Annex 3: Unpacking the complexity of data agents, data purposes, data protection and promotion

(for more information see https://unsceb.org/international-data-governance-pathways-progress)

1. The data economy – the "Datasphere"[1] – is a complex adaptive system capable of *generativity*, meaning that data can be used "for various ends, many of which were unanticipated or, if anticipated, would never have been thought to be valuable."[2]

2. There are many agents involved in generating, using and re-using data with different purposes. Different types and different potential uses of data require different levels of protection and promotion. This annex describes the interaction of multi-agent and multi-purpose data highlighting the different characteristics and regulations that (should) affect different types of data. The annex also introduces the concept of public good data, which has a different connotation of data for public good and elaborates on global public good data as a space where the UN system has a clear role.

Purpose for <u>generating</u> or re-using data	Data can be (re-)<u>used</u> for	Promotion	Protection				Example	Challenges	Title and comments
Private intent Typically generated by private sector	**Private interest** Generate profits by selling and re-selling data	Economic value of data to be realized from the collection, dissemination, trade, use and reuse of data	Respect privacy and consent if data relate to individual persons or business	Considerations for regulating the data market may be similar to regulations applied to other economic sectors and/or markets			Marketing data	Challenges in trading confidential/ individual data across entities (industries, institutions, countries) if different privacy regulations apply	**Private-intent data for private use** This space feeds the "data economy" where the economic value of data provides incentives to innovate
Private intent	**Public interest**	Data collected and "own" by the private		Define equal access regulations to			Big data generated by utility	To find the right balance to maximise	**Private-intent data for the public good**

---

[1] Datasphere – Wikipedia, https://en.wikipedia.org/wiki/Datasphere

[2] Zittran, J. (2007). *Saving the Internet*, Harvard Business Review. https://hbr.org/2007/06/saving-the-internet

Purpose for **generating** or re-using data	Data can be (re-)**used** for	**Promotion**	**Protection**			Example	Challenges	**Title and comments**
Typically generated by private sector		sector may have a unique value for the public. Data generated for private intent can be reprocessed and reused to maximise their benefit to the public.		ensure that the data "own" by the private sector that have public interest benefit people in an equitable manner, it reaches people in an equitable matter, and all can realize the potential of data, avoiding concentration		companies used for statistical purpose to define mobility patterns of people and design public transportation

Data generated from health research are made public to realize their maximum public health potential | public use while not undermining the economic incentive for producing the data | |
| **Public intent**

Typically, one or few public institutions together generate and manage data for the public good using public | **Public interest** | Data are generated or reused to design public policies | | | Open data policy, policy to encourage the use of data, personal identifiable data to be regulated for accessibility under conditions

Need to define | National statistical offices produce official statistics by generating new data (through surveys and censuses for example) or by reusing | It requires targeted public investments on data.

There could be lack of incentives for innovation | **Public-intent data for the public good** |

Purpose for __generating__ or re-using data	Data can be (re-)__used__ for	Promotion	Protection				Example	Challenges	Title and comments
resources					and apply a transparent quality framework to make public institutions accountable to the public for the generation of data that can effectively generate the knowledge needed for public policy and for the benefit of all people equally Equitable distribution of benefits from the data		administrative data		
Community intent Participatory approach to generate data Shared data generation (or reuse) for a shared value	**Community interest and public interest**					Equal commitment of all in the community to equally contribute to data for equal benefits. Need a common quality framework.	Countries contribute to generate UN data for the global monitoring of the SDG, following global standards agreed by the countries themselves. The UN serves as the "honest broker" that collects data from the members of the community (MS) and distribute the aggregated data back to the communities providing equal access to all	Not everyone may have equal power in the communities and some members could be more powerful in defining the data community objectives. It requires an honest broker that is empowered by the members of the community (data contributors) to ensure that the communities' rules (quality standards, equal contributions, …) are applied to the benefit of everyone in the	**Public good data (publicly-generated data for public good)** At global level: **Global public good data** The difference between public-intent and public generated data is that the public-intent data are typically managed by one or few institutions while data publicly generated is managed by a public community

Purpose for generating or re-using data	Data can be (re-)used for	Promotion	Protection				Example	Challenges	Title and comments
								community.	

3. Public goods are those that are available to all ("nonexcludable") and that can be enjoyed over and over again by anyone without diminishing the benefits they deliver to others ("nonrival"). The scope of public goods can be local, national, or global. Global public goods are those whose benefits affect all citizens of the world. They encompass many aspects of our lives: from our natural environment, our histories and cultures, and technological progress down to everyday devices such as the metric system.

4. Data has unusual properties. Data is different from goods and services, as a non-rivalrous resource that can - technically - be replicated and combined in numerous value chains without being depleted and under the right governance framework, can be treated as a public good. These actions must be underpinned by strong, ambitious governance measures that activate public, private and UN actors. Three examples:

- Climate - data is at the heart of climate action – both in terms of assessing and reducing emissions as well as boosting climate resilience globally. The UN system produces various authoritative datasets that are crucial for the global fight against climates, including emission inventories and meteorological data as well as global assessment of the progress through the Global Stocktake. These data are critical for measuring the gap between the goals and ambitions of the Paris Agreement and national action. They are also essential for charting the changing climate and feed prediction models, which give insight into the risks that lie ahead. Reliable meteorological data that are created and shared internationally, provide an essential evidence base required to build global consensus and garner the political will to act to minimize the extent and impacts of climate change. The same data feed daily weather forecasts predicting storms, heatwaves, floods and droughts days to months in advance, protecting lives and livelihoods. These data also support the economic prosperity of many weather-sensitive sectors - such as food systems, transport, and energy. UN organizations, such as the Intergovernmental Panel on Climate Change (IPCC), the World Meteorological Organization (WMO), UN Environment Programme (UNEP), the United Nations Framework Convention on Climate Change (UNFCCC) and others have been involved in the establishment of international agreements relating to these data for decades.

- Crime – with the increasingly transnational nature of crime and the global reach of organized criminal groups, disconnected national data systems can't identify the full nature of the criminal threats. Only data systems that can "connect the dots" of criminal activities across countries can provide the necessary information to dismantle transnational organized criminal groups. Such systems can function only if there is a shared responsibility to feed the system with quality data and maintain it as a global public good service.

- Health – in a hyper-globalized world, the spread of disease does not respect national boundaries or borders. The ongoing pandemic illustrates the point - within one month of the WHO being informed of a new strain of coronavirus, the disease had already been detected in 24 countries. Only by sharing data on genetic sequencing, cases, deaths, and containment measures could scientists and epidemiologists around the world develop vaccines and other controls to limit the spread of the disease and help governments to formulate appropriate policy responses. There will be other health emergencies – governments must recognize their shared responsibility to share timely, quality data and

maintain these data as global public goods. These include some extremely sensitive personal data, and therefore, responsibilities also include the need for robust confidentiality.