successful projects in this area, gender-responsive budgeting has not yet been widely adopted.

Data distortions occur when the classifications, categories, and instruments used to capture or analyse data embed certain gender biases; hence, technological affordances and choices play a significant role. Masters quotes Olson's work on library classification systems to illustrate this point. Olson describes how the designer of 1904 library classification standards envisaged a "singular public" who share the same cultural, social, or political interests. Because of the hierarchical nature of the library classification scheme, the introduction of a subdivision in those standards for "relations with women" served to emphasise a social hierarchy, positioning women in relation to men. This marginalisation within data structures can create marginalisation in the real world, shaping the layout of the library, the discoverability of information on women's experience, and marginalisation of the marginalised within library spaces. We live surrounded by a myriad of examples like this within data politics, built up over centuries and continuing to be created today.

Masters calls for "feminist data structuring processes, with equality as the goal", noting that this "would involve a great amount of reflection, articulation, and collaboration". This may seem demanding, but, over time, there have been efforts toward inclusive representation of women in data and classification, enabled and manifested by a shift away from hierarchical organisational technologies toward tagging systems. Open data can be transformed into new data structures, escaping the boundaries of the database within which it was originally collected. The legacy of prior bias is, nevertheless, hard to shift, as, for example, when tags for the writings of women authors still tend to be an expansion of traditional, male-centred classifications. Importantly, while open data does allow some new opportunities for engagement and experimentation with how data is structured, statistical classifications and formal datasets often use rigidly defined structures that may be qualitatively different from the free-form tagging of literary texts. As a result, it is important to think about how women are represented within the bodies setting data collection standards and about the kinds of actions needed to undo decades of accrued bias.

In addressing the representation of women within data, there will also be work needed that goes beyond the limitations of this chapter, addressing other biased classification systems of race, disability, and other areas of inequality. Indeed, work to include non-gender binary groups within data, and to sensitively balance a desire for data minimisation with a desire to understand the gendered dynamics of social policy, will need considerable thought and a commitment to dialogue.
Labour and leadership

Access and representation are issues that affect the entire data landscape, but there are also issues closer to home to be dealt with by the open data community. One of the most important of these relates to the role of women within open data organisations and projects. At the start of this decade, the stereotypical open data organisation was young, urban, white, and male. Anecdotal evidence suggests that this has shifted substantially, but there are still many barriers to gender equity within the institutions that work on open data, and we lack good data on diversity and equity across open data organisations that could be used to benchmark progress.

One indicative dataset to draw upon is the speaker list from the IODCs. From 2015 to 2018, the percentage of women on the speaker list grew from around a third to reach 49%; however, further analysis has found that “women filled 173 (51%) of 337 slots in the schedule, while the men filled 164. In other words, each man spoke, on average, in 1.36 sessions, while each woman spoke in 1.56 sessions or 15% more.” As this short study states, “This means that if you were a woman speaking at IODC, you were working 15% harder on average, a number coincidentally matching the gender pay gap of OECD [Organisation for Economic Co-operation and Development] countries.” Bursaries, travel funding, and active outreach can all be important to help challenge the reproduction of systemic gender biases in leadership development.

Important barriers to women becoming active in the open data field also include the limited availability of learning opportunities and the structure of many open data projects that rely on gaining skills during “free time”. Open data has parallels to open source, where researchers have argued that male participation is effectively “subsidised” by men’s access to higher paid employment and a lack of contribution to domestic labour in their free time, increasing their availability to contribute volunteer labour to “open” projects. This requires work on capacity building and data literacy for, and with, women, including raising awareness of the use of data for societal impact. Addressing such global imbalances will require increasing the resources allocated to building e-learning centres and tech hubs that women can use to gain access to internet and open data skills, especially in developing countries and rural areas.

It does not just matter whether women are present in open data organisations. The distribution of labour and the way it is valued also matters. There has been growing recognition that sustaining open data projects, which are often cross-organisational and cross-cultural collaborations, requires a substantial investment of emotional labour. This disproportionately falls to women and is often undervalued. Until the narratives around open data projects better account for the emotional labour and intellectual labour that makes for the success of projects, rather than focusing on solely technological innovation, recognition and reward is not likely to be equitably distributed.

When it comes to gender data projects specifically, women often work on and drive these projects alone. Ultimately, the issue of gender data work should not be solely placed on the backs of gender data organisations. Gender issues need to be better mainstreamed into the work of all open data organisations. There are promising signs of this in organisations such as ILDA which has sought to establish internal self-reflection processes around gender dynamics within a gender-mixed project team working on data and security. Building on this and working across
organisations to create good practice resources and handbooks that can be used to improve organisational policy and practice may be a good starting point.

In 2018, the year of the #MeToo movement, we cannot ignore the civic technology sector’s own challenges with respect to discrimination, harassment, and abuse. Events have revealed the need for change in individual attitudes and in organisational responses to harassment or abuse cases. More open data conferences now have codes of conduct in an effort to create safer spaces for women, but there is still a long way to go. Significant advances will only come with clear leadership within the open data field and critical reflection from everyone on how to create safe, inclusive, and empowering spaces. Gender is a personal issue; therefore, personal self-reflection among men is needed. This is especially true for men involved in technology and open data, who may see themselves as not contributing to the problem through direct discrimination, but who are contributing to the problem indirectly. This need for self-reflection is a part of a larger cultural challenge that needs to be addressed actively in both professional and personal capacities.

Conclusion

The chapter has explored key issues concerning gender equity in open data, including issues related to infrastructure and access, as well as the representation of women in datasets, open data labour, and open data leadership. It has also attempted to address where progress has been made to date and to sketch out key actions that need to be taken going forward. In particular, there is a need to build bridges between open data organisations and more specialised gender organisations. Action is needed at regional, national, and global levels. Datasets need to be generated, policies need to be adopted and implemented, opportunities need to be created and safe spaces need to be provided. As a community, open data stakeholders may want to consider a collective code of conduct against gender-based harassment that could be endorsed by organisations, governments, and companies. This would also strengthen any existing internal codes and policies and promote best practices.

The nature of empowerment must be explored by individuals and organisations in the attempt to better identify biases in current open data work and to understand what needs to be challenged, and, in turn, changed. Existing power dynamics in open data need to be actively addressed, especially those related to gender equity, race, and inclusion, in order to solve long-standing inequalities, or open data will continue to only empower the empowered.

As the drivers of a movement, the open data community must track the gender impacts of its own activities. Ultimately, patriarchal norms must be confronted to realise the full potential of gender equity in open data. Targeted interventions to address access and the needs of women and girls is vital, but there is a foundational need to change the perception of norms and the way people think about gender. This is not simply about the technology or data per se, but about the fundamental way we think about gender and gender dynamics.
Further reading

*Ada: A Journal of Gender, New Media, and Technology* is an open-access peer-reviewed journal featuring scholarship on gender, new media, and technology. http://adanewmedia.org


About the authors

**Ana Brandusescu** researches and advocates for a more transparent, inclusive, and responsible use of existing and emerging technologies through data and policy projects. These include initiatives on open data, data governance, open government, and digital equality.

**Nnenna Nwakanma** advocates for open data, open government, and the open web across the world, bringing together local and international stakeholders to advance an open and inclusive development agenda. Speaker, mentor, and leader, she inspires women to aspire more, learn more, and achieve more through technology.

The authors would like to acknowledge the contributions to this chapter of Tina Appiah, Dhanaraj Thakur, Fabrizio Scrollini, Nancy Salem, Paulina Bustos, Mor Rubinstein, Noémie Girard, Martin Bader, Selene Yang, Kate Vang, Joshua Powell, James McKinney, Manuel Acevedo, David Sasaki, Ruba Ishak, Policy Ignite, Gabriela Rodriguez, Jonathan van Geuns, and Steven Adler.

How to cite this chapter


This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence. It was carried out with the aid of a grant from the International Development Research Centre, Ottawa, Canada.
Endnotes


10. Open Heroines. (2016). Hello, we are Open Heroines. *Open Heroines Blog*, 16 May. https://openheroines.org/hello-we-are-open-heroines-8e7830e3b3a0

11. https://www.data2x.org/who-we-are/


Issues in Open Data I Gender Equity


28 Ibid.


37 Ibid.


