Last updated: September 13, 2024

Gender Sensitive Data Collection

Dr. Diana West
Emily Adams

Earth System Science Center, University of Alabama in Huntsville
SERVIR Science Coordination Office at NASA Marshall Space Flight Center

This slide deck is associated with a series of training materials developed by the SERVIR Science Coordination office to support Gender Equity and Social Inclusion into Applied Earth Sciences. Materials are available at the SERVIR / GESI-eo-training GitHub repository (link)

DOI: 10.5281/zenodo.13769571

Gender data gap is REAL

From cars that are **71% less safe for women than men** (because they've been designed using a 50th-percentile male dummy), t**o voice-recognition technology that is 70% less likely to accurately understand women than men** (because many algorithms are trained on 70% male data sets), to medication that doesn't work when a woman is on her period (**because women (or even women's cells) weren't included in the clinical trials**), we are living in a world that has been designed for men because for the most part, we haven't been collecting data on women.

the gender data gap is **not the product of a conspiracy by a group of misogynistic data scientists** - when **we** say *human,* 9 times out of 10, we mean men.

- Caroline Criado Perez

Gender data gap is REAL - algorithms, AI, machine learning, etc

When tech entrepreneur David Heinemeier Hansson complained to Apple that his wife was given a limit for her credit card **20 times lower than his**, despite having a higher credit score, he was informed by workers at the company that it was not discrimination, it was **“just the algorithm.”**

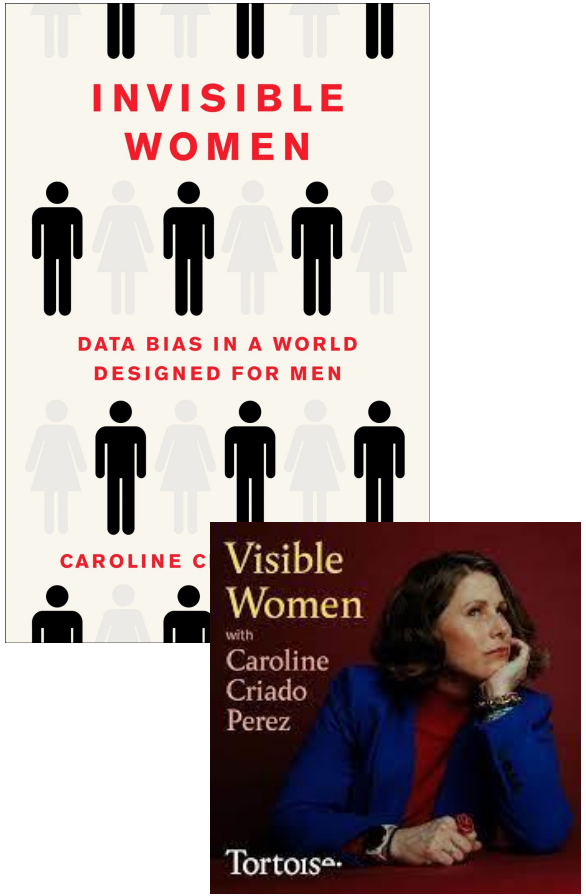
Algorithms are only as good as the **data we feed them**
Algorithms **amplify** our biases back to us.

One University of Washington study found that when an algorithm was trained on an image data set where **pictures of cooking were 33% more likely to feature women than men**, the algorithm increased the disparity to **68%**. That is, **pictures of men were labeled as female just because they were in front of a stove**.

How might this affect:

- Hiring algorithms?
- Medical diagnoses?
- Our work?

Gender data gap is REAL



"The gender data gap and its default male origins have been disadvantaging women for millennia, but in a world where we increasingly outsource our decision-making to algorithms trained on data with a great big hole in it, this problem is set to get a lot more serious very quickly. And if we don't choose to correct the mistakes of the past now, we will blunder into a future where we have literally coded them in."

- Caroline Criado Perez

Only 24 percent of available gender-specific indicators have data from the last 10 years.

World Bank Gender Data Portal

https://genderdata.worldbank.org

What groups are marginalized?

- Class (socio-economic)
- Disabilities
-

What kinds of problems do women face?

- Capability bias
- Medical bias
- Maternity bias
- Wage gap

What kinds of problems do women face?

- Inequalities of access: women often face barriers in terms of finances, knowledge, skills, and mobility
- Structural barriers: women may be disempowered by the regulatory landscapes; i.e. land rights, inheritance, home ownership structures, etc.
- Digital divide: women access digital tools and platforms differently from men and may not have the same access – example: table banking in Africa
- Sociocultural norms and practices: an intersectional approach that considers gender, age, class, and socioeconomic status is important to situate gender inequities

What kinds of problems do women face?

- Inequalities of access: women often face barriers in terms of finances, knowledge, skills, and mobility
- Structural barriers: women may be disempowered by the regulatory landscapes: i.e. land rights, inheritance, home ownership structures, etc.
- Digital divide: women access digital tools and platforms differently from men and may not have the same access
- Sociocultural norms and practices: an intersectional approach consider gender, age, class, and socioeconomic status is important to situate gender inequities

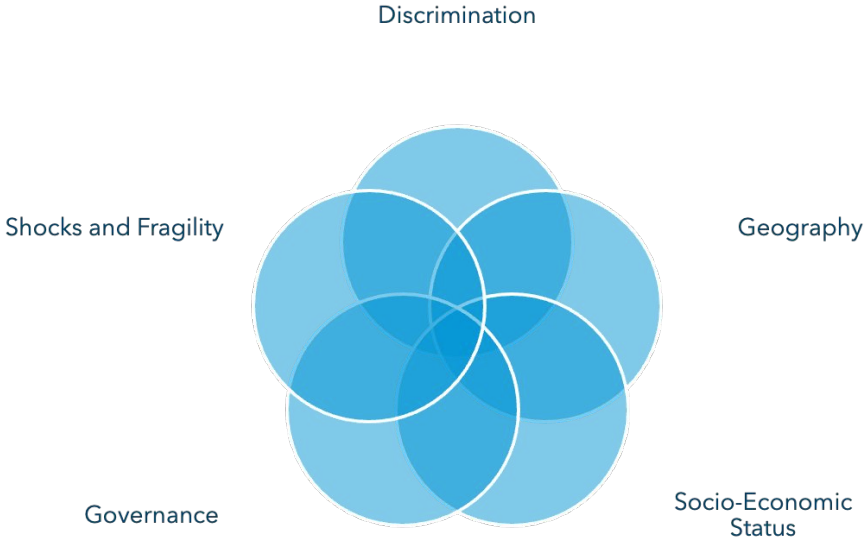
To evaluate these issues comprehensively, a full Gender Analysis targeted at answering project research questions must be conducted. However, we can also start thinking about gender-sensitive data collection before a full analysis can take place.

What do we need to think about if we want to design inclusive services?

- Regional specific cases (HKH, men are moving out of the mountains regions into the cities
- Anyone who is going to benefit from using the service and thinking about if there are people who won't use it why not
    - Use case scenarios - breaking down different users
- How people access information
- How we are distributing information about CB (sectors, work places, gender specific places)
- Vulnerability of different populations and how we define vulnerability - RS gives us risk but not exposure

Intersectional Framework

According to the UN: to understand what groups are excluded, why, and how to shape effective responses, five key factors should be assessed: **discrimination, geography, socio-economic status, governance and shocks and fragility.**



Intersectional Framework

All key groups must be identified and consulted

Ideally, multiple relevant identities should be represented in programme membership as well as in leadership

Organizational and programme inclusion policy should be based on incorporating as many diverse stakeholders as possible

Recognize the voices of marginalized identities and elevate these voices so they can be better heard

Data collection and analysis tools should be designed with multiple identities in mind; the more dimensions along which data can be disaggregated the more clear the picture will be

While a programme, project, or policy may focus on a particular disparity, being aware of related intersectional disparities and displaying solidarity with those movements will strengthen your programme and impact

GENDER // PRO
THE GEORGE WASHINGTON UNIVERSITY

Intersectional Framework

All key groups must be identified and consulted

Ideally, multiple relevant identities should be represented in programme membership as well as in leadership

Organizational and programme inclusion policy should be based on incorporating as many diverse stakeholders as possible

Recognize the voices of marginalized identities and elevate these voices so they can be better heard

Data collection and analysis tools should be designed with multiple identities in mind; the more dimensions along which data can be disaggregated the more clear the picture will be

While a programme, project, or policy may focus on a particular disparity, being aware of related intersectional disparities and displaying solidarity with those movements will strengthen your programme and impact

Service planning toolkit: at what stage does SERVIR do this?

GENDER // PRO
THE GEORGE WASHINGTON UNIVERSITY

Intersectional Framework

All key groups must be identified and consulted

Ideally, multiple relevant identities should be represented in programme membership as well as in leadership

Organizational and programme inclusion policy should be based on incorporating as many diverse stakeholders as possible

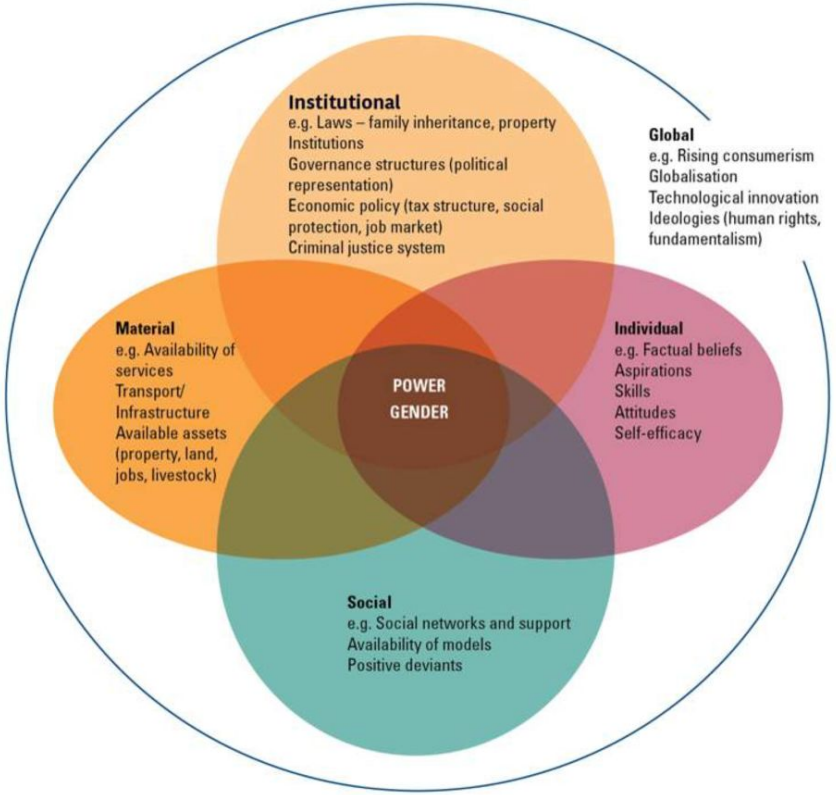
Recognize the voices of marginalized identities and elevate these voices so they can be better heard

Data collection and analysis tools should be designed with multiple identities in mind; the more dimensions along which data can be disaggregated the more clear the picture will be

While a programme, project, or policy may focus on a particular disparity, being aware of related intersectional disparities and displaying solidarity with those movements will strengthen your programme and impact

GENDER // PRO
THE GEORGE WASHINGTON UNIVERSITY

Intersectional Framework: Gender



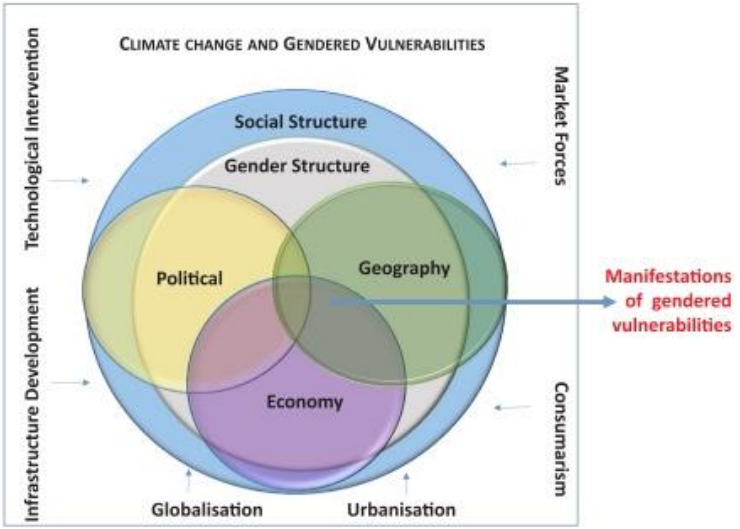


Fig. 2. Conceptual framework on Climate Change and Gendered Vulnerabilities.

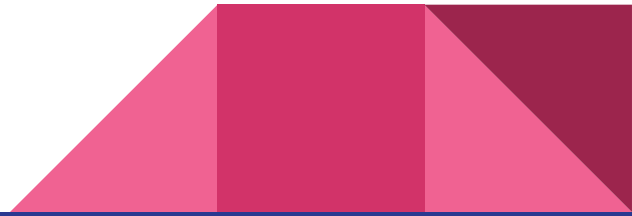
ICIMOD

Gender Statistics

Gender statistics reflect differences and inequalities in the situation of women and men in all areas of life.

1. **Sex-disaggregated** data:
    - data that are tabulated and presented separately for women and men, or girls and boys.

2. **Intersectionalities** framework:
    - Further disaggregation (for example, by age, level of education, place of residence, income, ethnicity, religion, disability status and sexual orientation) provides an even clearer picture of the relative status of women and men and is useful for identifying
    - vulnerable groups.



Gender Statistics: quiz

True or false?

If all of our data are sex-disaggregated, we already have gender statistics

Gender Statistics: quiz

True or false?

If all of our data are sex-disaggregated, we already have gender statistics

FALSE

The production of sex-disaggregated data is only one component of gender statistics. To be more effective, your organization should also produce statistics relevant to key gender issues in the country and that might affect only one sex (for example, data on the prevalence of early marriage among adolescent girls).

Gender Statistics: quiz

True or false?

If all of our data are sex-disaggregat[illegible]y have gender statistics

FALSE

The production of sex-disaggre[illegible]er statistics. To be more effective, your organization should[illegible]ey gender issues in the country and that might affect only one s[illegible]the prevalence of early marriage among adolescent girls).

Don’t let the GREAT be the enemy of the GOOD! Collect sex-disaggregate data whenever possible!!!!

Gender Statistics: quiz

True or false?

Collecting sex-disaggregated data is an extra lift for the Hubs and can be costly and time-consuming

Gender Statistics: quiz

True or false?

Collecting sex-disaggregated data is an extra lift for the Hubs and can be costly and time-consuming

FALSE!

Generally, producing gender statistics using existing instruments does not involve much additional cost. Often an extra question or a column indicating sex can be added to an existing survey. Furthermore, if a gender perspective is properly integrated and addressed from the early stages of service planning, additional or unexpected costs can be avoided.

Gender Statistics: quiz

True or false?

Collecting sex-disaggregated dat[illegible]tra lift for the Hubs and can be costly and time-consum[illegible]

But wait! This sounds a bit simplistic. You can have gender bias even if data is not "gender blind"!

FALSE!

Generally, producing gender stati[illegible]not involve much additional cost. Often an extra question or a column indicating sex[illegible]existing survey. Furthermore, if a gender perspective is properly integrated and ad[illegible]ssed from the early stages of service planning, additional or unexpected costs can be avo[illegible]

Gender bias in data collection

EXAMPLE:

A survey is collecting data on employment and labor in Kenya. They sex-disaggregated the data and discover that most rural women in the survey are listed as “unemployed”. Therefore, they conclude that only men generate income.

What is the problem?

What is the solution?

Gender bias in data collection

EXAMPLE:

A survey is designed to collect information from households in Brazil about how they make decisions about what crops to plant. You hold a focus group discussion with the farmers in the evening to discuss the impacts of climate shocks on their crop decisions. Only men show up to your discussions and you capture the data they provide.

What is the problem?

What is the solution?

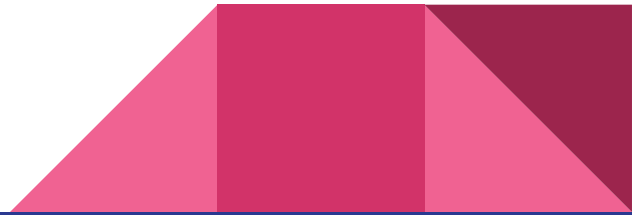
Gender bias in data collection

EXAMPLE:

An online survey was created to gather information about air quality and how it affects the users' mobility. SurveyMonkey was used to collect the data from 300,000 urban dwellers in Bangladesh. The findings were used to inform an air quality alert system.

What is the problem?

What is the solution?



Gender bias in data collection

EXAMPLE:

Emily has been tasked with supporting the design of a crop monitoring program in Botswana. RCMRD is taking in the process of expanding the existing crop monitor to the region and is engaged with the Ministry of Agriculture for this product. How can SCO best support the hub in applying a gender lens and an appropriate intersectional framework in designing this monitor?

Gender bias in data collection

EXAMPLE:

Emily has been tasked with supporting the design of a crop monitoring program in Botswana. RCMRD is taking in the process of expanding the existing crop monitor to the region and is engaged with the Ministry of Agriculture to conduct this activity. How can SCO best support the hub in applying a gender lens and an appropriate intersectional framework in designing this monitor?

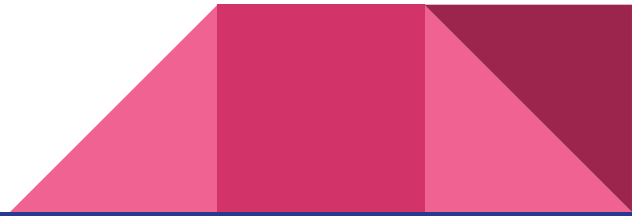
Emily's recommendations to RCMRD:

1. Ensure that the stakeholder engagement and consultations are inclusive, aware of intersectionalities, and are collecting gender statistics
2. Collect the in situ data from intersectional user groups for validation
3. Conduct a **gender analysis** to assess: gendered risks from climate shocks, gendered agricultural practices, and if gender gaps in the data impact existing crop monitor design

Think about your projects

What data collection opportunities exist?

What are your projects presently missing?



Funding for this work was provided by USAID's Gender Equity and Equality Action (GEEA) Fund and was implemented under the SERVIR program, a joint NASA- and USAID-led initiative. The cooperative agreement 80MSFC22N0004 between NASA and UAH facilitated this work.