“Big Data” for Gender:
Expanding Horizons and Recognizing Limitations

Emily Courey Pryor, Executive Director, Data2X

Global Forum on Gender Statistics,
Tokyo, Japan, November 16th, 2018

data2x

Data2X: What We Do

Data2X works to increase the availability and use of quality gender data.

- We build the case and mobilize action for gender data.
- We strengthen gender data production and use.

data2x

Our Worldview:
From Gender Data to Smarter Decisions

Produce more and better gender data

Analyze and derive actionable insights from that data

Use that data to drive smarter, gender-equitable decisions

data2x

Defining Big Data

What is Big Data?

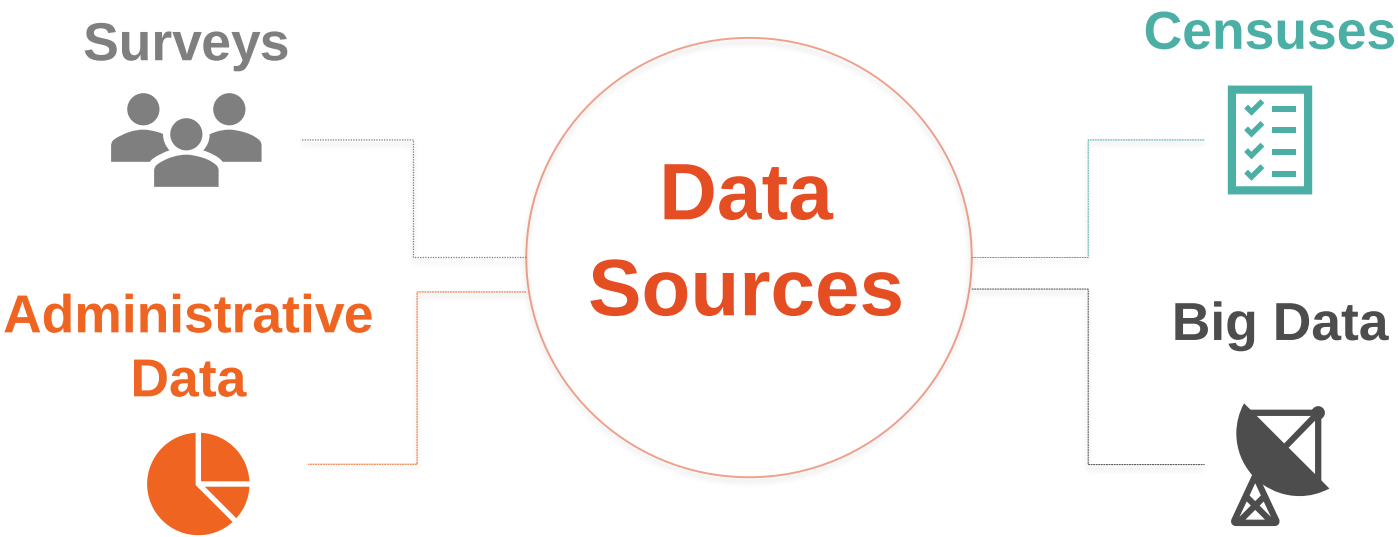
Large amounts of data

collected passively from digital interactions

with great variety and a high rate of velocity.

data2x

Big Data in the Data Ecosystem



data2x

Big Data: Risks and Considerations

- **Privacy**
- **Bias and access: Who does big data leave behind?**
  - Consider access, affordability, literacy, and other barriers
- **Country context: One size doesn't fit all**
- **Ground truth**
  - Digital data should enhance, not replace, information gathered from traditional sources like household surveys and censuses

data2x

Our mission in Big Data

- **Make women visible** within Big Data: avoid bias from the outset
- **Remove the risk:** figure out what works - and what doesn't!
- **Bridge communities:** Not 'traditional' vs. 'big' – instead, collaboration for greater & sustained impact

data2x

Big Data and the Well-Being of Women and Girls

data2x

Geospatial data can provide highly detailed, frequently updated information about the lives of women and girls.

#BigData4Gender
bit.ly/BigData4Gender-Report

data2x

Data gathered from social media can serve as an accurate and important source of information about the mental health of women and girls.

#BigData4Gender
bit.ly/BigData4Gender-Report

data2x

Cell phone and credit card records can reveal:

physical mobility

economic status

social network diversity

offering insights into the needs and priorities of women and girls.

#BigData4Gender
bit.ly/BigData4Gender-Report

data2x

Addressing Data Gaps: Big Data for Gender Challenge

10 projects representing **29** researchers from **20** global institutions across **8** countries

- **Gender-Differentiated Credit Scoring Algorithms Using Call Detail Records and Machine Learning**
  Leads: UC Berkeley, The World Bank
  Methods: Call detail records; machine learning algorithms

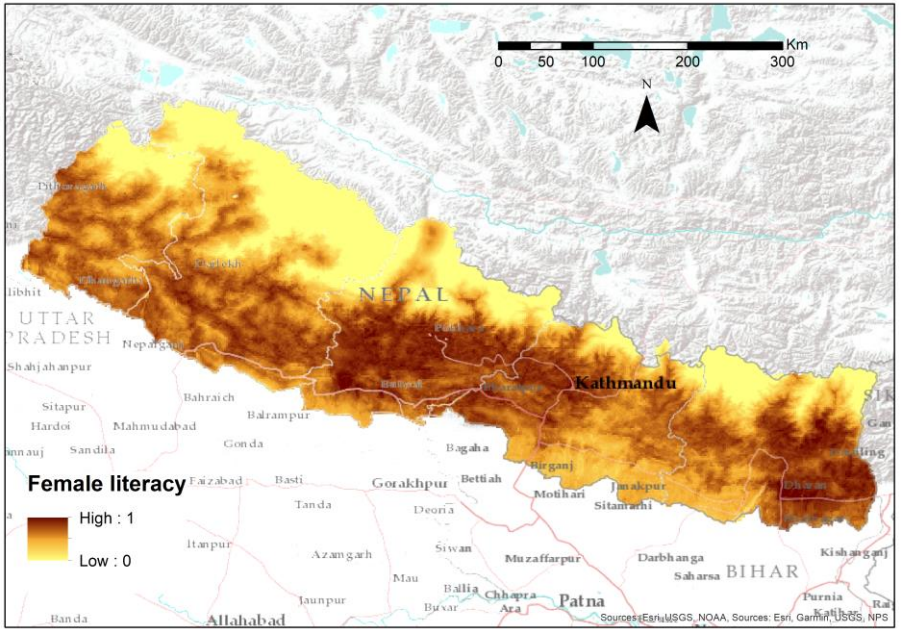
- **Women in the Gig Economy: A Data Gap with Implications for Informal Work, Time Use and Poverty**
  Leads: Overseas Development Institute, Ulula, Data-Pop Alliance
  Method: Mobile phone-based longitudinal survey

- **Safety First: Perceived Risk of Street Harassment and Educational Choices of Women**
  Lead: Girija Borker, PhD, Brown University
  Methods: Student surveys, Google Maps travel route data, mobile application data

data2x

Dynamic Wellbeing Mapping: Nepal

- **Aim**: Build dynamic maps of sex-disaggregated vulnerability indicators, e.g. population density, literacy, stunting, school enrollment
- **Data Sources**: Household surveys, GIS data, CDR data + phone surveys (ground truthing),
- **Method**: Predict sex-disaggregated indicators of wellbeing using household surveys combined with GIS data; + CDR data for mobility and migration mapping.



data2x

What we have learned so far…

- **Data access** is a common issue
- Forging **multi-stakeholder partnerships** is likely to yield better impact
- There are potentially **wide ranging applications** for Big Data to answer gender-relevant research questions
- There is appetite for a **community of practice** around Big Data and Gender
- Addressing **representativeness** is a key issue- ground truthing is crucial

data2x

Where do we go from here?

- 11 pilots will be complete and results shared in 2019
- Key questions & opportunities:
  - More methodological work?
  - Bringing select projects to scale?
  - Mobilizing more people & resources?
  - Demonstration of impact?

data2x

@Data2X

www.data2x.org/big-data-challenge-awards/