

Use of big data in measuring development impact

Douglas Glandon, Lead Evaluation Specialist, 3ie

March 17th 2021





What is big data?

- “...umbrella term referring to the large amounts of digital data continually generated by the global population” (1)
- “...generated very fast, huge in volume, and too large to be analyzed on a single computer” (2)
- ...characterized by 3Vs:
 - *Volume*
 - *Variety*
 - *Velocity*
- Or 4 Vs, 5Vs, 7Vs, 10Vs, 42Vs...

What's going on? ◀

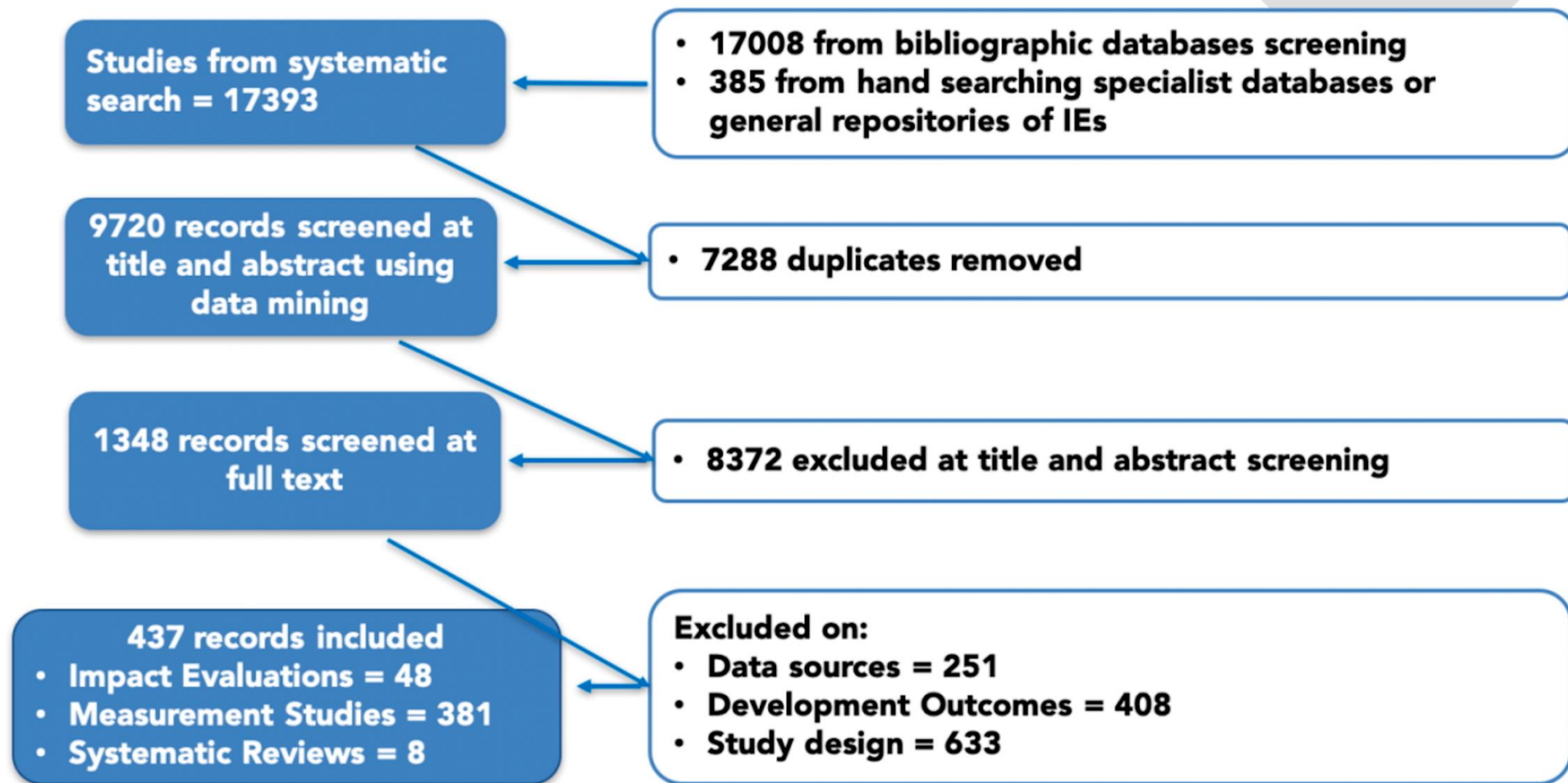
How are people using big data for measuring and evaluating development outcomes?

How does this vary by geography or outcome?

What are the main technical, methodological, and ethical issues emerging?

What lessons can be derived for impact evaluation going forward?

Search results and screening



Outcomes

Economic development and livelihoods	Sustainable agriculture and food security	Health and well-being	Education_	Governance and human rights	Water and sanitation.	Energy, industry and infrastructure provision	Urban Development	Environment sustainability	Global partnerships
--------------------------------------	---	-----------------------	------------	-----------------------------	-----------------------	---	-------------------	----------------------------	---------------------

Data Sources

- Social networks
- Internet searches
- Mobile data content
- Citizen reporting and crowd sourced data
- Data produced by public agencies
- Data produced by businesses
- Cellphone call records data (CDR)
- Data from fixed sensors
- Data from mobile sensors (tracking)
- Data from satellites

25 Impact evaluations

Forest Conservation and Slippage: Evidence from Mexico's National Payments for Ecosystem Services Program
Impacts of payments for ecosystem services programme in Mexico
Indigenous Land Rights and Deforestation: Evidence from the Brazilian Amazon
Forest Cover Impacts of Chinese Development Projects in Ecologically Sensitive Areas
Strict versus mixed-use protected areas: Guatemala's Maya Biosphere Reserve
Paper park performance: Mexico's natural protected areas in the 1990s
Clear, but don't invest: protected areas discourage some land uses more than others
The Impacts of World Bank Development Projects on Sites of High Biodiversity Importance
The Local Impacts of World Bank Development Projects Near Sites of Conservation Significance
Titling community land to prevent deforestation: An evaluation of a best-case program in Morona-Santiago, Ecuador
Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy
Using Satellite Imagery to Assess Impacts of Soil and Water Conservation Measures: Evidence from Ethiopia's Tana-Beles Watershed
Evaluating whether protected areas reduce tropical deforestation in Sumatra
Cash for Carbon: A Randomized Controlled Trial of Payments for Ecosystem Services to Reduce Deforestation
Quantifying Heterogeneous Causal Treatment Effects in World Bank Development Finance Projects
Does subway expansion improve air quality?
Using big data to evaluate the impacts of transportation infrastructure investment: the case of subway systems in Beijing
The Effects of Land Use Regulation on Deforestation: Evidence from the Brazilian Amazon
Assessing the causal impact of Chinese aid on vegetative land cover in Burundi and Rwanda under conditions of spatial imprecision
Harnessing transparency initiatives to improve India's environmental clearance process for the mineral mining sector
Impacts of community monitoring of socio-environmental liabilities in the Ecuadorian and Peruvian Amazon
A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests
The Ecological Impact of Transportation Infrastructure
Evaluating the environmental impact of payments for ecosystem services in Coatepec (Mexico) using remote sensing and on-site interviews
A Top-Down Approach to Estimating Spatially Heterogeneous Impacts of Development Aid on Vegetative Carbon Sequestration

● Impact evaluations

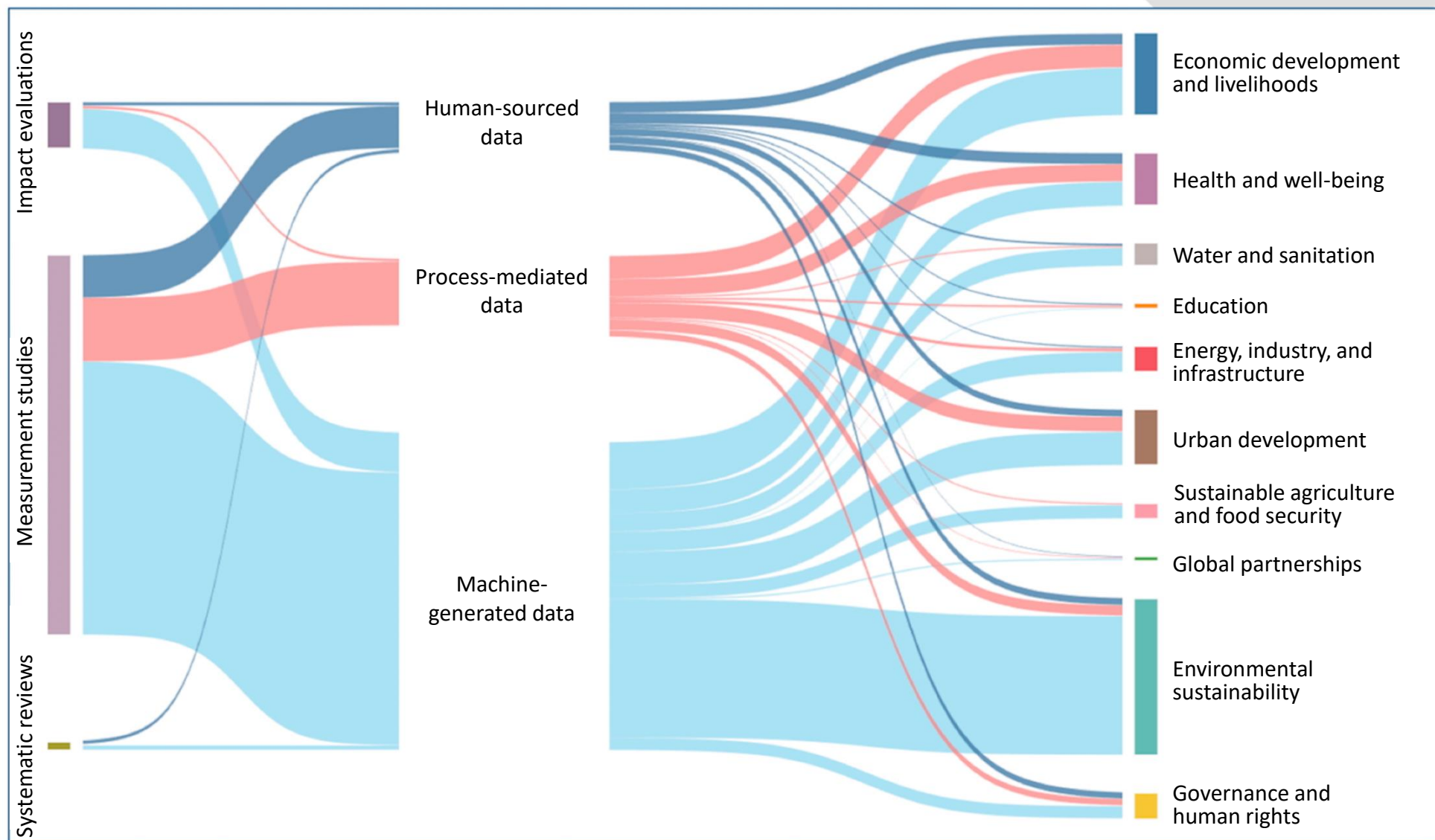
● High confidence

● Medium confidence

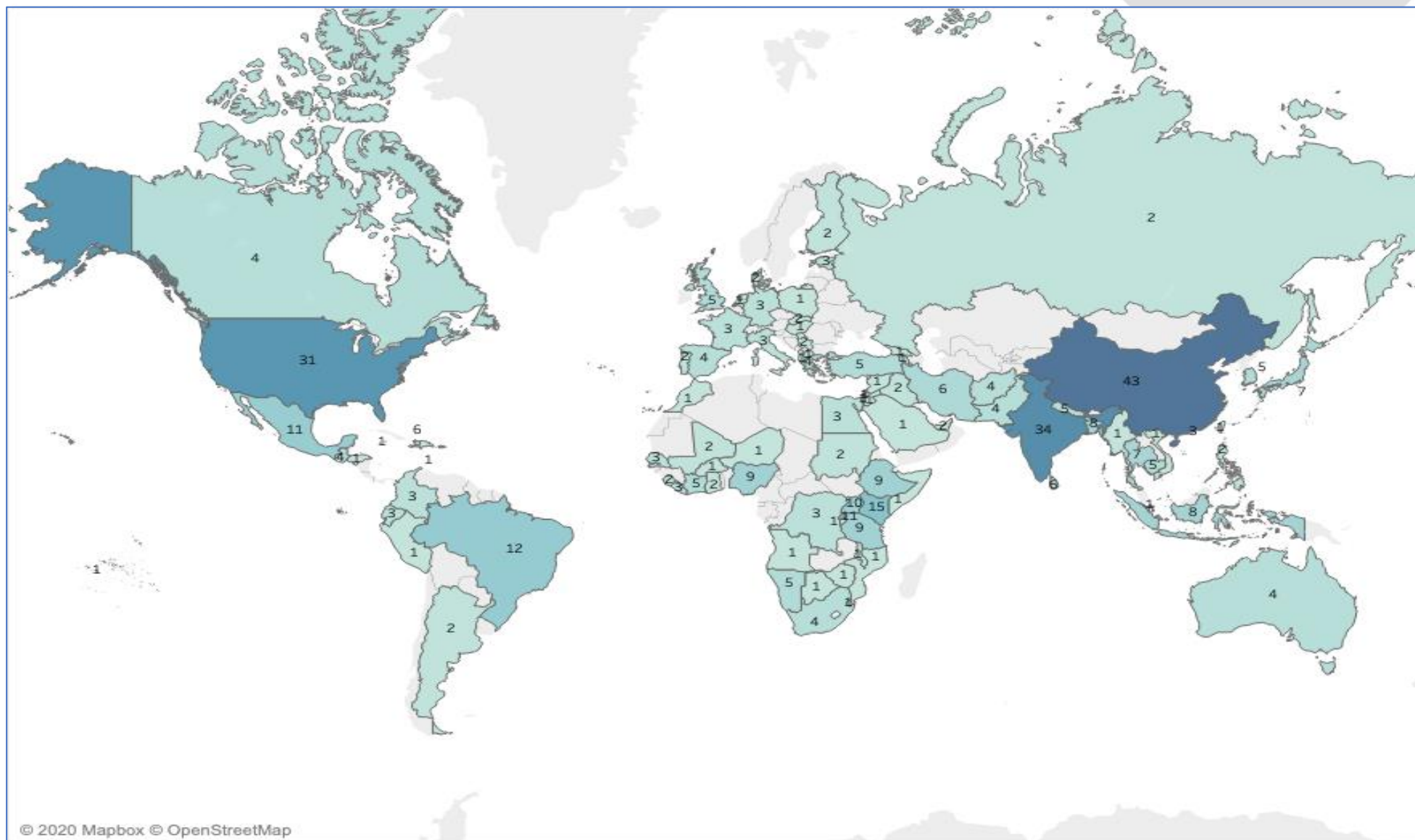
● Low confidence

● Measurement Study

Overview of the map ◀



Overview of the map ◀



- ▶ What are some key considerations and challenges to keep in mind?



A word cloud centered on the slide, featuring the word "Data" in large red font. Other words include "privacy" (blue, top), "security" (purple, middle), "quality" (light blue, right), "Unintended" (blue, middle), "exclusion" (blue, middle), "Transparency" (light blue, bottom), "Consent" (purple, bottom), and "Ethics" (green, left, rotated). The words are arranged in a circular pattern around the central "Data" word.

privacy
Data
security quality
Unintended
exclusion
Transparency
Consent
Ethics

► What does all this mean going forward?

- Increased and more diverse uses of big data in impact evaluation
- More attention to ethics, privacy, bias
- Not a replacement for traditional data collection
- Even greater need for meaningful stakeholder engagement
- Support for local research capacity, voice, and ownership



Outcomes

Data Sources

