



UNIVERSITAS  
ISLAM  
INDONESIA

Data Science Talk and Workshop Series #2

# Big Data and Social Inequalities

Ahmad Raf'ie Pratama, Ph.D.



## Ahmad Raf'ie Pratama, Ph.D.



Assistant Professor  
Secretary of Department of Informatics  
Universitas Islam Indonesia

[ahmad.rafie@uii.ac.id](mailto:ahmad.rafie@uii.ac.id)

# Social Inequalities

the existence of **unequal opportunities and rewards** for **different social positions or statuses** within a group or society

- Oxford Dictionary



UNIVERSITAS  
ISLAM  
INDONESIA

# Dimensions of Social Inequalities



**Income**



**Power**



**Occupation**



**Race/Ethnicity**



**Wealth**



**Education**



**Nationality**



**Ancestry**



UNIVERSITAS  
ISLAM  
INDONESIA

# Why inequalities are BAD?

Figure 70.1 Health and social problems are worse in more unequal countries



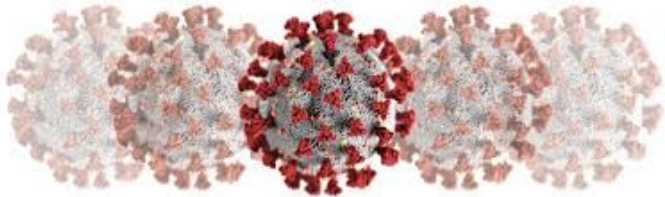
Source: Wilkinson and Pickett (2009).



UNIVERSITAS  
ISLAM  
INDONESIA

# Takes COVID-19 for instance

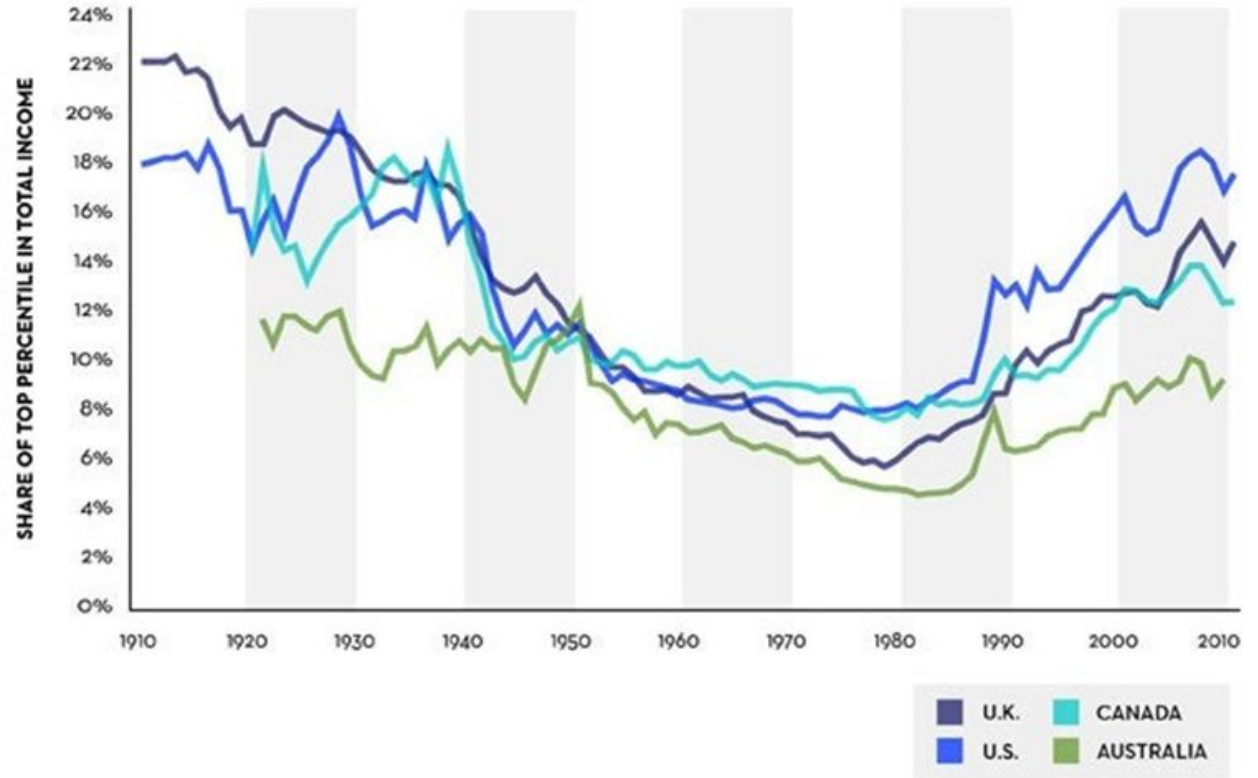
- The worst insured, the least hygiene, and the most ignorant person can determine the quality of the whole public healthcare system.
- Those from low SES (the working class) are usually the ones least likely to have the ability to work remotely.
- No matter how healthy and how good of healthcare the privileged do have, the virus will eventually knock the door once everybody else has already infected.





UNIVERSITAS  
ISLAM  
INDONESIA

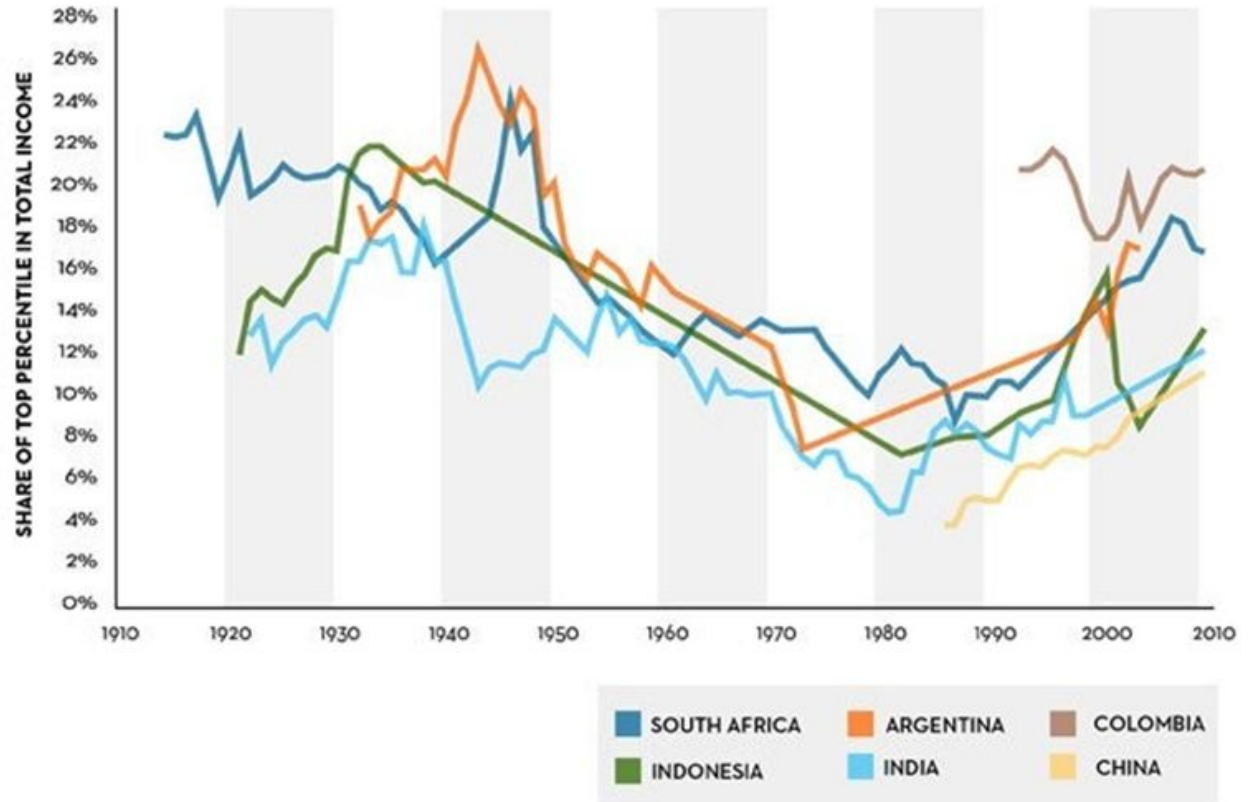
## INCOME INEQUALITY IN ANGLO-SAXON COUNTRIES, 1910-2010





UNIVERSITAS  
ISLAM  
INDONESIA

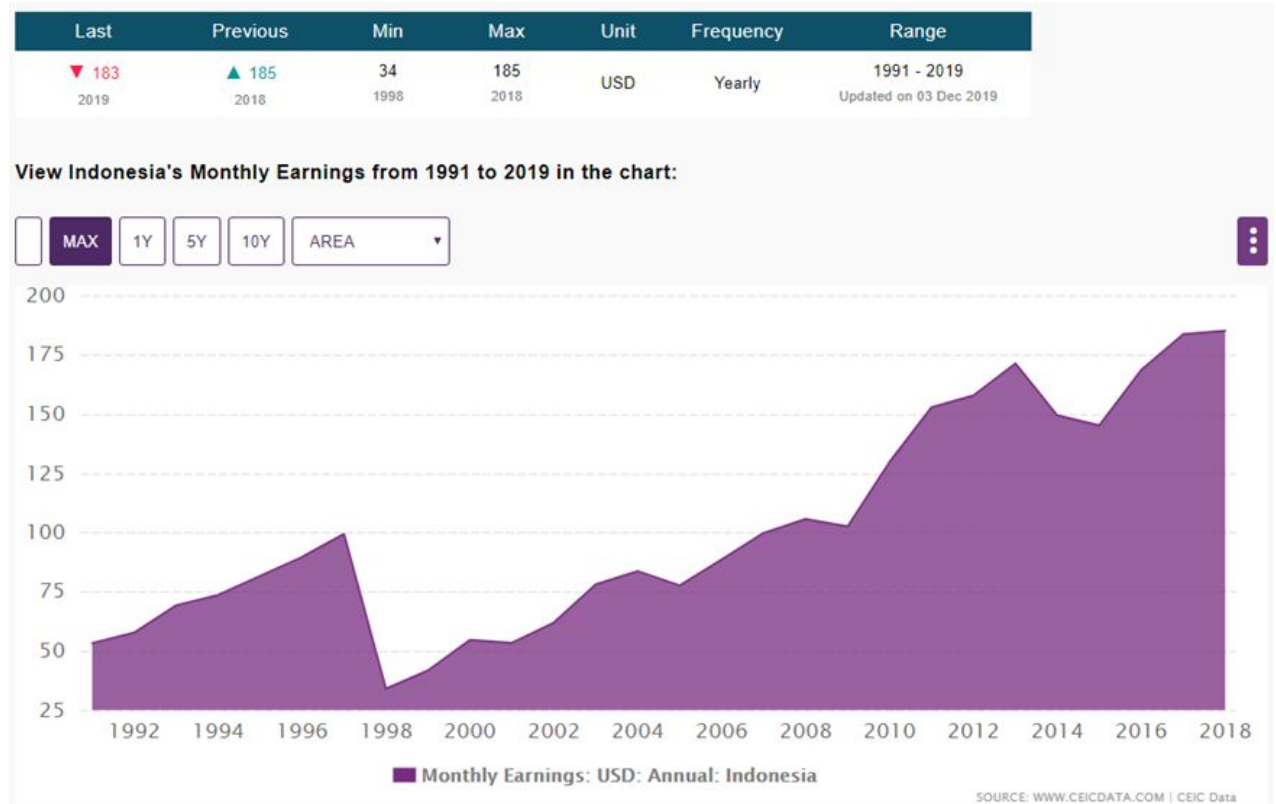
## INCOME INEQUALITY IN EMERGING COUNTRIES, 1910-2010







UNIVERSITAS  
ISLAM  
INDONESIA



# Big Data

**extremely large data sets** that may be analyzed computationally to **reveal patterns, trends, and associations**, especially relating to **human behavior and interactions**

- Oxford Dictionary



UNIVERSITAS  
ISLAM  
INDONESIA

# Positive Effects of Big Data



The rise of open science



Transform human abilities to tackle  
global challenges



UNIVERSITAS  
ISLAM  
INDONESIA

# Concerns over the use of Big Data



The unsustainable nature of the digital data landscape.



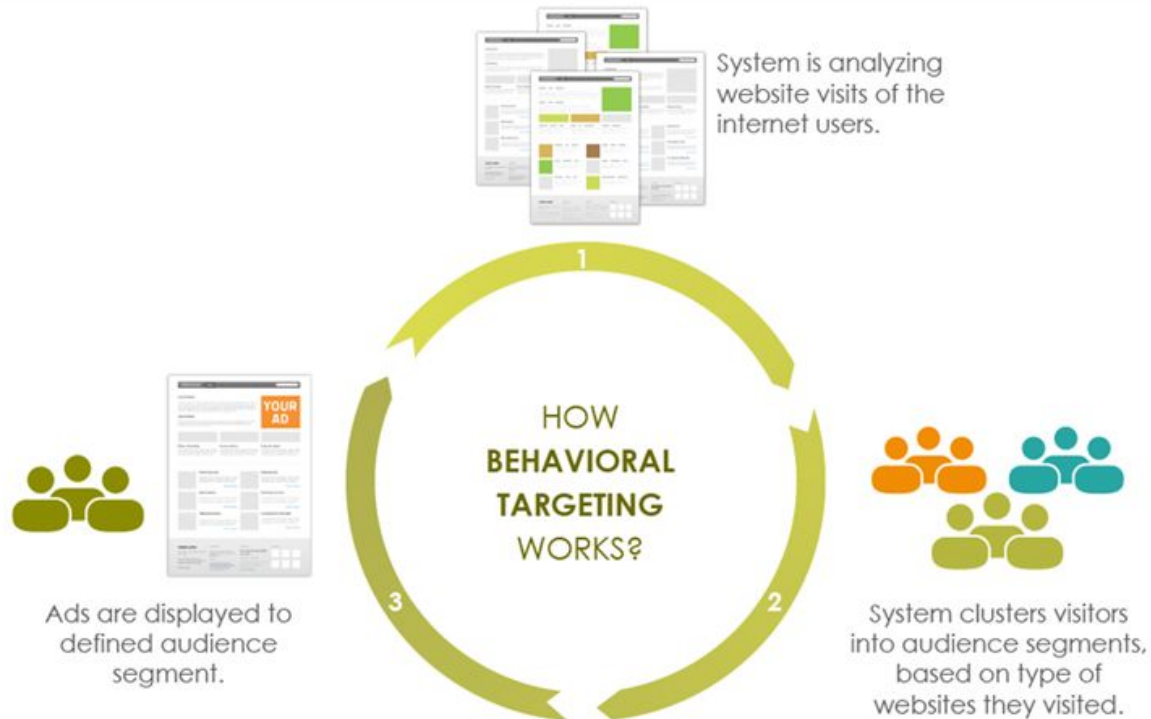
The quality and credibility of the data themselves.



The processes used to transform those data into knowledge.



The extent to which big and open data are reinforcing existing social divides.



# Case 1: Data Brokers Enable Targeting of Financially Vulnerable Communities

- Behavioral profiling allows advertisers to offer goods at different prices (price discrimination).
- Behavioral profiling is also used by especially seedy companies to target a variety of financial and economic scams at vulnerable populations most likely to fall prey to their offers.
- Users undervalue the personal data they provide and most users don't even know their data is being shared with third parties.



UNIVERSITAS  
ISLAM  
INDONESIA

THE WALL STREET JOURNAL.

TECH

## Facial-Recognition Software Suffers From Racial Bias, U.S. Study Finds

Report amplifies concerns about the rollout of algorithms that analyze images

By [Asa Fitch](#)

Dec. 19, 2019 9:01 pm ET



SHARE



TEXT

## MIT Technology Review

Artificial Intelligence Dec 20, 2019

### A US government study confirms most face recognition systems are racist



Almost 200 face recognition algorithms—a majority in the industry—had worse performance on nonwhite faces, according to a [landmark study](#).



UNIVERSITAS  
ISLAM  
INDONESIA

## **Case 2:** In Online Searches, Big Data Systems Reproduce Racial Bias

- Algorithms discriminate based on data they lack, causing minority groups to suffer as a results.
- Even Google ads discriminate based on the name of the person searched \*

\* Latanya Sweeney, Discrimination in Online Ad Delivery, 11 Queue 10 (2013), [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2208240](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2208240).





UNIVERSITAS  
ISLAM  
INDONESIA

The screenshot displays the Entelo search platform interface. At the top, a blue navigation bar contains the 'entelo' logo and several icons: a magnifying glass for 'Search', a sonar icon for 'Sonar', a list icon for 'Lists', an envelope icon for 'Track', a folder icon for 'Reports', and a user profile icon. Below the navigation bar, the main content area is divided into a left sidebar and a right main panel.

**Left Sidebar:**

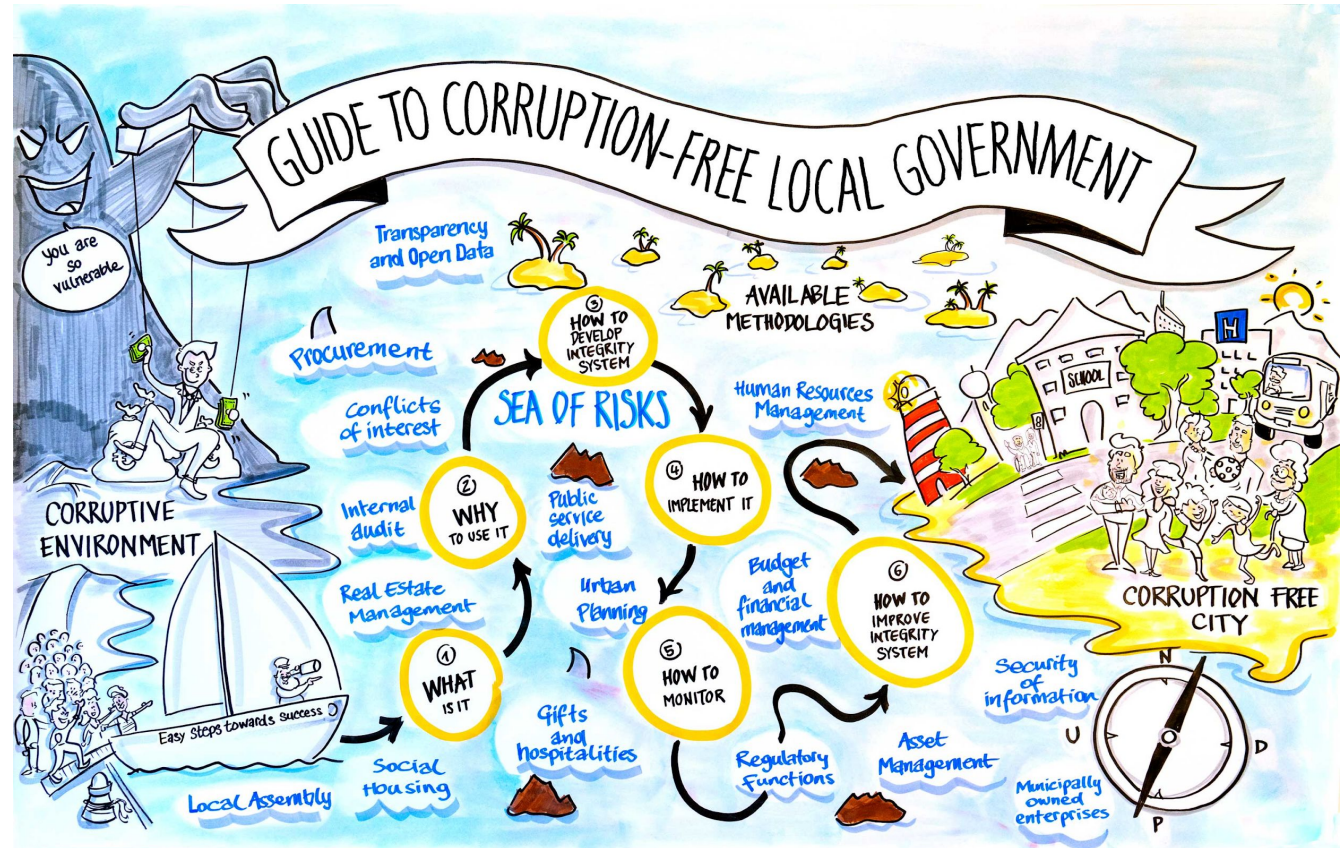
- Search:** A search bar containing the text 'UX DESIGNER' with a 'Save' button to its right.
- Filters:**
  - Keyword:** A dropdown menu with a magnifying glass icon.
  - Name:** A dropdown menu.
  - San Francisco, CA 94101:** A dropdown menu.
  - Edit:** A dropdown menu with 'Everywhere' selected.
  - More Likely to Move:** A button with a star icon.
  - Email Available:** A button with a star icon.
  - Diversity:** A button with a star icon.
  - Companies & Positions:** A button with a star icon.
  - Company:** A text input field.
  - Current only:** A checkbox.
  - Position:** A text input field.
  - Current only:** A checkbox.
  - Years of experience:** A text input field.
  - Min. months at current job:** A text input field.
  - Exclusions:** A button.
  - Social:** A button.
  - School & Field of Study:** A button.
  - Add profile:** A button.

**Right Main Panel:**

- Stephanie Anderson:** A profile card showing a placeholder image, name, location (SAN FRANCISCO BAY AREA, US), current role (UI Designer at XYZ Co. - over 8 years), previous role (Senior Interaction and Visual Designer at MNO Co. - 11 months), education (Westwood College of Technology, Bachelor of Arts, Visual Communications), and skills (web design, design, HTML, Music, photography, typography). A 'More Likely to Move' button is present.
- John Doe:** A profile card showing a placeholder image, name, location (SAN FRANCISCO BAY AREA, US), current role (UI Designer at ABC - over 4 years), previous role (UX designer and Visual Designer at the MPC - 7 months), education (Eina School of Technology, Bachelor of Arts, Visual Communications), and skills (web design, design, HTML, Music, photography, typography). A 'More Likely to Move' button is present.
- Samuel Thompson:** A profile card showing a placeholder image, name, location (SAN FRANCISCO BAY AREA, US), current role (Senior Mobile Product at Visa - over 2 years), previous role (Product Manager at Consultant - over 1 year), education (Westwood College of Technology, Bachelor of Arts, Visual Communications), and skills (web design, design, HTML, Music, photography, typography). A 'More Likely to Move' button is present.
- Mary Ann:** A profile card showing a placeholder image, name, location (SAN FRANCISCO BAY AREA, US), current role (Experience Designer at Microsoft - over 2 years), previous role (Director at Addcherry - 11 months), education (Arizona State University, Bachelor of Arts, Visual Communications), and skills (web design, design, HTML, Music, photography, typography). A 'More Likely to Move' button is present.

## Case 3: Hiring Algorithms May Put Jobs Out of Reach

- Institutional, structural, and other forms of bias are critical aspects of any equity analysis, especially when it comes to employment.
- Without active measures to mitigate them, biases will arise in predictive hiring tools by default.
- While predictive hiring tools rarely make affirmative hiring decisions, they often automate rejections.



## Case 4: The Temptation to Abuse Government Databases

- Governments at every level collect personal data, with varying degrees of scope and detail.
- Private companies often play a central role in helping governments gather and store data.
- The availability of personal data—and government's inability to properly manage and oversee its own staff's activities—has led to abuse.

# Solution?

Substantial investment in **Data Governance and Stewardship**

Priority in supporting **Fairness in Data Handling**

Promotion in **Interdisciplinary** and **Interdomain Collaboration**



# What needs to be done?

- The identification of exclusions and inequalities built into data pipelines.
- Investment in adequate data expertise and skills
- The development of intelligent and ethical strategies for data sharing and the development of algorithms.
- Mechanisms to promote the quality and trustworthiness of data sources and analytic tools.
- Creative solutions to global challenges in dialogue with relevant publics.

## What needs to be done? (cont)

- The understanding that whether and how data should be open needs to be decided on a case-by-case basis.
- Sustainable data infrastructures and related expertise.
- Resources that are interoperable, long-term, internationally coordinated, and publicly accountable.
- Critical scrutiny of research across different audiences.



UNIVERSITAS  
ISLAM  
INDONESIA

# Thank You!

## Any Questions?

**Ahmad Raf'ie** Pratama, Ph.D.

---