AI4SDG – roadmap to a Global Data Commons to achieve the Sustainable Development Goals (1/2)

We invited global technology leaders to submit their views on what it would take to make Global Data Commons a reality

The Global Data Commons aims to deploy AI to help achieve the SDGs

The position papers discuss several challenges that prevent the implementation of a GDC

1. Access to quality data: 30
2. Technical challenges: 26
3. Legal challenges: 23
4. Political/regulatory challenges: 22
5. Social challenges: 21
6. Business/commercial challenges: 11

Example quotes from position papers:

- Capitalizing on the immense volume of data available and use AI to tackle the world’s greatest challenges
- Detect, present and help scale-up use cases for AI enabling the 17 SDGs
- The use of AI for Sustainable Development Goals will allow us to:
  - Monitor progress towards the achievement of SDG
  - Simulate implications
  - Predict outcomes of measures taken
  - Provide recommendations for policy makers

Our review of the position papers indicates that we can follow an iterative and systematic approach to creating the Global Data Commons

- Define use case and scope
- Identify stakeholders, associated risks and rewards
- Identify legal and regulatory context
- Identify rights and responsibilities for using data
- Articulate collaboration examples and limitations
- Identify relevant ethical and associated characteristics
- Define interoperability with existing data commons
- Identify use case-specific data considerations, i.e., anonymization, pseudonymization
- Map out end-to-end data lifecycle, i.e., platform from originators to end user distribution
- Identify technology requirements for present data value chain, e.g., storage, analytics, distributed ledgers
- Define data control issues such as access rights and usage, etc.
- Define required technology architecture and systems
- Determine technology requirements and select best practice technology solutions/implementation strategies available
- Define governance model
- Define stakeholder management model
- Define trust management and regulatory framework
- Define key requirements for contracts to be implemented
- Tailor AI standards templates to specific use case context

Scope definition and risk assessment
Data and data flow and workflow
Contracts and agreements definition

1. Data and data flows, and data lifecycle, i.e., data from origination to end
2. Technology solution and infrastructure
3. Geopolitical transparency, accountability and insight
4. Governance solution and infrastructure
5. Business and agreements definition
6. Scope definition and risk assessment

1. Project outcomes and lessons learned
2. Conceptualization
3. Year 2-3
4. Technology solution and infrastructure
5. Contracts and agreements definition
6. Data and data flow and workflow
7. Scope definition and risk assessment
8. Build
9. Scope
10. Scale
AI4SDG – roadmap to a Global Data Commons to achieve the Sustainable Development Goals (2/2)

We invited global technology leaders to submit their views on what it would take to make Global Data Commons a reality.

Survey across all technology leaders indicates consensus on some key high-level themes...

A. Scope definition

- Define the end goal (link to specific SDG) to rally stakeholders; detail platform capabilities and create a 60 second crisp narrative to explain GDC to external stakeholders

- Choose a pilot use-case that solves a specific SDG objective, is a trans-national challenge, and has a ready set of public, private and social sector partners who have low conflict of interest (monetary or IP related)

B. Governance

- Set up supervisory body (e.g., PMO) to define high level scope and attract initial partners

- Identify representative partners across sectors (e.g., govt (UAE), multilateral (space agencies), NGO (citizen science))

GDC platform capabilities:

- GDC is a combination of data repository & data directory
- GDC should leverage only non-human data (e.g., satellite, weather)

GDC’s scope of data assets:

- GDC is a holistic data repository of multiple datasets
- GDC should also leverage personally identifiable data with advanced encryption

GDC architecture:

- Decentralized or distributed (data stored & processed at multiple sites)
- Centralized (data stored in central location)

...while technical infrastructural details need to be further investigated

1. Align GDC aspirations with social goals
2. Expand the scale of GDC to create impact
3. Craft innovative incentive mechanisms
4. Institutionalize trust through technology
5. Ensure long-term sustained commitment

Respondents believe implementing GDC is a good idea despite challenges
Respondents believe SDGs are a good starting point for a GDC
Respondents believe structured use case driven approach is best way forward

>80%
~90%
~70%