Artificial Intelligence and Big Data for Anti-corruption

Yasmine Hamdar
Computational Social Scientist, DPA

June 16, 2022
Data-Pop Alliance
Data-Pop Alliance is a non-profit “Think-and-Do-Tank” created in 2013 by the Harvard Humanitarian Initiative, MIT Connection Science and ODI that brings together researchers, practitioners and activists who want to “Change the World with Data” through projects in and with Global South countries. DPA is incorporated as a Charity or equivalent in the US (2016), Mexico (2018), Senegal (2021) and France (2022).
Our 6 Thematic Programs

1. Just Digital Transformations
2. AI and Statistics for the SDGs
3. Resilient Livelihoods and Ecosystems
4. Data Feminism
5. Geographies of Inequalities
6. Technology and Democracy
“Overview and Outlook 2021-23” Report
18 out of 22 Arab states have a corruption perception index < 50*

* 100 is very clean
* 0 is highly corrupt
Artificial Intelligence: The next frontier in anti-corruption?

Figure 1. World Bank Blogs - “Arab Barometer Report: Perceptions of corruption on the rise across MENA”

<table>
<thead>
<tr>
<th>Country</th>
<th>Percent</th>
<th>Paying a Bribe for Better Healthcare Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>MENA OVERALL</td>
<td>47</td>
<td>% saying highly or somewhat necessary</td>
</tr>
<tr>
<td>Algeria</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Jordan</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Kuwait</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Libya</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Palestine</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Yemen</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Weighted Estimates
Source: Arab Barometer, Wava V
AI-ACT Approaches: Top-down versus Bottom-up

Example:
Human auditor identifies and prosecutes corruption

Top-down Anti Corruption Government introduces checks & balances to reduce corruption

Example:
Machine Learning tool predicts risk of corruption from data

Traditional anti-corruption

Corruption Abuse of entrusted power for private gain

AI-supported anti-corruption

Example:
Citizens organizing protest against corruption

Bottom-up anti-corruption Identify and support existing efforts within society to reduce corruption

Example:
Tweetbots motivate citizens' action by flagging suspected corruption

Figure 2. AI-ACT Approaches. Adapted from "The promise and perils of using artificial intelligence to fight corruption," by N. Köbis, C. Starke & I. Rahwan, 2022
Requirements for Successful AI-ACT

Input data
- Quality and biases
- Size
- Source

Algorithmic design
- Accuracy
- Efficiency

Institutional implementation
- Social context
- Awareness
Use of AI: Natural Language Processing for Social Media Data

Output: flagged or not

Response

NLP Model

Text:
Last night I again saw @constables taking bribe, near the same location (on the border). All hopes of my fight against corruption, shattered... Unfortunately I didn't have my phone at that time, otherwise I would have recorded their faces. (4)
Use of AI and Big Data Analytics: Revealing Fraud in Public Procurement

Figure 3. Investigating bidding patterns. Adapted from "Using Data and Transparency to Fight Corruption in Public Procurement" by S. Zimmermann - data modeling using WBG and national data.
Challenges and Concerns

- **Transparency**
- **Open Data**
- **Ethical challenges:** privacy concerns and data leaks
Digital Preparedness

- Cutting-edge Tech
- Custom-tailored Data
- Public AI Algorithms
- Publicly Available Data
Thank you

yhamdar@datapopalliance.org

