Big Data & Gender

An Overview of the Global Landscape

Bapu Vaitla
January 26, 2022
How do we define big data?

- A form of digital data
- Information finely disaggregated over space and across time
- When used in concert with traditional data
  - Allows nuanced, effective response to fast-moving situations—pandemics, economic shocks, natural disasters
  - Allows micro-targeting of vulnerable and marginalized groups
Gendered threats to physical safety

Safety incident records → Geospatial analysis → High-resolution surface

“Street Harassment and Educational Choices: A Geospatial Analysis,” Girija Borker
The promise of big data

Big data offers unique insights into the lives of women and girls

- Sex-disaggregated individual data; passive traces of behavior

Gender-sensitive big data is ready to scale and integrate with traditional data

- Many sources already at scale; traditional data always needed
The perils of big data

We must identify and correct bias in big datasets

- Gendered access to tech, biased algorithms, cultural bias

Privacy and security issues must be confronted head-on and immediately

- Specific risks to women/girls, but safe data sharing is possible
Representation is critical

Women and girls must be central to data governance

• Avoid the past failures of technology
• Representation requires participation: who has the power?
Mobile phone & geospatial data

Percentage of women engaged in agricultural livelihoods

Province level, based on survey data

1 sq km resolution, based on geospatial and CDR correlates

“Towards high-resolution sex-disaggregated dynamic mapping,” Flowminder Foundation
The current landscape of big gender data

The current environment

COVID-19/recession is catalyzing groundbreaking gender research

Where we are progressing

Mobility, health, and social norms (among other topics) are increasingly well-studied through a combo of big & traditional data
The current landscape of big gender data

Where we are lagging

Gaps remain around economic opportunity, human security, public participation

Addressing bias

Gender-representative samples still a challenge, but some progress using big data
The current landscape of big gender data

Emerging opportunities
Researchers are finding ways to reuse data from large private companies safely; very useful for policymaking

Practical change
The emerging model of data collaboratives holds great promise