GENDER EQUALITY AND BIG DATA

GLOBAL PULSE

UN WOMEN

This report provides background context on how big data can be used to facilitate and assess progress towards SDG 5 "Achieve gender equality and empower all women and girls". It examines successes and challenges in the use of big data to improve the lives of women and girls, and identifies concrete data innovation projects that have considered the gender dimension from across the development sector.

ISBN: 978-1-63214-109-5

Produced by UN Women Innovation Facility
Authors: Claudia Abreu Lopes and Savita Bailur
Design: Dammsavage Inc.
Cover Photo: Pulse Lab Kampala developed a radio content analysis tool to analyze radio talk in Uganda. UN Global Pulse.

Gender Equality and Big Data

MAKING GENDER DATA VISIBLE

January 2018

GLOBAL PULSE

UN WOMEN

TABLE OF CONTENTS

ACKNOWLEDGMENTS

This research report was developed by UN Women Innovation Facility with support from UN Global Pulse. UN Women would also like to thank Denmark for its generous financial support towards preparation of this report.

This report was written by Claudia Abreu Lopes, PhD and Savita Bailur, PhD. We thank Navin Haram from UN Women, and Felicia Vacarelu Josiane Toundzi Dzouankeu from UN Global Pulse for their helpful comments and guidance. We also gratefully acknowledge help and advice from many others, especially Fiona Bayat-Renoux from UN Women, Amanda Zerbe, Miguel Luengo-Oroz, Mila Romanoff, and Anoush Tatevossian from UN Global Pulse.

INTRODUCTION

In a rapidly evolving world, global challenges are becoming increasingly complex and interlinked. Innovative approaches are needed in order to improve the way the United Nations system responds and adapts to this changing environment, where established methods are not always sufficient to ensure progress towards goals in sustainable development and humanitarian response.

Data and information gaps are often impediments to achieving these goals. The 2030 Agenda explicitly calls for a data revolution for sustainable development (United Nations Department of Public Information, 2015). The Cape Town Global Action Plan for Sustainable Development Data was launched on 15th January 2017 at the inaugural United Nations World Data Forum, and it calls for a commitment by governments, policy leaders, and the international community to undertake key actions under six strategic areas of data for sustainable development: its dissemination and use; coordination and strategic leadership; innovation and modernization of national statistical systems; strengthening of basic statistical activities and programmes; multi-stakeholder partnerships; and mobilizing resources and coordinating efforts for capacity building.

The UN plays a very important role in ensuring that the data revolution for sustainable development is inclusive, not only by incorporating big data[1] and analytics into planning and decision making, but also by working with governments, policy leaders, and the broader international community to address the gaps in women's access to ICT and other tools and activities that generate new sources of data. Several pilot projects have demonstrated the feasibility of using big data analytics for sustainable development. For example, insights regarding mobility patterns, social interactions, sentiments and attitudes, economic activity, early warning, and community well-being can be derived from aggregated, privacy-protected datasets.

UN Women – the agency leading the UN's work on women's empowerment and gender equality – recognizes the potential of integrating big data and analytics into programmes and policies, and aims to identify applications of big data that could lead to impactful solutions in its areas of work. Global Pulse is a UN innovation initiative that works to implement and enable data science for development and humanitarian action in a responsible way, protecting individual privacy. Therefore, it is natural for these organizations to collaborate to ensure that the data revolution for sustainable development responds to women's needs and issues of data inclusion.

The objective of this collaboration is to identify strategic options in order to establish a joint programme of work around the use of big data and analytics to improve the lives of women and girls. UN Women and UN Global Pulse will work together to: (i) make programmatic recommendations that harness big data and analytics, including both "quick wins" and impactful long-term projects on gender; and (ii) identify potential advocacy projects and policy-oriented interventions to improve the lives of women and girls.

1. For the purposes of this report, big data refer to data sources that require new tools or methods to capture, curate, manage and process them in an efficient way.

ABOUT THIS REPORT

The potential of big data for sustainable development lies primarily in the application of insights from new data sources to inform policy interventions on the three pillars of sustainable development: economic, social, and environmental. There is still progress to be made in this area. Although pilot projects have shown the feasibility of using big data to assess and facilitate progress towards the Sustainable Development Goals (SDGs), there remains a dearth of examples that have scaled or become sustainable. Also, the methodological and technical expertise required to implement big data projects is not evenly distributed across geographies and organizations.

This report provides background context on how big data can be used to facilitate and assess progress towards the SDGs, and focuses in particular on SDG 5 – "Achieve gender equality and empower all women and girls". It examines successes and challenges in the use of big data to improve the lives of women and girls, and identifies concrete data innovation projects from across the development sector that have considered the gender dimension.

This report is the result of a literature review focused on big data and gender, interviews with colleagues from UN Women and UN Global Pulse, interviews with individuals and organizations working in the field of big data (LIRNEAsia, IDRC, World Wide Web Foundation, and the University of Southern California), and answers to a short questionnaire posed to six UN Women country offices.

A "gendered" DATA REVOLUTION

The data revolution has been driven by explosions in the volume of data, the speed with which data is produced, the number of producers of data, and the range of issues on which there is data. This data is generated by new technologies such as mobile phones and the emerging "internet of things," but also includes other data sources, such as social media and citizen-generated data[2]. In the context of sustainable development, this also suggests the integration of these new sources of data with traditional data in order to produce high-quality information that is more detailed, timely, and relevant. These shifts also create opportunities for greater openness and transparency, which must be leveraged in a way that protects privacy and human rights, taking into account data "inclusivity."

Big data presents opportunities for improved programme planning and implementation based on real-time feedback (Laney, 2001). Big data analytics can facilitate:

- real-time situational awareness;
- the ability to "shine a light on the invisible," by improving information on the lives of women and girls;
- new information on mobility, social interactions; sentiment and cultural beliefs, and economic activity;
- early warning of emerging issues and crises;
- improved understanding of community well-being;
- understanding of both local impacts and larger geographic patterns;
- identification of trends and correlations within and across large datasets that would otherwise be unknown;
- data visualization for more nuanced and accessible insights;
- opportunities for participatory monitoring, real-time feedback, and learning loops;
- the ability to recalibrate and iterate within the implementation of a programme; and
- improvements in accountability and transparency.

2. Defined by CIVICUS as data that people or their organizations produce to directly monitor, demand or drive change on issues that affect them (http://civicus.org/images/ER%20cgd_brief.pdf).

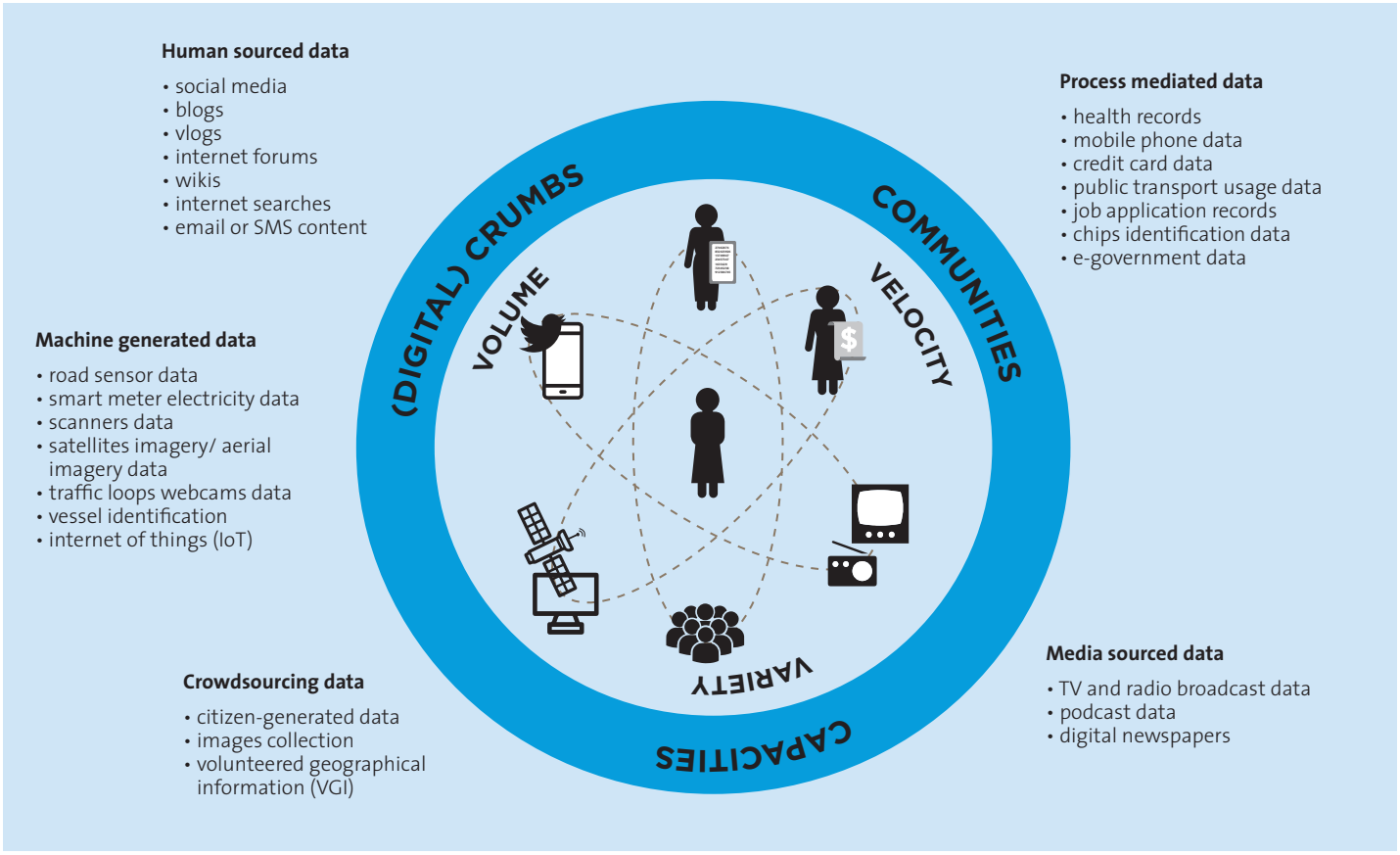
Big data was initially defined in terms of the "3 Vs" – volume (amount of data), variety (different types of data), and velocity (speed at which data is generated and transmitted) (Laney, 2001). While these properties emphasize the characteristics of the data, the social context of big data is better captured as an ecosystem of 3 Cs, as advanced by Letouzé (2015): digital crumbs (digital translations of human action and interactions captured by digital devices); powerful capacities to collect, aggregate and analyse data; and communities involved in generating, governing, and using data (including data generators, end-users, policy makers, privacy advocates, and civic hacker communities).

The realization of these benefits depends on an effective approach to data acquisition, analysis, and use, ensuring that big data serves as a complement to traditional data sources. Ultimately, the data revolution can lead to enhanced understanding and advocacy, more informed planning and decision-making, and more agile programme implementation and monitoring, all for improved results and greater impact.

Traditional data (i.e., household surveys, institutional records, or censuses) are often collected with a specific intention, following a structured format, and with valid and reliable measurements. While big data is not always collected in this way, the many forms of big data (illustrated in Figure 1.1) can help to close the gender gap by providing more granular, near real-time information, especially in locations where other sources of data are lacking (Vaitla et al., 2017).

FIGURE 1.1

Types, physical characteristics and social dimensions of big data[3]



Big data generally:

- is produced as a byproduct of people's digital behaviour;

- is not gathered in a way guided by a research question, and therefore requires interpretation after the event;

- imperfectly matches the entire universe of cases. Representativeness is generally not a factor in big data collection, as there is no sampling strategy involved, and non-coverage is often a concern when assessing data quality;

- is often accessible in real time (at the time that the data is produced). However, data analytics may require some days or weeks;

- can be analyzed by combining different data sources; multiple datasets can illuminate new insights and/or validate indicators, or help with triangulation of data sources;

- can be harnessed to improve decision-making. Appropriate guidance and frameworks can help to translate insights from big data into value for organizations, governments or businesses.

Adapted from Hilbert (2016)

There are several issues to consider in using big data to foster a more gender-responsive data revolution for sustainable development. These include:

- Addressing the gender gaps in "traditional" data – including areas where women's activities, women's needs, women's interests, and threats women face are largely invisible – to get a richer, more nuanced understanding of gender equality and women's empowerment issues.
- Real-time monitoring of gender indicators and progress (or women's perceptions on progress) on gender equality across the SDGs, including but also extending beyond SDG 5.
- Understanding the social norms and political realities around gender equality and women's empowerment in order to effectively interpret big data analytics. For example, what women are comfortable saying online may not reflect their opinions. Alternatively, there may be digital-world threaths faced by women that prevent them from feeling comfortable expressing themselves online, such as issues of privacy and security or the risk of online harassment and abuse.
- Addressing the gaps in women's access to ICTs and other technologies that generate data. The gender gap in internet access is 11.6% globally and 32.9% in Least Developed Countries (LDCs) (ITU, 2017). Moreover, there are differences in the way women and men use technology. There is also a gender gap in the degree of sophistication of use, as well as in the degree of control that women and men typically have over these resources and tools.
- Improved understanding by the development community of methodological issues of studies focused on gender that limit generazability such as self-selection of cases and context the data is produced.
- Understanding and properly addressing potential risks of harms that may result from the use of data, even in its de-identified form. Many of such risks may develop unexpectedly depending on a particular context. For example, gender by itself is considered sensitive information by many privacy related regulations due to the fact that knowing or being able to identify gender, even without knowing other demographic information, may lead to discrimination of an individual.

BIG DATA FOR SUSTAINABLE DEVELOPMENT

What is the added value of leveraging big data for gender?

In September 2015, Member States of the United Nations adopted, by consensus, the 2030 Agenda for Sustainable Development. The 17 SDGs and 169 targets contained in the 2030 Agenda constitute a transformative plan for people, planet, prosperity, partnerships and peace.

The data revolution was recognized as an enabler of the 2030 Agenda. It can not only help to monitor progress towards the SDGs, but it also engages multiple stakeholders to advance evidence-based policies and programmes aimed to reach the most vulnerable and leave no one behind.

> "Official statistics must move from sample survey paradigm of the past 70 years to a mixed data sources paradigm for the future; official statistics mindset must start with user needs for information for policy development, monitoring, evaluation, and understanding societal trends, and work backwards to best combination of data sources."
>
> Constance Citro (2015)
> -Constance Citro, Director of Committee on National Statistics, United States National Academy of Sciences

Big data and data analytics present myriad opportunities relevant to sustainable development priorities, especially in countries where traditional data is scarce or non-existent. There are also pragmatic aspects to consider: big data can be wider, deeper, more rapidly provided, and (in some cases), more cost-effective than traditional data. However, big data analytics also has its limitations and risks, such as lack of representativeness, self-selection bias, spatial auto-correlation bias, inadequacy for causal attribution and spurious correlations (Meltcaf et al., 2016).

One of the most promising aspects of big data is its potential to be integrated with other data sources. When combined with for example surveys, big data has the potential to illuminate societal trends and global patterns, or to shed light on more distal enablers of an event. However, it is also important to consider and mitigate the privacy risks that can arise from combining multiple datasets.[3] Recognizing the current challenges of surveys – which can include low response rates, differential population coverage, and privacy concerns – Constance Citro (2015), Director of the U.S. National Academy of Sciences Committee on National Statistics has advocated for a new paradigm in official statistics (see box on the right).

Given the potential of big data and data analytics to offer rapid feedback on a large scale, it can add value by yielding knowledge that can be used to inform interventions throughout their life cycles, from planning to implementation, and evaluation of development programmes. Big data can address multiple types of research questions, including those that are descriptive, predictive, and prescriptive (Letouzé, 2015).

3 These include the "mosaic effect," wherein combining multiple datasets can lead to the re-identification of individuals or groups, or to heightened re-identification risk. For more information, see Huber (2014).

Descriptive value: what is happening?

In cases where traditional data sources relevant to a SDG exist but the quality of the source cannot be verified (e.g., administrative records to measure domestic violence), analysis of big data can be used to validate indicators. Big data may also fill gaps where information does not exist: for example, mobile phone spending may serve as a proxy for income level in areas where this data is not available through other means.

Predictive value: what will happen?

Big datasets with a temporal dimension elucidate societal dynamics – such as a change in social norms – that cannot be directly shown by traditional indicators. Understanding temporal dynamics is critical in order to build and test robust theories of change that predict how a desired outcome evolves and whether it is affected by individual and systemic factors.

Prescriptive value: why did it happen?

Analysis of big data can also contribute to monitoring and achieving the SDGs by revealing the processes that led to particular results. Combining multiple data sources related to the same entity (communities, groups) may shed light on associations and patterns, revealing multiple aspects of a phenomenon. Although some associations may be spurious, consistent patterns (e.g., over time or across geographic contexts) may help to corroborate or refute hypotheses related to barriers to or enablers of a specific development target. Programme evaluation or theory building can benefit from a greater availability of big datasets, with the potential to illustrate the social mechanisms that bring about particular SDG outcomes.

Where in the 2030 Agenda does big data represent opportunities vis-à-vis gender?

Roughly a quarter of all SDG indicators (53 out of 230) explicitly or implicitly address gender equality.[4] In addition to the targets in SDG 5 –"Achieve gender equality and empower all women and girls"–, the targets in the other goals encompass a comprehensive set of gender-related issues, including the gender dimensions of poverty, hunger, health, education, water and sanitation, employment, safe public spaces, climate change, and peace and security.

However, there remains a general lack of statistical data in these areas, arising both from failure to prioritize gender equality in data collection and from lack of resources to collect and analyse data. Lack of investment often translates into pilot and research projects with limited reach, and which involve only a small number of stakeholders.

Of the 14 proposed indicators to monitor SDG 5, there are only three for which internationally accepted standards for measurement exist and for which data is regularly collected by most countries (referred to as Tier I indicators). Of the remaining 11 indicators, five have internationally accepted standards, but data collection by most countries is largely irregular (referred to as Tier II indicators). For the remaining six (referred to as Tier III indicators), international standards do not yet exist and most countries do not regularly collect the data (see Table 1).

Big data can have a role to play for all three tiers. For Tier 1 indicators, big data can help with questions of measurement and validation (for example, via bias correction, disaggregation, and higher-frequency measurement). For Tier 2 indicators, big data can help with indicator estimation in countries where data collection is irregular, via triangulation with other data sources. For Tier 3 indicators, big data can be used to analyse international and national trends.

4. "Gender-sensitive SDG indicators" here refers to (1) all indicators under SDG 5; and (2) indicators across the framework that explicitly refer to sex, gender, women and girls and/or are specifically or generally targeted at women and girls.

Which gender-related indicators in the 2030 Agenda could benefit from new data analytics?

TABLE 1:
SDG 5 indicators and relevant data sources

Tier	SDG 5 indicator	Traditional data sources	Big data sources
3	5.1.1 Number (%) of countries with legal frameworks that promote gender equality and non-discrimination against all women and girls	Policy analysis, CEDAW[5]	
2	5.2.1 Proportion of ever-partnered women and girls aged 15 years and older subjected to physical, sexual or psychological violence by a current or former intimate partner in the previous 12 months, by form of violence and by age.	Admin data, DHS[6] and VAW[7] surveys	Social media data, mobile phone surveys
2	5.2.2 Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner, since age 15, by age and place of occurrence	Admin data, DHS and VAW surveys	Social media data, mobile phone surveys
1	5.3.1 Percentage of women aged 20-24 years who were married or in a union before age 15 and before age 18 (i.e., child marriage).	Admin data, DHS and MICS[8] surveys	Mobile phone surveys
1	5.3.2 Percentage of girls and women aged 15-49 years who have undergone female genital mutilation/cutting, by age group	DHS and MICS surveys	Medical records, mobile phone surveys
2	5.4.1 Average daily (24 hours) spent on unpaid domestic and care work, by sex, age and location (individuals 5 years and above)	Time-use surveys	Mobile phone surveys
1	5.5.1 Proportion of seats held by women in national parliaments and local governments	Administrative data and household surveys	Radio broadcast data social media
1	5.5.2 Proportion of women in managerial positions	Administrative data and household surveys	Job applications data
3	5.6.1 Proportion of women aged 15-49 years who make their own informed decisions regarding sexual relations, contraceptive use and reproductive health care	DHS and MICS surveys	Medical records, social media data, mobile phone surveys
3	5.6.2 Number (%) of countries with laws and regulations that guarantee women aged 15-49 years access to sexual and reproductive health care, information and education	Country reports	
3	5.a.1 (a) Proportion of total agricultural population with ownership or secure rights over agricultural land, by sex; and (b) share of women among owners or rights-bearers of agricultural land, by type of tenure	Agricultural census and surveys	Mobile phone surveys and satellite imagery
3	5.a.2 Proportion of countries where the legal framework (including customary law) guarantees women's equal rights to land ownership and/or control	National laws	
2	5.b.1 Proportion of individuals who own a mobile telephone, by sex	Household surveys (e.g, ITU)	Call records data
3	5.c.1 Proportion of countries with systems to track and make public allocations for gender equality and women's empowerment	Policy analysis	

5. Convention on the Elimination of All Forms of Discrimination against Women
6. Demographic and Health Surveys
7. Violence Against Women Surveys
8. Multiple Indicator Cluster Survey

Similarly, many of the gender-related indicators to track progress toward the other goals – such as gender and poverty, gender pay gaps, women's representation in local governments, violence and abuse, access to justice, and indicators related to gender and the environment – currently lack compatible methodologies for comprehensive and periodic monitoring. A disaggregation by sex and age is needed for all indicators (since different age groups may have different experiences; for example, inequality is experienced differently by women and girls at different points in their lives).

Table 2:
Types of data suitable for other gender-related SDG indicators

Goal	Indicators	Traditional data sources	Big data sources
SDG1 End poverty in all its forms everywhere	1.1.; 1.2; 1.3; 1.4	Household surveys	Aggregated call records or other mobility data and satellite imagery
SDG2 End hunger, achieve food security and improved nutrition and promote sustainable agriculture	2.1; 2.2	Gallup World Poll/FAO, Survey, Household surveys	Crowdsourced data (such as SMS surveys), mobile money transaction data, social media data, radio broadcast data
SDG3 Ensure healthy lives and promote well-being for all at all ages	3.1; 3.4; 3.7; 3.8; 3.9	Administrative, census and household surveys	Social media data, SMS surveys, social media and radio broadcast data
SDG 4 Ensure inclusive and equitable quality education and promote life-long learning opportunities for all	4.1-4.7	Administrative and household surveys	SMS surveys, social media data, TV and radio data, digital news data, and citizen-generated data
SDG 7 Ensure access to affordable, reliable, sustainable and modern energy for all	7.1	Household surveys	Data from energy meters, satellite data, call records or other mobility data, and citizen-generated data
SDG 8 Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all	8.5	Household and establishment surveys	Web search data, TV and radio data, digital news data, and job applications data, social media data
SDG 10. Reduce inequality within and among countries	10.1	Household surveys	Credit card data or other financial transaction data
SDG 11. Make cities and human settlements inclusive, safe, resilient and sustainable	11.7	Household surveys, Administrative reports	Social media, mobility data from call records, satellite imagery, mobile surveys, TV and radio broadcast data, digital news data, citizen-generated data, and crowdsourced data
SDG 16. Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels	16.1, 16.2, 16.7	Administrative and household surveys	Social media, SMS surveys, TV and radio broadcast data, digital news data, and citizen-geneated data
SDG 17. Strengthen the means of implementation and revitalize the global partnership for sustainable development	17.8	Household surveys	Web search data,social media data, TV and radio broadcast data, digital news data.

POLICY AND GENDER

Assessment of policy issues related to big data and gender

To ensure that global sustainable development goals are realized and that progress is appropriately and responsibly tracked using big data, applications of big data analytics should respect individual privacy and the highest norms of ethical conduct.

While the use of big data has many benefits, especially in the context of the SDGs, this use raises questions about the potential harms posed to individuals and to identifiable groups or populations. The key to ensuring that data is utilized responsibly is to properly assess the likelihood of risks and related harms in proportion to the expected benefits of data use (United Nations Development Group, 2017). An appropriate risk assessment must consider the risks and harms that data use and non-use pose to fundamental human rights (UN Global Pulse, n.d., "Risks, harms, benefits assessment"). A balance needs to be struck between protecting privacy and maximizing the utility of big data for safeguarding civil rights, ensuring fairness, and preventing discrimination.

There are a number of long-standing data privacy principles (A/RES/45/95) that have been used and implemented through various regulatory frameworks around the world and by many international organizations. However, many issues that arise with the use of big data analytics, in particular, remain yet to be solved. For example, what is the threshold at which de-identified data is no longer personal, if the world of big data and connectivity no longer guarantees irreversible de-identification? When is it feasible (and how practical is it) to seek consent in situations of emergency, development response or when data is de-identified? When can data be repurposed for sustainable development? How to properly mitigate risks that are unique to the world of big data and sustainable development? How to ensure data is accurate and unbiased?

Data teams need to understand how insights drawn from the data may impact people's lives, and they must be thoughtful and responsible in the ways they present any conclusions to responders, civil society leaders or policy makers. In market research contexts, companies have already adopted some tools and practices with regard to big data; however, there are differences between this use and the use of big data for sustainable development and humanitarian action. A few possible mitigation strategies have already been offered by privacy and ethical frameworks that are adaptable to sustainable development and humanitarian contexts, such as the "UNDG Guidance Note on Big Data for SDGs: Data Privacy, Data Protection and Data Ethics" (United Nations Development Group, 2017) and the "UN Global Pulse Data Privacy and Data Protection Principles" (n.d.).

RESPONSIBLE USE OF BIG DATA: KEY ISSUES

Legitimacy and Informed Consent

Legitimate data use is one of the key pillars of responsible data handling. While consent is one of the main bases for legitimate data use, it is not always possible to obtain it. Furthermore, people whose data is being used may not be in a position to make informed choices or provide consent. They may not be aware of the implications (positive or negative) of their data being collected or used, have little or no awareness of their digital rights, and lack power to influence the process of data gathering (Antin et al., 2014). In this context, development agencies have a responsibility to protect the best interests of their beneficiaries. The key principles in these cases are to do no harm and to ensure accountability. Approaches to ensure responsible data use in such situations include transparency and education about the possible risks and benefits of data use and non-use; conducting risk assessments of the use of data in a given context; and understanding how well-informed consent is.

Data Bias

Incomplete representation in datasets – due to the exclusion of certain populations – can lead to biases in conclusions or recommendations based on data analytics. This in turn can lead to unfair or discriminatory decision-making (see O'Neill, 2016). Incomplete representation in data may perpetuate social inequalities. To reduce bias, data teams need to ensure proper data quality and ask themselves: Who is not represented in a given dataset? Whose realities are not reflected in this data? What is the impact of these exclusions when using data to inform programmes, conduct advocacy or implement policy?[9]

Data Access and Availability

To ensure adequate data-driven decisions, policymakers need accurate data, but most big datasets are currently held by the private sector. However, data collection and access is not always easy. Due to limitations posed by existing regulations, lack of incentives to share data for the private sector, and lack of capacity on the public sector side. While some open data initiatives and new data ecosystems are emerging, supported by networks like the Global Partnership for Sustainable Development Data and the Responsible Data Forum, there are not many examples of solutions for public-private data collaborations.

Data Re-identification and Group Harms

Interconnected environments and endless production of new data makes it hard, if not impossible, to anonymise data completely. This poses re-identification risks. While the data used in development contexts is mainly at the community level (minimizing the risk of individual re-identification), data teams still need to be mindful that even data used at the community level may still cause harm to identifiable groups (for example, if such groups are targeted in specific contexts).

Data Aggregation and Data Utility

Data aggregation is one of the most common ways to both protect individual privacy and present conclusions useful to policy makers. Data teams need to be aware of the risks of data use while also ensuring that the level of aggregation does not diminish the quality of the data. (UN Global Pulse, 2015). Data teams should employ the principles of data minimization, necessity, and proportionality when aggregating data – thus ensuring that only a minimum necessary level of detail in a data set is used to achieve the intended positive outcome of the data use.

Data Security and Data Capacity

Well-managed technical systems and well-trained human resources are important factors that influence data security. To prevent data breaches, data teams need to perform ongoing vulnerability assessments of their systems and undergo regular trainings on data privacy and data security.

9 These points were made in the interviews with the World Wide Web Foundation and LIRNEAsia.

COMPENDIUM OF DATA INNOVATION PROJECTS

This report provides a list of data innovation projects that were identified through a literature review and interviews with domain experts from various UN agencies. This list is not exhaustive, but includes a number of pertinent examples.

The projects listed below include examples of big data for sustainable development projects sourced from UN Global Pulse, the UN Global Working Group on Big Data for Official Statistics, the UNECE/Sandbox, Data2X, the NYU Governance Lab, the Data Science for Social Good programme at the University of Chicago, Flowminder, UNICEF, and the World Bank Group.

Projects related to SDG 5 focus on overall gender discrimination (5.1), gender-based violence (5.2), early marriage and Female Genital Mutilation (FGM) (5.4), and sexual and reproductive decision-making (5.6). Other projects were relevant to the gender dimensions of other SDGs, including projects focused on economic resilience to disasters, income inequality, financial inclusion and migration (SDG 1) and projects focused on maternal health (SDG 3). Other projects focus more on methods and tools that could be applicable to understanding the gender dimensions of several SDGs (for example, a Twitter tool to label gender, or a tool to analyse speech on radio broadcasts in local languages).

PROJECT 1. Using machine learning to analyse public radio content in Uganda (UN Global Pulse, 2016, "Making Ugandan community radio machine-readable using speech recognition technology")

Country	Year	Organization/ Partners	Data sources	SDG
Uganda	Ongoing	UN Global Pulse, Makerere University (Uganda), and Stellenbosch University (South Africa).	Radio	Cross-cutting

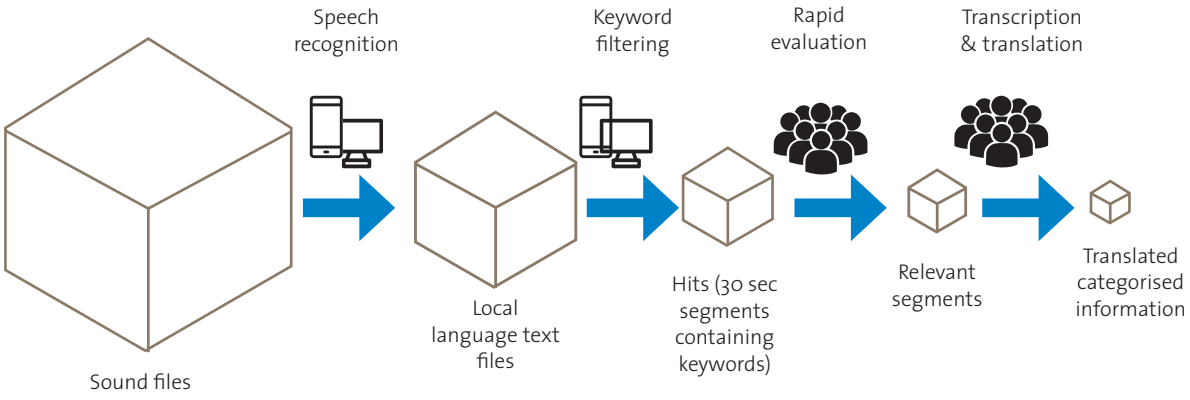
UN Global Pulse, through its lab in Kampala, has been working to leverage big data from radio conversations to inform the development of projects and programmes on the ground, especially in countries where radio has remained a main source of information-sharing.

The Radio Content Analysis tool was developed as part of a project conducted by Pulse Lab Kampala, in collaboration with the Stellenbosch University in South Africa. The tool works by converting public discussions that take place on radio in various African languages into text. Once converted, the text can be searched for topics of interest. The tool is now fully functional in the Northern and Central regions of Uganda and available for three languages: Luganda, Acholi and English (as spoken in the country).

To assess its validity, the Lab worked with the government and UN partners on a number of case studies. By sampling different indigenous languages, types of broadcasters, and locations within Uganda, the pilot studies assessed the potential uses of radio talk across five topics: perceptions towards refugees in Uganda, the impact of small-scale disasters on livelihoods, perceptions on the delivery of healthcare services, understanding the spread of infectious diseases, and monitoring the effectiveness of awareness raising radio campaigns.

Results yielded that relevant information can be extracted from opinions expressed over radio regarding for example the state of healthcare service delivery or the levels of acceptance of host communities towards refugees. Additionally, data analysis of radio content showed it is a valuable source for evaluating the effectiveness of behavioural change campaigns broadcasted on radio.

Figure 2:
Overall radio mining process for converting a large quantity of unstructured audio data into a small quantity of relevant, structured information.



Overall radio mining process for converting a large quantity of unstructured audio data into a small quantity of relevant, structured information.

Website: *http://radio.unglobalpulse.net/uganda/*

PROJECT 2. Mapping indicators of women's welfare at high spatial resolution (literacy, stunting, poverty and maternal health indicators) (Bosco et al., 2017)

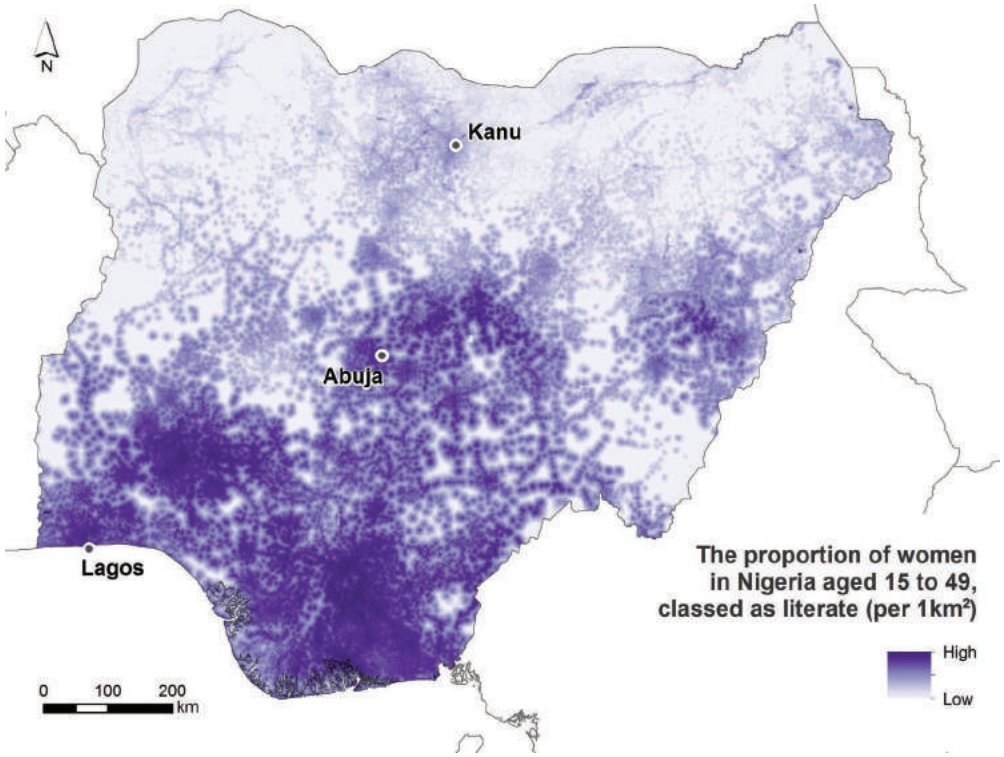
Country	Year	Organization/ Partners	Data sources	SDG
Bangladesh, Nigeria, Haiti, Kenya, and Tanzania	Ongoing	World Pop/Flowminder UN Foundation and Data2x program	Satellite and GPS-located household surveys	Cross-cutting

This project develops modelling techniques that use publicly-available high-resolution geospatial data to infer similarly high-resolution patterns of social and health phenomena across entire countries.

The approach takes advantage of the fact that many types of social and health data are correlated with geospatial phenomena. These relationships can predict social and health outcomes in areas where surveys have not been performed but correlated geospatial data is available.

The project has generated a series of highly detailed maps that clearly illustrate landscapes of gender inequality, including in stunting and literacy. The project also assessed use of modern contraceptives by women in some countries.

Figure 3:
Female Literacy Levels in Nigeria. Credits: Bosco, C. et al. 2017



Female Literacy Levels in Nigeria. ©Bosco et al., 2017
http://www.worldpop.org.uk/about_our_work/projects/index.php?sheet=Mapping-literacy

PROJECT 3. Discovery of complex anomalous patterns of sexual violence in El Salvador (De-Arteaga & Dubrawski, 2017).

Country	Year	Organization/ Partners	Data sources	SDG
El Salvador	2016	Carnegie Mellon University (Maria De-Arteaga and Artur Dubrawski)	Administrative data (reported rapes)	5.2

This project applied data mining techniques to unveil complex anomalous spatio-temporal patterns of sexual violence. The researchers analysed all reported cases of rape in El Salvador over a period of nine years, identified patterns, and explained how such analyses could be conducted in real-time, enabling early detection of emerging patterns that would allow rapid response by policymakers and law enforcement.

Figure 4:
Rate of total reported cases of sexual violence by 10,000 inhabitants in El Salvador.

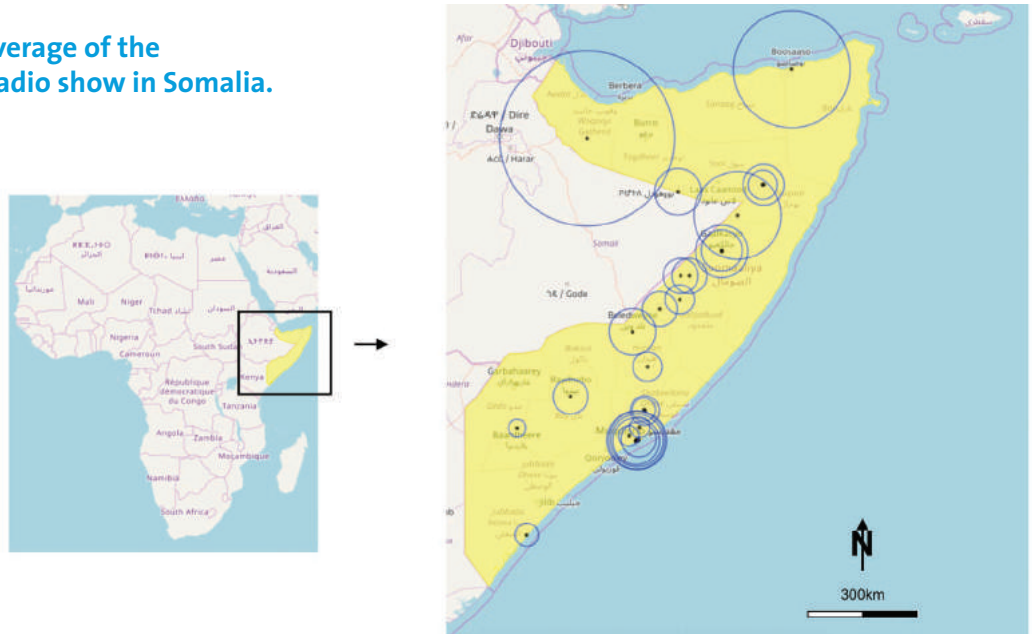
Rate of total reported cases of sexual violence by 10,000 inhabitants in El Salvador. ©De-Arteaga & Dubrawski, 2017
https://mariadearteaga.files.wordpress.com/2016/05/discovery-complex-anomalous1.pdf

PROJECT 4. Two-way radio: Using radio and mobile phones to engage with Somali women and youth (Africa' s Voices Foundation, 2017)

Country	Year	Organization/ Partners	Data sources	SDG
Somalia	Ongoing since 2017	UNICEF Somalia and Africa's Voices Foundation	SMS	Cross-cutting

Through interactive broadcasts on 26 radio stations in Somalia, Africa's Voices Foundation (a start-up from the University of Cambridge) is gathering SMS messages in local dialects from radio audiences. Data has been gathered on food security, maternal and child health, domestic violence, child marriage, FGM, and education, with over 15,000 people participating via SMS from all provinces of Somalia. The objective is to generate insights about cultural beliefs and practices that can be incorporated into UNICEF Somalia's Communications for Development (C4D) programmes and used to map health vulnerability areas. The partners have also developed interactive dashboards that can show the differences in how women and men responded to different questions.

Figure 5:
Radio coverage of the Kalkaal radio show in Somalia.



Radio coverage of the Kalkaal radio show in Somalia. ©Lewis, A. (2017)
http://www.africasvoices.org/case-studies/unicef-somalia-child-protection-gender-equality/

Project 5: Analysing attitudes towards contraception and teenage pregnancy using social media data (UN Global Pulse, 2014)

Country	Year	Organization/ Partners	Data sources	SDG
Uganda	2014	UN Global Pulse, UNFPA, UNICEF	Facebook posts and messages from U-report	5.6.1

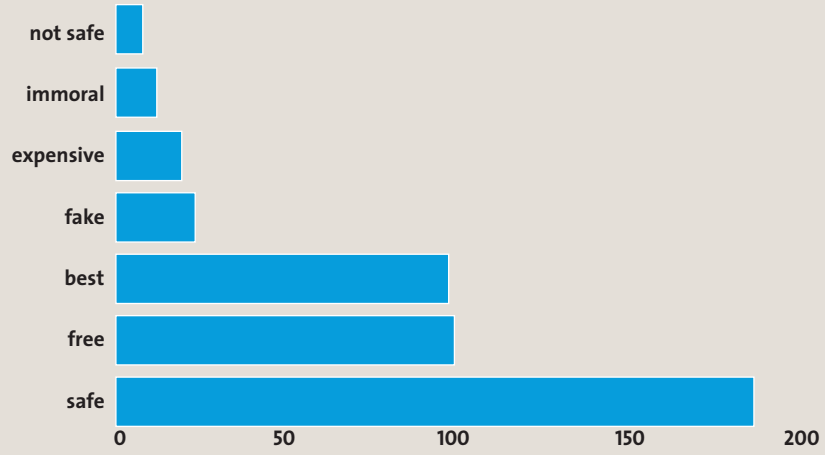
This project explored the use of digital data to understand conversations among young Ugandans related to contraception and teenage pregnancy, and to analyse their perceptions towards different types of contraception. The data used for the study was extracted from public Facebook posts and UNICEF's U-report platform, an SMS-based polling system for Ugandan youth.

The analysis was developed into an interactive dashboard that provides snapshots of the volumes of conversations around different methods of contraception over time, and of popular discussion topics associated with contraception and teen pregnancy.

For example, the graphic below illustrates conversations on the topic of condoms. The words most associated with the topic were 'safe,' 'free' and 'best,' an indication of positive attitudes towards condoms as a contraceptive method.

The study revealed the potential of digital data to help track attitudes and perceptions related to reproductive health and rights in real time, which could support monitoring and evaluation of development programmes.

Figure 6:
Frequency of terms used in connection with condoms



https://www.unglobalpulse.org/projects/UNFPA-social-data

Project 6: Feasibility Study: Identifying trends in discrimination against women in the workplace in social media (UN Global Pulse, 2014)

Country	Year	Organization/partners	Data sources	SDG
Indonesia	2014	UN Global Pulse, the International Labour Organization (ILO) and the Government of Indonesia	Tweets Social media	5.1

This feasibility study analysed online data as a source of real-time signals of discrimination against women in the workplace in Indonesia. Public tweets (over 100,000 tweets over three years) were filtered using keywords related to discrimination, revealing four topics with a considerable volume of discussion: (1) discrimination in job requirements; (2) permission for women to work; (3) perceptions on appropriateness of different types of work for women; and (4) the multiple burdens of working women. While most of the tweets identified corresponded to discriminatory job requirements, weak signals from messages coming from directly affected populations may imply that it is common for women workers to keep silent about their experiences related to discrimination and violence in the workplace (for various reasons, including fear of losing one's job or of facing further discrimination). Website: http://www.unglobalpulse.org/sites/default/files/UNGP_ProjectSeries_Women_Workplace_Discrimination_2014.pdf

Project 7: Monitoring in real time the implementation of HIV mother to child prevention programme (UN Global Pulse, n.d.)

Country	Year	Organization/ partners	Data sources	SDG
Uganda	Ongoing	UN Global Pulse, the Ministry of Health of Uganda, and UNAIDS	Digital live records from patients going for ANC visits, diagnosed with HIV/AIDS and receiving antiretroviral treatments	3.3

As part of Uganda's National Prevention of Mother-to-Child Transmission (PMTCT) programme, UN Global Pulse (through Pulse Lab Kampala) developed an application that tracks in real-time the implementation of Option B+, a programme for prevention of mother-to-child HIV transmission.

This interactive platform tracks the indicators received from health centres across Uganda. The information is analysed using several variables such as: number of patients going for regular antenatal care visits, number of HIV/AIDS cases, and number of patients receiving the option B+ treatment. The dashboard also highlights relevant correlations between treatment drop out rates and different relevant factors. For example, the application analyses the relation between stock-outs of medical supplies and drop out rates.

The application was developed to enable the timely flow of information to address bottlenecks on the rollout of the Option B+ treatment in Ugandan districts.

Pulse Lab Kampala working on a project to monitor in real time the implementation of the HIV mother -to -child prevention programme, Kampala, Uganda. UN Global Pulse, n.d

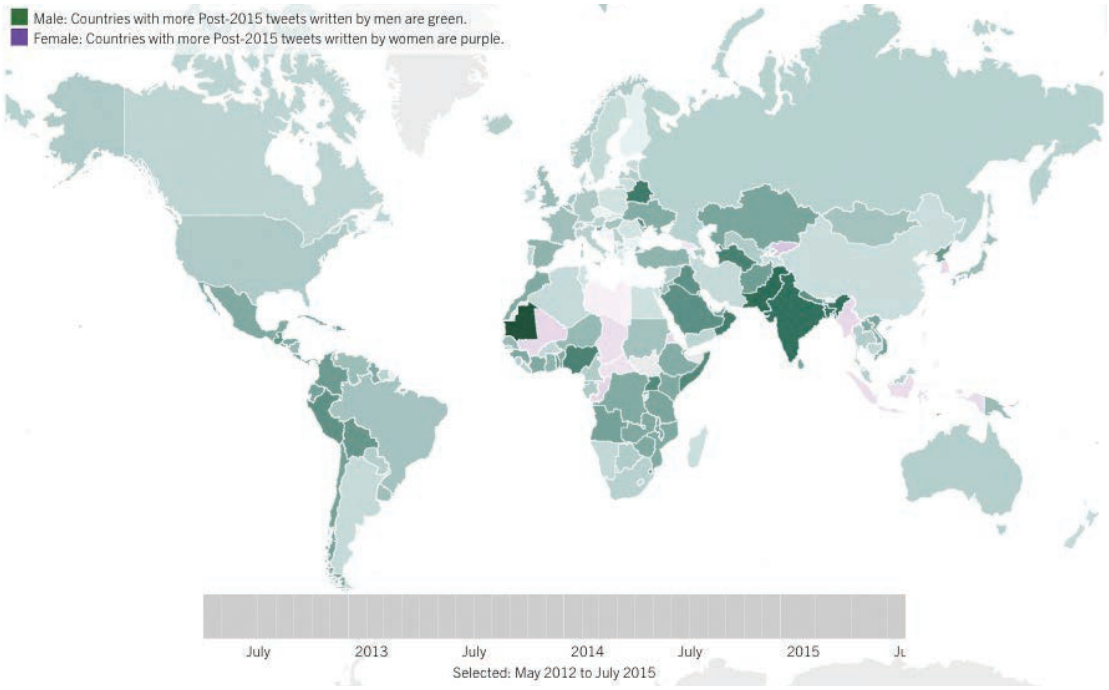
https://www.unglobalpulse.org/projects/monitoring-hiv-mother-child-prevention-programme

PROJECT 8: Sex disaggregation of social media posts (UN Global Pulse, 2016)

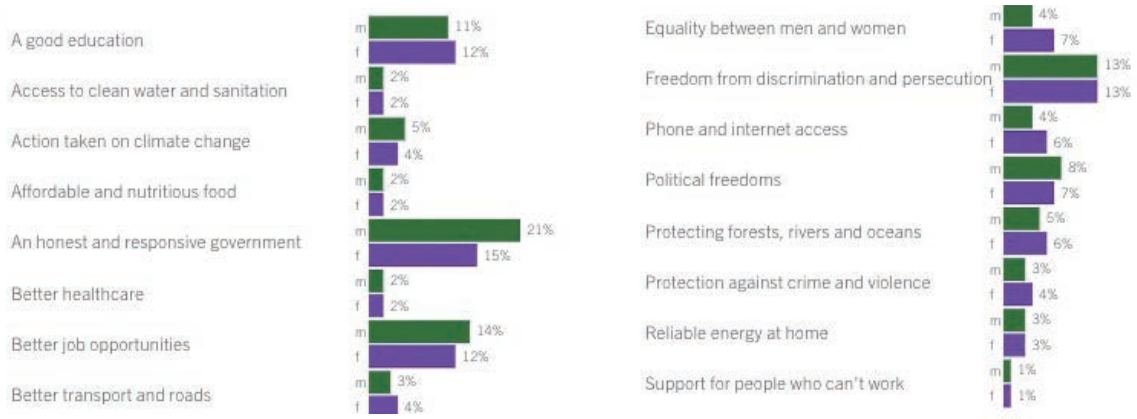
Country	Year	Organization/partners	Data sources	SDG
Global	2016	UN Global Pulse, Data2X and the University of Leiden	Twitter	Cross-cutting

UN Global Pulse collaborated with Data2X and the University of Leiden to develop and prototype a tool to infer the sex of users. Using open source software, the sex-disaggregation tool automates the process of looking up public information from Twittera profiles by analysing user-names from a built-in database of predefined names (from sources such as official statistics) that contain gender information. As in most cases user name is insufficient to discern sex, the tool also analyses profile photos, using face recognition software.

UN Global Pulse used the tool to improve an existing real-time online dashboard showing the volume of tweets around priority topics related to SDGs. The tool was tested on more than 50 million Twitter accounts.



All Countries *(Compare how likely men and women are to tweet about each of the 16 topics.)*



Online dashboard comparing the volume of tweets related to SDGs topics for male and female. UN Global Pulse 2016
http://www.unglobalpulse.org/projects/sex-disaggregation-social-media-posts

PROJECT 9: Using financial transaction data to measure economic resilience to natural disasters (UN Global Pulse, 2016, "Measuring Economic Resilience with Financial Transaction Data")

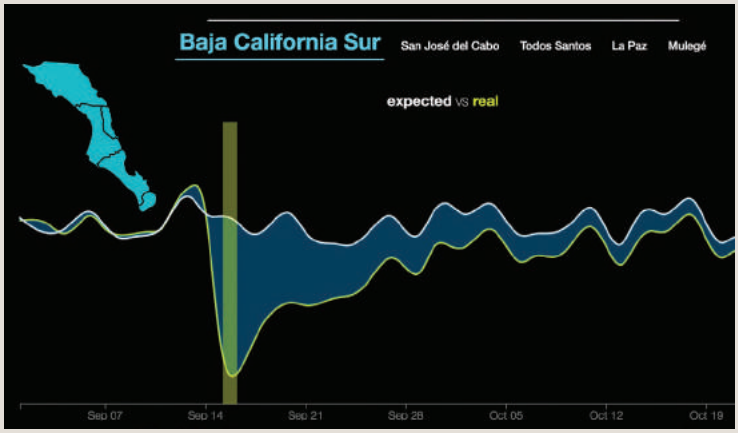
Country	Year	Organization/ Partners	Data sources	SDG
Mexico	2016	UN Global Pulse, BBVA	Financial records	1.5

This project explored how financial transaction data can be analysed to better understand the economic resilience of people affected by natural disasters. 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 in 2014. The project measured daily Point of Sale transactions and ATM withdrawals at high geo-spatial resolution to gain insight into the way people prepare for and recover from disaster.

The study revealed that people spent 50% more than usual on items such as food and gasoline in preparation for the hurricane and that recovery time ranged from 2 to 40 days depending on characteristics such as gender or income. Women increased expenditures in preparation of the hurricane twice as much as men. However, recovery times for women were measured as consistently longer than for men.

Although the study did not focus explicitly on gender, it revealed some important differences in how women and men are impacted by and recover from natural disasters.

Website: https://www.bbvadata.com/odile/



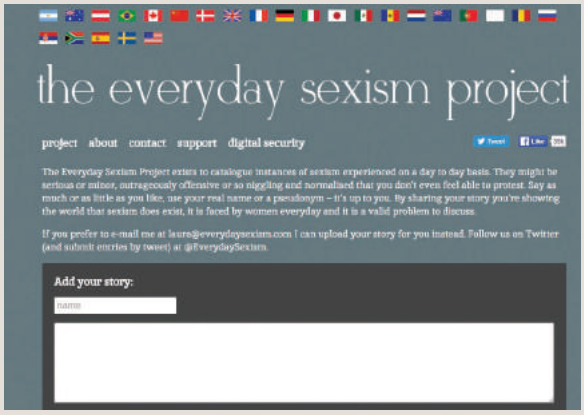
Dashboard to compare the economic activity in Baja California Sur under normal conditions (not affected by the hurricane) with real projections. UN Global Pulse 2016 http://www.unglobalpulse.org/projects/using-financial-transaction-data-measure-economic-resilience-natural-disasters

PROJECT 10: Semantic map of sexism: topic modelling of the Everyday Sexism content5 (Oxford Internet Institute, n.d.)

Country	Year	Organization/ Partners	Data sources	SDG
Global	Ongoing	Oxford Internet Institute	Crowdsourced reports	5.1

A group of researchers from the Oxford Internet Institute are working on a project to create a map of sexism across the globe. The project, which applies a natural language processing approach to analyse data collected from the *Everyday Sexism Project* website, provides an analysis of a large-scale crowdsourced dataset on sexism.

Knowledge gained from this study will advance both the sociological understanding of women's lived experiences of sexism and methodological understandings of the suitability of computational topic modelling for conducting this type of research.

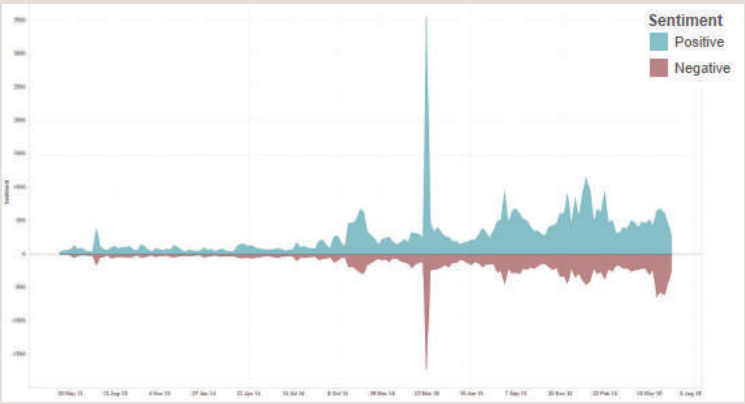


The everyday sexism project website http://everydaysexism.com/

PROJECT 11: Sentiment analysis of reproductive health content of SMS and social media by Kenyan youth (Well Told Story, 2016)

Country	Year	Organization/ Partners	Data sources	SDG
Kenya and Tanzania	2017	Well Told Story and Bill & Melinda Gates Foundation	SMS and social media	5.6.1

Figure 7:
Volume of SMS related with contraception with positive and negative sentiment



Africa's Voices Foundation, 2016

Well Told Story is a socially-motivated research and media company, delivering large-scale, positive, social and behaviour change among 6+ million youth in East Africa through a free, data-driven, multimedia (comic, radio, events, digital media) brand called Shujaaz. As a component of the Shujaaz media, guided, dynamic digital conversations among Shujaaz fans on Facebook, WhatApp and via SMS create a rich dataset that helps inform learning and strategy.

A recent sentiment analysis of the millions of on-going digital media exchanges among hundreds of thousands of Shujaaz fans produced encouraging insights on the power of collective discussions to stimulate massive normative changes around contraception. A longitudinal panel study linked the normative shifts to behaviour change at scale.

Website: *http://www.welltoldstory.com/can-we-see-new-norms-form-new-technology-allows-us-now-only-to-watch-the-process-but-also-to-understand-and-influence-it/*

PROJECT 12: Big data for financial inclusion and poverty mapping (United Nations Global Working Group on Big Data for Official Statistics, n.d.)

Country	Year	Organization/ Partners	Data sources	SDG
Democratic Republic of the Congo, Côte d'Ivoire, Ghana, Uganda, Zambia	Ongoing	The World Bank Group	CDR and mobile financial transaction data	1.4

The aim of this project is to better understand international financial coporation (IFC) client profiles and to determine profiles of likely users. The project ultimately aims to identify lists of individuals who are likely to be interested in financial services from IFC. The IFC will then use this information to develop products and marketing that appeal to potential clients, to ultimately increase the use of financial services by those who were previously unbanked.

Researchers intend to compile call detail records and financial transaction data (for transactions conducted via mobile phones) from both network operators and financial institutions. They will also draw on in-house information about IFC clients.

The project also aims to use household surveys together with big data to assess how financial services affect household expenditure, and to produce poverty maps.

Website: http://www.ifc.org/wps/wcm/connect/e99c35804bc94332956edd7cbf6249b9/Field+Notes+4+-+Big+Data+-+Final.pdf?MOD=AJPERES

PROJECT 13: Big Data and the Cloud – Piloting "eHealth" for community reporting of community performance-based financing in Ghana (United Nations Global Working Group on Big Data for Official Statistics, n.d.)

Country	Year	Organization/ Partners	Data sources	SDG
Ghana	Ongoing	World Bank Group	Android phone-based surveys	3

The Government of Ghana has commenced implementation of a pilot performance-based financing project in 4 regions with particularly poor Maternal Child Health Nutrition outcomes, and to simultaneously assess the effectiveness and cost-effectiveness of this project using a new approach in Ghana which incentivizes community health teams to improve health outcomes. In this system, performance payments are based on achieving key MCHN indicators (quantity and quality), reported and verified quarterly.

The project uses android-based survey tools to assess performance directly from beneficiary communities. The government is considering scaling this project up to a national programme, should it work at the pilot level.

Service delivery at community level in Ghana. © Anyhony Seddoh
Website: https://unstats.un.org/bigdata/inventory/?selectID=WB43

PROJECT 14: #Menstruation Matters... to all of us (Wiederer, 2017).

Country	Year	Organization/partners	Data sources	SDG
Global	2017	UNICEF	Mobile, social media	5.1, 5.6

U-Report is UNICEF's text-message based innovation that amplifies the voices and views of young people in developing countries.

In a poll by UNICEF's Global Innovation Centre and WASH section through U-Report, almost 45,000 young girls and women aged 13-25 from 19 countries, including Cote d'Ivoire, Mozambique, France, Mexico, Senegal, United Kingdom and Uganda, used SMS and Facebook to answer questions about menstruation. Some countries – like Indonesia – also polled boys and men.

The poll showed that in both high-income countries as well as lower-income countries, girls and women face real challenges to attending school or work when menstruating. Among those who responded:

- On average, 13% of girls from low- and middle-income countries missed school or work due to their period. This number was markedly higher in Pakistan and Indonesia at 28%. In high-income countries, this figure was 8%.
- Almost half of girls in Pakistan and 31% of girls in Cote d'Ivoire said that they didn't know about menstruation until they started their first period.
- Overall, 30% of girls and women said they missed school or work because of pain during their periods. Other factors cited were the lack of sanitary materials, lack of toilets, and embarrassment.

Website: http://unicefstories.org/2017/06/01/menstruationmatters-to-all-of-us/

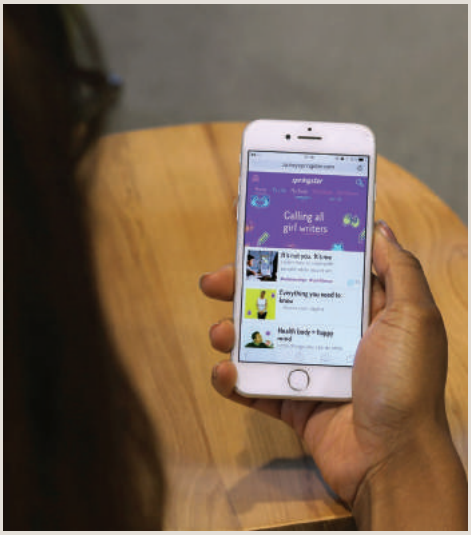
PROJECT 15: Girl Effect's Springster mobile platform

Country	Year	Organization/partners	Data sources	SDG
Global	Ongoing	Girl Effect	Google Analytics, online surveys, comments, social media	Cross cutting

Girl Effect's Springster is a global mobile-first platform targeting 14-16 year old vulnerable girls in an online magazine and discussion format. It aims to build users' confidence and skills online to help them fulfil their potential offline by tackling issues such as how girls can help their families financially, how to make meaningful friendships, and how to navigate puberty. Springster does this in an informative and entertaining way, encouraging girls to come back and discover more.

Researchers from Girl Effect are using a range of methods to better understand how users' engagement and participation relate to changes in girls' lives before and after they read content on Springster and engage in conversations with other users. By using several data sources, including Google Analytics, comment analysis, online surveys and social media analytics, researchers are combining big data with traditional approaches to learn more about the intersection of digital behaviour, participation in community spaces and engagement with content and how participation in this online safe space influences change.

Website: http://www.girleffect.org/what-girls-need/articles/2017-insights-inspiration/our-vision-to-enable-100-million-girls-to-find-their-voice-online/



Girl accessing Springster platform on a mobile phone.

SUMMARY

Big data has an essential and vital role to play in achieving the SDGs, including SDG 5 – "Achieve gender equality and empower all women and girls". Sources such as social media trails, call data records, radio data, satellite imagery and other "digital exhaust", both alone and combined with traditional data sources can shed light on the lives of women and girls.

However, we should also be cautious of the limitations of big data. There remain concerns with its use, including that careless interpretation of big data might lead to disproportionate representation of those who are capable of producing these trails (for example, it is possible to fail to include those who do not have access to technology, are not online or prefer not to engage). Data use also requires adequate normative frameworks, stronger implementation mechanisms for protection of privacy, including tools for mitigating risks of harms to not only individuals but also groups of individuals. Access to big data may also be a challenge, as it may be expensive and requires technical knowledge to retrieve, handle and store. In addition, analytical skills are expensive – in some cases, the data itself might be free, but deriving insights may have high associated costs, especially for budget-constrained organizations.

Bearing these concerns in mind, and with appropriate applications of big data to certain populations (for example, using radio data for offline communities), big data analytics hold huge potential for UN Women as a fast-evolving source of knowledge that can provide valuable insights about women and girls and for women and girls. The contribution in particular to helping monitor the SDGs is important, as demonstrated by the compendium of projects in this report.

REFERENCES

A/RES/45/95 (1990, 14 December). Guidelines for the regulation of computerized personal data files. Available at http://www.un.org/documents/ga/res/45/a45r095.htm

Antin K., Byrne R., Geber T., van Geffen S., Hoffmann J., Jayaram M., Khan M., Lee T., Weinberg F., Wilson C., Rühling B., Rahman Z., Simeoni C. (2014). Shooting your hard drive into space and other ways to practise responsible development. Available at https://responsibledata.io/wp-content/uploads/2014/10/responsible-development-data-book.pdf

Africa's Voices Foundation (2016). UNICEF Somalia. Available at http://www.africasvoices.org/case-studies/unicef-somalia/.

Bellagio (2014). Big data and positive social change in the developing world: A white paper for practitioners and researchers. Oxford: Oxford Internet Institute.

Bosco, C., Alegana, V., Bird, T., Pezzulo, C., Bengtsson, L., Sorichetta, A., ... & Wetter, E. (2017). Exploring the high-resolution mapping of gender-disaggregated development indicators. Journal of The Royal Society Interface, 14(129).

Citro, C. (April 2015). From multiple modes for surveys to multiple data sources for estimates. Keynote address at UNECE workshop on data collection and communication. Washington DC. Available at: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2015/mtg1/S0-1_USA__Citro__P.pdf

Couper, M. P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. Survey Research Methods, 7(3): 145-156.

De-Arteaga, M., & Dubrawski, A. (2017). Discovery of complex anomalous patterns of sexual violence in El Salvador. Available at Zenodo at http://doi.org/10.5281/zenodo.571551.

Groves, R. M. (2011). The three eras of survey research. Public Opinion Quarterly, 75(5): 861-871.

Hilbert, M. (2013). Big data for development: From information- to knowledge societies. Available at SSRN: https://ssrn.com/abstract=2205145

Hilbert, M. (2016). Big data for development: A review of promises and challenges. Development Policy Review, 34: 135–174.

Huber, B.R. (2014, 20 November). 'Mosaic Effect' paints vivid pictures of tech users' lives, Felten tells privacy board. Available at http://wws.princeton.edu/news-and-events/news/item/mosaic-effect-paints-vivid-pictures-tech-users-lives-felten-tells-privacy

ITU (2017). ICT Facts and Figures 2017. Available at http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf

Laney, D. (2001, 6 February). 3-D Data Management: Controlling Data Volume, Velocity and Variety. META Group Res Note 6. 6. Available at http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

Letouzé E. (2015). Big data and development: An overview. Pop Alliance White Paper Series. Data-Pop Alliance, World Bank Group, Harvard Humanitarian Initiative. Available at http://datapopalliance.org/wp-content/uploads/2015/12/Big-Data-Dev-Overview.pdf

Metcalf, Jacob, Keller, E. F. and Boyd, D. (2016). *Perspectives on Big Data, Ethics, and Society*. Council for Big Data, Ethics, and Society. Accessed December 7, 2017. http://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/

O'Neill, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy.* New York: Crown Publishing Group.

Organisation for Economic Co-operation and Development. (2013). New data for understanding the human condition: International perspectives. OECD Global Science Forum Report on Data and Research Infrastructure. Available at http://www.oecd.org/sti/sci-tech/new-data-for-understanding-the-human-condition.pdf

Oxford Internet Institute. (n.d.) Semantic map of sexism: Topic modelling of the everyday sexism content. Available at https://www.oii.ox.ac.uk/research/projects/semantic-map-of-sexism/

UN Global Pulse (n.d.). Data privacy and data protection principles. Available at http://www.unglobalpulse.org/privacy-and-data-protection-principles

UN Global Pulse. (n.d.). Monitoring in real time the implementation of HIV mother-to-child prevention programme. Available at http://unglobalpulse.org/projects/monitoring-hiv-mother-child-prevention-programme

UN Global Pulse (n.d.). Privacy tools. Risks, harms, benefits assessment. Available at http://www.unglobalpulse.org/privacy/tools

UN Global Pulse. (2012). Big data for development: Opportunities & challenges. Available at http://www.unglobalpulse.org/projects/BigDataforDevelopment

UN Global Pulse. (2013). Big data for development: A primer. Available at http://www. unglobalpulse.org/bigdataprimer

UN Global Pulse. (2014). Analyzing attitudes towards contraception & teenage pregnancy using social data. Global Pulse Project Series, no. 8. Available at http://www.unglobalpulse.org/sites/default/files/UNGP_ProjectSeries_Family_Planning_2014_0.pdf

UN Global Pulse. (2014). Feasibility study: Identifying trends in discrimination against women in the workplace in social media. Global Pulse Project Series no. 11. Available at http://www.unglobalpulse.org/projects/indonesia-women-employment

UN Global Pulse. (2015). Mapping the risk-utility landscape of mobile data for sustainable development & humanitarian action. Global Pulse Project Series no. 18. Available at http://www.unglobalpulse.org/projects/mobile-data-privacy

UN Global Pulse (2016). Making Ugandan community radio machine-readable using speech recognition technology. Global Pulse Tool Series 1, 2016. Available at http://unglobalpulse.org/resource-library/projects-and-tools. Website: http://radio.unglobalpulse.net/uganda/

UN Global Pulse (2016). Measuring Economic Resilience with Financial Transaction Data. Global Pulse Project Series no. 24. Available at http://unglobalpulse.org/sites/default/files/2016%20BBVA%20Project%20Brief.pdf

UN Global Pulse. (2016). Sex disaggregation of social media posts. Global Pulse Tool Series no. 3. Available at http://data2x.org/wp-content/uploads/2017/03/Big-Data-and-the-Well-Being-of-Women-and-Girls.pdf.

United Nations, Department of Public Information. (2015, 10 October). *Implementing 2030 agenda requires 'data revolution' to address gaps in collection, quality, Secretary-General tells development committee, urging new partnerships*. SG/SM/17208-ENV/DEV/1589-STAT/454. Available at http://www.un.org/press/en/2015/sgsm17208.doc.htm

United Nations Development Group. (March 2017). Guidance Note on Big Data for SDGs: Data Privacy, Data Protection and Data Ethics. Available at https://undg.org/document/undg-guidance-note-on-big-data-for-achievement-of-the-2030-agenda-data-privacy-ethics-and-protection/

United Nations Global Working Group on Big Data for Official Statistics. (n.d.). Big Data Project Inventory. Available at https://unstats.un.org/bigdata/inventory/

UN Women (2015, September). Monitoring gender equality and the empowerment of women and girls in the 2030 agenda for sustainable development: Opportunities and challenges. New York. Available at http://library.pcw.gov.ph/sites/default/files/monitoring%20ge%20and%20empowerment%20of%20women%20and%20girls.pdf

Vaitla, B. et al. (April 2017). Big Data and the Well-Being of Women and Girls: Applications on the Social Scientific Frontier. Data 2X. Available at: http://www.data2x.org/wp-content/uploads/2017/11/Big-Data-and-the-Well-Being-of-Women-and-Girls_.pdf

Wijesekera, S. (2017, May 27). #MenstruationMatters… to all of us. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/entry/menstruationmatters-to-all-of-us_us_59298550e4b07d848fdc0415

UN WOMEN IS THE UN ORGANIZATION DEDICATED TO GENDER EQUALITY AND THE EMPOWERMENT OF WOMEN. A GLOBAL CHAMPION FOR WOMEN AND GIRLS, UN WOMEN WAS ESTABLISHED TO ACCELERATE PROGRESS ON MEETING THEIR NEEDS WORLDWIDE.

UN Women supports UN Member States as they set global standards for achieving gender equality, and works with governments and civil society to design laws, policies, programmes and services needed to implement these standards. It stands behind women's equal participation in all aspects of life, focusing on five priority areas: increasing women's leadership and participation; ending violence against women; engaging women in all aspects of peace and security processes; enhancing women's economic empowerment; and making gender equality central to national development planning and budgeting. UN Women also coordinates and promotes the UN system's work in advancing gender equality.

UN WOMEN

Planet 50-50 by 2030
Step It Up for Gender Equality

220 East 42nd Street
New York, New York 10017, USA
Tel: 646-781-4400
Fax: 646-781-4444

www.unwomen.org
www.facebook.com/unwomen
www.twitter.com/un_women
www.youtube.com/unwomen
www.flickr.com/unwomen