From Big Data to Big Impact for Social Good

In the age of information abundance, data has emerged as a byproduct of digital life and a potent force for transformation. Yet, as promising as it is, the use of data to solve the world's most pressing social, environmental, and economic problems, commonly framed as "Data for Good," remains far from its full potential.

The Promise of Data for Good

Data generated from mobile phones, digital transactions, sensors, satellites, and social media platforms offers unprecedented opportunities to address critical challenges from pandemics and poverty to climate change and disaster response. As the UN's 2024 *Data for Development* report emphasizes, data can inform sustainable development, enable responsive governance, and power scientific innovation **if governed, shared, and applied responsibly**.

Similarly, Abbasi et al. (2023) argue that while data science has evolved through phases of business intelligence and machine learning into today's AI-driven paradigm, **its alignment with social good objectives has paradoxically weakened**. Their empirical review shows that the proportion of data science research focused on issues like health equity, climate resilience, and poverty has declined even as the field booms.

Where Potential Meets Reality

The COVID-19 pandemic was a real-time test of the Data for Good agenda and a humbling one. Buckee et al. (2022) recount how, during the pandemic, tech companies released aggregated mobility data to inform lockdown strategies. However, **most of these efforts failed to translate into actionable insights for decision-makers**, especially in low-resource settings. The reasons? A fragmented data pipeline, lack of trusted intermediaries, and institutional unpreparedness.

These shortcomings are not unique to health crises. Across sectors, there exists a persistent **gap between data availability and real-world impact**.

Barriers to Impact

1. **Access and Governance**
    - Most valuable data resides in private platforms, inaccessible to public institutions.
    - Ethical and legal frameworks for responsible sharing remain underdeveloped.
2. **Analytical and Translational Capacity**

   - Complex data streams require high-level technical expertise, often missing in public sector agencies.
   - There's a shortage of what Buckee et al. call "**data bilinguals**" professionals who can bridge data science and public policy.
3. **Institutional Readiness**
   - Many local governments lack the digital infrastructure, interoperability standards, and decision-making frameworks to operationalize data insights in time-critical settings.

A Framework for Action

Abbasi et al. offer a **DSSG (Data Science for Social Good) Framework** that integrates:

- Research genres (e.g., computational design, theory-building)
- Social good challenges (e.g., equity, sustainability, justice)
- Socio-technical contexts (people, organizations, institutions)

This tripartite model urges academia to rethink how data science research can be both **rigorous and relevant** by designing systems and insights with clear links to real-world societal outcomes.

In parallel, the UN calls for **global data governance mechanisms** that harmonize ethics, privacy, and economic considerations while fostering multi-stakeholder collaboration.

From Principles to Preparedness

To unlock data's potential for good, we need to act on several fronts:

- **Build data infrastructure with inclusion at the core**: Open data standards, ethical frameworks, and privacy-first protocols must be globally adopted, particularly empowering the Global South.
- **Train the next generation of "data bilinguals"**: Equipping professionals who can contextualize, translate, and ethically apply data in complex policy environments is key to lasting impact.
- **Institutionalize data preparedness**: Just as disaster drills prepare communities, *data drills* must prepare institutions, setting up agreements, methods, and roles before crises strike.

Conclusion: Big Data Must Serve Big Purpose

The data revolution is here, but its benefits are not automatic. Without thoughtful governance, ethical foresight, and interdisciplinary collaboration, data risks deepening the very inequities it promises to fix. But with the right investment in people, policy, and platforms, data can become not just a tool for prediction, but a force for justice,

sustainability, and human flourishing. “Data preparedness must become an integral component of disaster preparedness.” That logic applies more broadly: **data must be embedded not just in technology stacks, but in the DNA of institutions committed to the public good.**

References:

Abbasi, A., Chiang, R. H. L., & Xu, J. J. (2023). *Data science for social good*. Forthcoming in the *Journal of the Association for Information Systems*. https://doi.org/10.17705/1jais.00849

United Nations Conference on Trade and Development (UNCTAD). (2024). *Data for development: Report of the Secretary-General* (E/CN.16/2024/2). United Nations Commission on Science and Technology for Development, Twenty-seventh session. https://unctad.org/meeting/commission-science-and-technology-development-twenty-seventh-session

Buckee, C., Balsari, S., & Schroeder, A. (2022). Making data for good better. *PLOS Digital Health*, 1(1), e0000010. https://doi.org/10.1371/journal.pdig.0000010