WEBINAR SERIES

Data for Sustainable Development

Session 4:
Big Data and the global gender gap: The promises and perils of digital information

Featuring:
Rebecca Furst-Nichols
Deputy Director of Data2X

Bapu Vaitla
Data2X Fellow

Date: Tuesday, 20 March, 2018
Time: 4:00 PM CET
#SDTalks
Big Data and the Global Gender Gap: The Promises and Perils of Digital Information

Rebecca Furst-Nichols
Deputy Director, Data2X

Bapu Vaitla
Fellow, Data2X

UN System Staff College
Knowledge Centre for Sustainable Development
SD Talks Special Series on Data for Sustainable Development
March 20th, 2018
Data2X: Our Mission

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

• **Promote** expanded, unbiased, and innovative gender data collection.
• **Identify** gender data gaps for priority attention.
• **Lead** partnerships to close gender data gaps.
• **Advocate** for better gender data and its use in decision making.
• **Educate** about how to improve and use gender data to improve lives and outcomes for all.
What is Gender Data?

• Data that is *disaggregated by sex*, such as primary school enrollment rates for girls and boys.

• Data that *pertains specifically to girls and women* as a result of biology or their social roles, such as maternal mortality rates or unpaid care work.
Where are the Gender Data gaps?

There are over 28 identified gaps in gender data based on need, population coverage, and policy relevance across five domains:
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.
Partnerships: Big Data for Gender

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

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

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

#BigData4Gender
bit.ly/BigData4Gender-Report

#BigData4Gender
bit.ly/BigData4Gender-Report
1. Obtain survey point data on well-being (e.g., DHS)
2. Obtain geospatial data at same locations: population density, infrastructure, vegetation type (satellite, etc.)
3. Correlate the two sets of info
4. Predict landscape of data on well-being
1. Obtain survey point data on well-being (e.g., DHS)
2. Obtain geospatial data at same locations: population density, infrastructure, vegetation type (satellite, etc.)
3. Correlate the two sets of info
4. Predict landscape of data on well-being
Credit Card/Cell Phone Data: Economic Lifestyles
Di Clemente, Gonzalez, et al. (MIT)

- Anonymized credit card data from 150k users, with age, sex, location info
- Subset: cell phone data
- Portraits of economic lifestyles: mobility, access, preferences
- Could illuminate how women cope with economic/environmental shocks & stresses
Credit Card/Cell Phone Data: Economic Lifestyles
Di Clemente, Gonzalez, et al. (MIT)

- Anonymized credit card data from 150k users, with age, sex, location info
- Subset: cell phone data
- Portraits of economic lifestyles: mobility, access, preferences
- Could illuminate how women cope with economic/environmental shocks & stresses
Cell Phone Data: Epidemiology
Wesolowski et al. (Carnegie Mellon, Harvard, etc.)

- Locations of ~15m cell phone subscribers; travel maps
- Malaria prevalence map based on existing data
- Source/sink maps based on human movement and parasite prevalence
- Allows precisely targeted interventions, in time and space

Fig. 3. Sources and sinks of people and parasites. Kernel density maps showing ranked sources (red) and sinks (blue) of human travel and total parasite movement in Kenya, where each settlement was designated as a relative source or sink based on yearly estimates. (A) Travel sources and sinks. (B) Parasite sources and sinks.
Locations of ~15m cell phone subscribers; travel maps
- Malaria prevalence map based on existing data
- Source/sink maps based on human movement and parasite prevalence
- Allows precisely targeted interventions, in time and space
Cell Phone Data: Epidemiology
Wesolowski et al. (Carnegie Mellon, Harvard, etc.)

- Locations of ~15m cell phone subscribers; travel maps
- Malaria prevalence map based on existing data
- Source/sink maps based on human movement and parasite prevalence
- Allows precisely targeted interventions, in time and space

Fig. 3. Sources and sinks of people and parasites. Kernel density maps showing ranked sources (red) and sinks (blue) of human travel and total parasite movement in Kenya, where each settlement was designated as a relative source or sink based on yearly estimates. (A) Travel sources and sinks. (B) Parasite sources and sinks.
Cell Phone Data: Epidemiology
Wesolowski et al. (Carnegie Mellon, Harvard, etc.)

- Locations of ~15m cell phone subscribers; travel maps
- Malaria prevalence map based on existing data
- Source/sink maps based on human movement and parasite prevalence
- Allows precisely targeted interventions, in time and space
1. Key signal phrases (validated with diagnostic test)
2. Analyze positive/negative affect, cognitive attributes, lexical density, social concerns, etc.
3. Compare men/women, control versus mental illness disclosure samples
1. Key signal phrases (validated with diagnostic test)
2. Analyze positive/negative affect, cognitive attributes, lexical density, social concerns, etc.
3. Compare men/women, control versus mental illness disclosure samples
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
What’s Next? Big Data for Gender Challenge

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

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

Gender and Urban Mobility: Addressing Unequal Access to Urban Transportation for Women and Girls

**Leads:** The GovLab, UNICEF, ISI Foundation, Universidad del Desarrollo, Telefónica, DigitalGlobe

**Method:** High-resolution satellite data; call detail records

Safety First: Perceived Risk of Street Harassment and Educational Choices of Women

**Lead:** Girija Borker, PhD, Brown University

**Method:** Student surveys, Google Maps travel route data, mobile application data
WITHOUT DATA EQUALITY THERE IS NO GENDER EQUALITY

@Data2X

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

Data for Sustainable Development

Session 4:
Big Data and the global gender gap: The promises and perils of digital information

Featuring:
Rebecca Furst-Nichols
Deputy Director of Data2X

Bapu Vaitla
Data2X Fellow

Date: Tuesday, 20 March, 2018
Time: 4:00 PM CET
#SDTalks