Leveraging Big Data to Advance Gender Equality

By Ahmed Nauraiz Rana

Gender equality and the empowerment of women and girls, one of the Sustainable Development Goals, is a highly complex and challenging undertaking. We must address multiple issues—discrimination, violence, education, employment, economic resources, and technology—and work across economic sectors, from agriculture to financial services. Achieving gender equality will require significant amounts of accurate data about the situations and struggles of women and girls. Globally, however, there is a major gap in data that is disaggregated by sex, and this gap often renders women's societal, cultural, and economic contributions and obstacles practically invisible. It can also exacerbate existing gender divides, feeding and reinforcing biases in social programs, access to financial and other services, economic opportunities, and even development programs designed to address gender inequality. Part of the solution may be in the form of big data, which, if used effectively, can provide the volume of data needed to portray women and their situations accurately, which in turn can inform the creation of evidence-based solutions.

The social and economic integration of women into society is increasingly becoming part of all development discourse. Various mechanisms are being employed around the world to shed light on the issues and inherent biases that women are subjected to, and numerous interventions focused on greater gender equality are being implemented. But the success of all these efforts is dependent on data that is verifiable, reliable, and ensures integrity. Just 21 percent of the data required to monitor 54 gender-specific indicators within the Sustainable Development Goals (SDGs) is current.[1]

Additionally, the overrepresentation of men in the tech field filters into content creation, with recommendation algorithms often trained on male-majority data. As a result, disaggregating data based on gender is critical to understanding how developing countries can help women living on the border between poverty and prosperity. Gender equality is a fundamental prerequisite for multiple development goals, so it is imperative to emphasize the fact that progress will falter without a data-driven focus.

Discrimination against women is a multifaceted phenomenon that spans economic sectors and is ingrained in societal practices. Issues such as land rights, access to education, financial inclusion, healthcare, gender-based violence, family planning, and many others can only be correctly addressed if evidence-based policies are formulated and progress is monitored in a quantifiable manner.

This is contingent on the use of big data—extremely large data sets that can be analyzed for patterns and trends. However, up-to-date data exists for only a small fraction of indicators that were developed to evaluate SDG #5, Gender Equality. As a result, most countries have never been able to measure more than three of the 14 indicators that were created to assess progress toward this goal.[2] Clearly, innovative approaches are needed to ensure effective data collection that is disaggregated by sex, which is defined by the United Nations as "data collected and tabulated separately for women and men."[3]

About the Author

Ahmed Nauraiz Rana, Associate Digital Economy Officer, Gender and Economic Inclusion Group, Economics and Private Sector Development, IFC. His email is arana4@ifc.org.

IFC | International Finance Corporation WORLD BANK GROUP

Creating Markets, Creating Opportunities

For private sector businesses, sex-disaggregated data is a necessity where the aim is to build consumer-centric business strategies and enhance the company's value proposition to specific market segments, including the women's market. It allows for the recognition of customer segmentation and the market opportunity.[4]
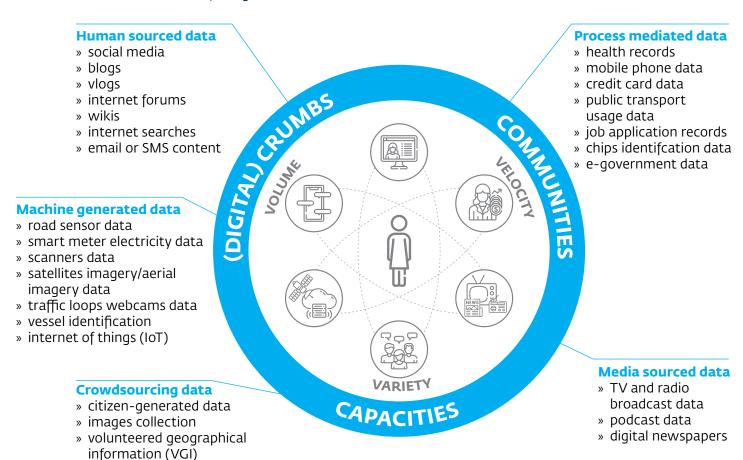
Inclusive approaches to gender data collection and usage have been shown to yield greater revenues, as well as numerous other nonfinancial positive outcomes such as employee retention, and operational replicability and scalability. Furthermore, a number of existing studies have identified various avenues through which sex-disaggregated data can assist women entrepreneurs in overcoming barriers to market access, financial inclusivity, and identification of prospective opportunities, and address value-chain bottlenecks using gender-specific information.

McKinsey Global Institute estimated that an increase in the level of gender equality alone will add $28 trillion to global GDP by 2025.[5] The revenue share was estimated to be higher for tech companies, start-ups, and industries where innovation is the key to growth, highlighting the fact that greater gender equality is not just desirable but is an integral part of successful revenue-generating businesses.

The data revolution currently underway can—and in certain places already is—being leveraged to achieve sustainable development goals pertaining to gender rights and equality. In 2017, for example, the Bill & Melinda Gates Foundation launched a tech initiative in partnership with UN Women to help countries improve the quality of gender-specific data collected globally.[6] Similarly, the United Nations Foundation spearheaded Data2X, a technical platform initiative to help close the daunting gender data gap in five development domains—health, education, economic opportunity, political participation, and human security—by collaborating with government agencies and the private sector.[7]

FIGURE 1 Current Data Sources for Big Data

Human sourced data
» social media
» blogs
» vlogs
» internet forums
» wikis
» internet searches
» email or SMS content

Machine generated data
» road sensor data
» smart meter electricity data
» scanners data
» satellites imagery/aerial imagery data
» traffic loops webcams data
» vessel identification
» internet of things (IoT)

Crowdsourcing data
» citizen-generated data
» images collection
» volunteered geographical information (VGI)

Process mediated data
» health records
» mobile phone data
» credit card data
» public transport usage data
» job application records
» chips identifcation data
» e-government data

Media sourced data
» TV and radio broadcast data
» podcast data
» digital newspapers

(DIGITAL) CRUMBS · COMMUNITIES · CAPACITIES
VOLUME · VELOCITY · VARIETY

Source: United Nations Entity for Gender Equality and the Empowerment of Women, also known as UN Women.

IFC | International Finance Corporation WORLD BANK GROUP

Creating Markets, Creating Opportunities

How Can Big Data Support Gender Equality?

Traditional data collection methods include surveys, interviews, and focus groups, among others. However, these methods have shown limitations in their ability to collect quality data about female subjects and target groups. They have also been found to incorporate stereotypes and other socio-cultural factors that induce gender biases. Such barriers raise valid concerns about construct and external validity, and lead to incomplete or inaccurate information that obstructs both the formulation of evidence-based policies and the determination of the root causes of gender discrimination.

Big data allows researchers and policy makers to transcend ineffective means of data collection. It offers policymakers and investors alike an additional evidence base and complements traditional forms of data collection. In recent years big data has evolved as a parallel source for understanding gender perceptions and forms of discrimination and marginalization due to its ability to detect large-scale data patterns and generate predictive models. A study commissioned by UN Women to understand the role of big data in evaluating women's political participation and leadership found that the analysis and interpretation of conversations within a cultural context can be significantly enhanced by focus groups with social media users.[8]

Areas of Intervention – Current Trends

Geospatial data accessed via satellite imagery can be leveraged to predict women's wellbeing. A number of research studies have found social and health indicators such as child stunting, infertility, literacy, and access to contraception to be correlated with geospatial factors such as climate, elevation, aridity, and geographical location.[9] And it is relatively easy to map such geospatial factors across countries and complement that data with information gathered via demographic and health surveys (DHS). This allows for the modelling of possible wellbeing indicators where data from DHS is not readily available, transforming inaccessible and neglected areas into a continuous landscape of information for gender wellbeing.

Similarly, mobile phone data can be leveraged to make better-informed decisions regarding women's social protection. A study commissioned by the Massachusetts Institute of Technology in collaboration with UN Global Pulse used data from credit cards and mobile phones to identify patterns of women's spending and physical mobility.[10] Using these to project women's economic status, researchers were able to identify seven economic lifestyle clusters among women in the dataset. The categorization of clusters allowed development professionals to understand individuals' economic status and needs. The technique helped identify women who were more vulnerable to economic downturn and shock, and improved targeting methods and interventions for extending socioeconomic protections.

Big data extracted from patterns of Internet use also aids in monitoring women's mental health. Most publicly available mental health data does not include sex-disaggregated information, yet there is a significant need for gender specific mechanisms for data collection and analysis. Internet use data can be leveraged to analyze the expression of thoughts on social media platforms and provide insights into women's mental health and welfare. To this end, researchers at UN Global Pulse successfully employed artificial intelligence machine-learning techniques to determine if expressions of distress or anxiety on social media accurately identified mental health disclosures. They analyzed publicly available tweets from women and girls in India, South Africa, the United Kingdom, and the United States to locate signals of depression and make appropriate and targeted recommendations.[11]

Of the many challenges that women face, one of the most staggering is access to financial services. For example, in

---

BOX 1 Coping With COVID-19 Through a Gender Lens

The COVID-19 pandemic poses a serious threat to women's employment and livelihoods due to its potential to exacerbate preexisting inequalities and reinforce gender gaps in social, political, and economic systems. From a lack of access to health services, social protections, and digital technologies, to a significant rise in domestic violence and unpaid care work, the impacts of COVID-19 are being felt acutely by women around the world. Recognizing the range of gender-based differences in the implications of the crisis is critical to ensuring that men and women have equal opportunity to benefit from response efforts and can participate in the eventual economic recovery.

As economies around the world gradually recover, it is advisable for private sector firms to fill information gaps by leveraging sex-disaggregated data as they begin the work of rebuilding a resilient economy. Assessing the gender-differential business impact of COVID-19 is absolutely essential to creating effective strategies and designing crisis solutions that meet the needs of both female and male entrepreneurs. Similarly, in order to develop a diverse and inclusive workforce that is resilient to another such crisis, it is imperative to collect, analyze, and utilize sex-disaggregated data to better understand gaps and how they might lead to lower productivity or profitability at the firm level.

IFC | International Finance Corporation WORLD BANK GROUP

Creating Markets, Creating Opportunities

Senegal some 87 percent of women lack access to any formal financial products or services.[12] This financial exclusion can be attributed both to gender discrimination and to a lack of data that may prevent financial institutions from making sound decisions. This lack of financial inclusion thus directly impacts women's ability to improve their earning capacity and engage in entrepreneurial activities, making them more vulnerable to economic shocks and downturns. An analysis of the Senegal River Valley rice value chain by Feed the Future revealed that financial institutions were hesitant to lend to female rice farmers due to perceived high risk, a lack of available collateral, and most important, a lack of information that prevents financial institutions from being able to evaluate risk and tailor lending terms to female consumers.[13]

Such a lack of information is not unique to agricultural lending; it occurs across economic sectors. Readily available data on key transaction factors can help address many of the concerns of financial services suppliers.

Data Innovation Projects Advancing Gender Equality

Various other projects using big data to address gender inequality issues have been undertaken. The following are a few of them.

Discovery of Complex Anomalous Patterns of Sexual Violence in El Salvador. A Carnegie Mellon University project applied data mining techniques to uncover complex anomalous spatio-temporal patterns of sexual violence. The researchers identified patterns after analyzing data on rape incidents in El Salvador over a period of nine years. These patterns helped formulate real-time early detection that would allow law enforcement agencies to initiate early rapid response.[14]

The Everyday Sexism Project is an initiative by the Oxford Internet Institute to create a semantic map of sexism via the application of a natural language processing approach that analyzes a large-scale, crowd-sourced dataset on sexism collected from the project website. The goal is to develop an advanced sociological understanding of women's life experiences of sexism and compute a methodological, evidence-based approach for modelling relevant interventions.[15]

Two-way Radio: Engagement with Somali Women. Africa's Voices, a University of Cambridge tech startup, designed five interactive radio shows on gender and child protection issues to address the gaps in data and evidence pertaining to (i) female genital mutilation/cutting, (ii) child marriages, (iii) girls' access to education, and (iv) juvenile justice. The show was broadcast across 27 stations and was leveraged to collect data via SMS messages in local dialects from radio audiences. The data that was collected was instrumental in generating insights regarding cultural beliefs and practices and was further used to develop gender equity programs and map health vulnerability areas.[16]

Identifying Trends in Workplace Discrimination. This combined initiative by UN Global Pulse, the International Labour Organization, and the Government of Indonesia was designed to address discriminatory behavior against women in the workplace in Indonesia. Researchers collected and analyzed data from publicly available tweets over three years to identify real-time signals of discrimination. The data helped understand women's unwillingness to report experiences related to discrimination and violence in the workplace.[17]

Mapping Indicators of Women's Welfare: Literacy, Stunting, Poverty, and Maternal Health. An initiative of World Pop/Flowminder and Data2X is developing modelling techniques that use high-resolution geospatial data to infer similar high-resolution patterns of social and health phenomena across entire countries. Based on deduced correlation, the method is used to predict social and health outcomes where surveys have not been performed, generating a series of highly detailed maps that illustrate landscapes of gender inequality, including stunting, literacy, and the use of modern contraceptives.[18]

Big Data and the Cloud: Piloting eHealth. The Government of Ghana, in partnership with the World Bank Group, implemented a performance-based financing project in four regions with particularly poor maternal child health nutrition outcomes to incentivize community health teams to improve female health outcomes. The quantity and quality of performance-based indicators was assessed by means of data collected directly from beneficiary communities, using an Android-based survey tool.[19]

Girl Effect's Springster Mobile Platform. Researchers from Girl Effect have been using a wide range of techniques to comprehend how user engagement affects girls' lives after they engage in conversations with other users, and/or read content on Springster, a global mobile-first platform developed for confidence and skill development of vulnerable girls aged 14 to 16. Researchers have been using Google Analytics, comment analysis, online surveys, and social media analytics to combine big data with traditional approaches to find innovative and improved ways of influencing positive change and strengthening female image building in an online space.[20]

Measuring Economic Resilience with Financial Transaction Data. The project, undertaken by BBVA Analytics, explored ways that financial transaction data could be analyzed to determine the level of economic resilience of people affected

IFC | International Finance Corporation
WORLD BANK GROUP

Creating Markets, Creating Opportunities

by natural disasters. It did not focus solely on women but did incorporate a gender lens that revealed differences in how women and men are impacted, including their differences in coping and recovery mechanisms. The project used the Mexican state of Baja California Sur as a case study to assess the impact of Hurricane Odile on livelihoods and economic activities over a period of six months. The project team measured daily point-of-sale transaction and ATM withdrawals at high geo-spatial resolution to gain insights into the way people prepare for and recover from disasters. Findings revealed that women increased their spending twice as much as men in preparation for the disaster, and that they took much longer than men to fully recover from it. These findings are instrumental as a starting point to design programs to assist and rehabilitate women after natural disasters.[21]

Conclusion

Big Data can play a fundamental role in achieving gender equality and empowering women and girls across the globe by identifying multi-sectoral gaps in the provision of equal opportunities and the protection of female rights, and by aiding in the implementation of evidence-based policies and interventions. Data from sources such as radio transmissions, social and digital media, call records and mobile network operations, and satellite imagery, both alone and in combination with traditional data sources, can help highlight the needs and concerns of women and girls. However, it is important to be cautious of the limitations and concerns that big data poses, including careless interpretations that can lead to disproportionate representation and biased recommendations.

Data collection and processing requires an adequate framework, extensive digital infrastructure, stringent regulations for privacy protection, and tools to mitigate risks of harm to data subjects. Access to big data is also challenging due to the high associated costs and technical expertise required to retrieve scattered information. However, given these considerations, big data analytics has enormous potential for policy makers and investors as a fast-evolving source of information that can be leveraged to gain valuable insights regarding women and girls.

Please see the following additional reports and EM Compass Notes about technology and its role in emerging markets: *Reinventing Business Through Disruptive Technologies - Sector Trends and Investment Opportunities for Firms in Emerging Markets* (March 2019); *Blockchain: Opportunities for Private Enterprises in Emerging Markets* (January 2019); *Accelerating Digital Connectivity Through Infrastructure Sharing* (Note 79, February 2020); *Artificial Intelligence and the Future for Smart Homes* (Note 78, February 2020); *Artificial Intelligence and 5G Mobile Technology Can Drive Investment Opportunities in Emerging Markets* (Note 76, December 2019); *How Artificial Intelligence is Making Transport Safer, Cleaner, More Reliable and Efficient in Emerging Markets* (Note 75, November 2019); *Bridging the Trust Gap: Blockchain's Potential to Restore Trust in Artificial Intelligence in Support of New Business Models* (Note 74, Oct 2019); *Artificial Intelligence: Investment Trends and Selected Industry Uses* (Note 71, Sept 2019); *The Role of Artificial Intelligence in Supporting Development in Emerging Markets* (Note 69, July 2019).

1   Mohapatra, Arti and Lauren Shields. 2018. "Can Data Improve the Lives of Women Around the World?" GreenBiz, 9 August 2018. https://www.greenbiz.com/article/can-data-improve-lives-women-around-world.

2   Suzman, Mark. 2017. "Data Driven Gender Equality." World Economic Forum, 18 September 2017. https://www.weforum.org/agenda/2017/09/gender-equality-it-starts-with-data.

3   United Nations Statistics Division. "Gender Statistics Manual – Integrating a Gender Perspective into Statistics." https://unstats.un.org/unsd/genderstatmanual/What-are-gender-stats.ashx

4   FAO. 2019. "Sex-disaggregated data in agriculture and sustainable resource management: New approaches for data collection and analysis." Rome. 2019. http://www.fao.org/3/i8930en/i8930en.pdf.

5   Executive Briefing. "The Power of Parity." McKinsey Global Institute. https://www.mckinsey.com/featured-insights/employment-and-growth/the-power-of-parity-advancing-womens-equality-in-the-united-states.

6   United Nations Entity for Gender Equality and the Empowerment of Women.

7   Noble, Eva. 2018. "Without Data Equality, There will be no Gender Equality." Women Deliver, 11 June 2018. https://womendeliver.org/2018/without-data-equality-there-will-be-no-gender-equality/.

8   Lopes, Claudia and Savita Bailur. 2018. "Gender Equality and Big Data." UN Women, January 2018. https://undg.org/wp-content/uploads/2018/02/Gender-equality-and-big-data-en.pdf.

9   Rogers, Kelli. 2017. "3 Ways Gender Data Could Go Big." Devex, 21 March 2017. https://www.devex.com/news/3-ways-gender-data-could-go-big-89862.

10 Rogers, Kelli. 2017.

IFC | **International Finance Corporation** WORLD BANK GROUP

Creating Markets, Creating Opportunities

[11] Rogers, Kelli. 2017.

[12] World Bank Group. 2016. "Enhancing Financial Capability and Inclusion in Senegal - A Demand-Side Survey." Finance & Markets Global Practice, No. ACS18885, June 2016. http://documents.worldbank.org/curated/en/371101467006421447/pdf/ACS18885-WP-P151555-PUBLIC-SENEGAL-Enhancing-Financial-Capability-and-Inclusion-Final-20160615.pdf.

[13] Miklyaev, Mikhail, Majid Hashemi, and Melani Schultz. 2017. "Cost Benefit Analysis of Senegal's Rice Value Chains." Development Discussion Paper Vol.4. https://cri-world.com/publications/qed_dp_301.pdf.

[14] De-Arteaga, Maria and Artur Dubrawski. 2016. "Discovery of Complex Anomalous Patterns of Sexual Violence in El Salvador." Data for Policy, 2016. https://mariadearteaga.files.wordpress.com/2016/05/discovery-complex-anomalous1.pdf.

[15] Everyday Sexism. http://everydaysexism.com/.

[16] Africa's Voices. 2017. "Child Protection and Gender Equality in Somalia." UNICEF. https://www.africasvoices.org/case-studies/unicef-somalia-child-protection-gender-equality/.

[17] Lopes, Claudia and Savita Bailur. 2018. "Gender Equality and Big Data." UN Women, January 2018. https://www.unglobalpulse.org/sites/default/files/Gender-equality-and-big-data-en-2018.pdf.

[18] Lopes, Claudia and Savita Bailur. 2018.

[19] UN. (no year). "Big Data Project Inventory." Big Data UN Working Group. https://unstats.un.org/bigdata/inventory/?selectID=WB43.

[20] Golant, Farah Ramzan. 2017. "Our Vision to Enable 100 Million Girls to Find Their VoiceOnline." Girl Effect, 18 October 2017. https://www.girleffect.org/stories/our-vision-to-enable-100-million-girls-find-their-voice-online/.

[21] UN Global Pulse. 2015. "Measuring Economic Resilience to Natural Disasters with Financial Transaction Data." Project Series, No. 24. https://www.unglobalpulse.org/projects/using-financial-transaction-data-measure-economic-resilience-natural-disasters.

Additional Selected EM Compass Notes Previously Published by IFC Thought Leadership

IFC International Finance Corporation WORLD BANK GROUP

Creating Markets, Creating Opportunities

IFC | International Finance Corporation WORLD BANK GROUP

Creating Markets, Creating Opportunities

IFC
2121 Pennsylvania Avenue, N.W.
Washington, D.C. 20433 U.S.A.

ifc.org/ThoughtLeadership

IFC | **International Finance Corporation**
WORLD BANK GROUP

Creating Markets, Creating Opportunities