



United Nations

Department of Economic and Social Affairs
Statistics

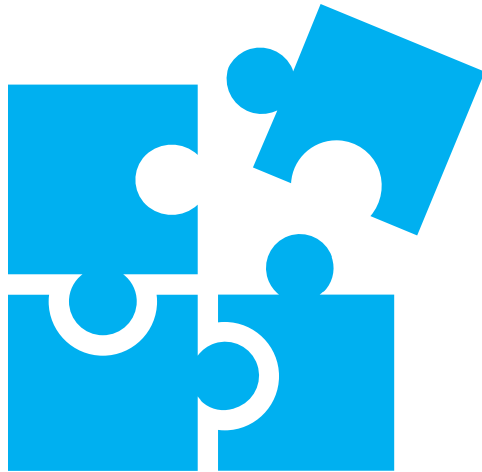
UNSD-DFID Project on SDG Monitoring
National Reporting Platforms:
Interoperability and statistical standards



SDG data ecosystem is characterized by multiple tensions

- Global vs local data needs
- Top-down vs bottom-up data producers
- Structured data exchange vs organic data sharing processes
- Sectoral vs. Cross-cutting data applications

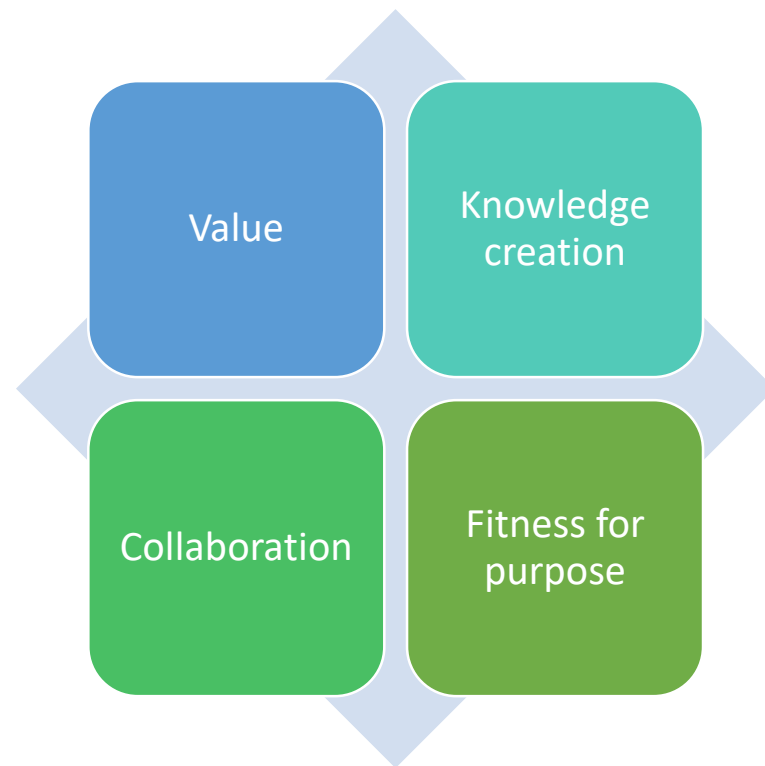
Data interoperability challenge



It's difficult to share, integrate and work with the wealth of data that is available in today's digital era:

- Divergent needs and capabilities of multiple internal and external constituencies
- Disparate protocols, technologies and standards
- Fragmented data production and dissemination systems

Interoperability is a characteristic of good quality data



Data interoperability for the SDGs



- There are many unrealized opportunities to extract value from data that already exists to meet information needs of the 2030 Agenda



- Investing time and resources in the development and deployment of data interoperability solutions will help us make better use of the data that currently sits in sectoral and institutional silos to implement and monitor the SDGs

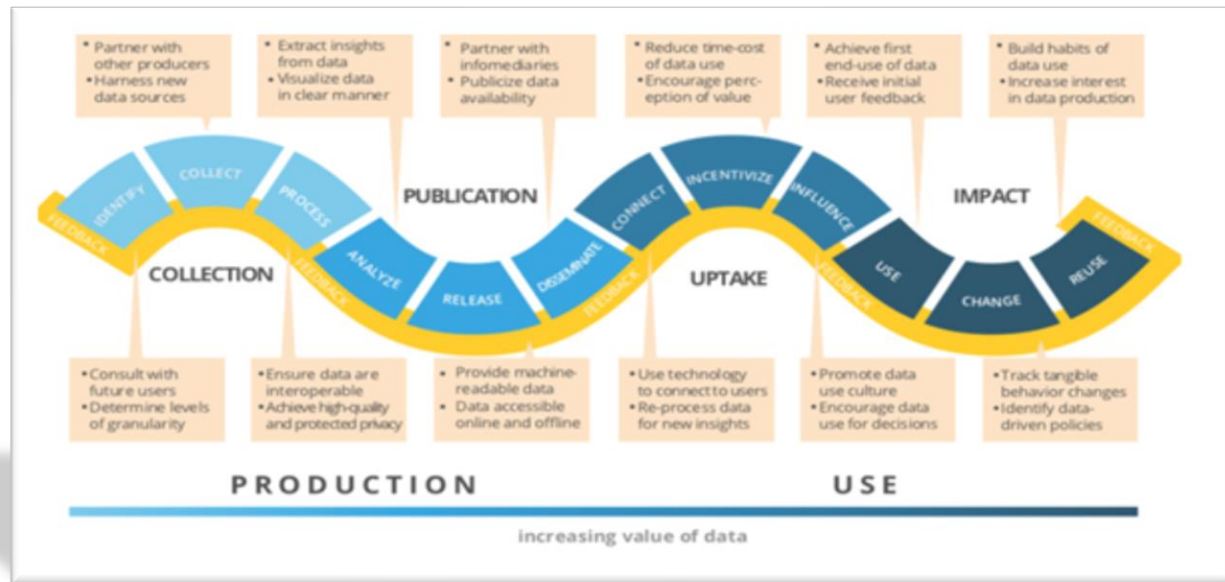


Interoperability and Open Data

- Many organizations are now adopting open data policies that authorize and facilitate the reuse of their data assets
- Open data requires data to be interoperable not only from a technical perspective, but also from a legal and institutional perspective



Interoperability in the data value chain



- Interoperability is a key element in data collection
- Planning for interoperability should take place early in the data life cycle
- But interoperability considerations also should inform every step in the data value chain

Source: Open Data Watch



Getting the **governance** and **institutional framework** right



Designing **data structures** with users in mind



Standardizing the **data content**



Providing **standard interfaces** to access and work with data



Disseminating **linked open data** for knowledge creation

Data platforms management and governance

- Very often interoperability is less a technology problem and more a data management and governance issue
- To be effective, data platforms need oversight and accountability across their lifecycle.
- Institutional frameworks play a key role in creating the environment where data, technology, and business processes fit with each other

Data management and governance

- Legal and regulatory frameworks are crucial to interoperability
 - They set the boundaries of what is acceptable conduct and what is not
 - They specify how data can be shared across organizations (e.g., standards for data reporting, security and protection)
 - They determine what data can, or cannot, be shared and integrated (for example, data protection and privacy laws).



Data and metadata models

- SDG data and metadata is often represented in variety of (usually incompatible) ways across different platforms within the National Statistical System
 - ✓ Prioritizing internal operational needs
 - ✓ Having a specific applications in mind



Data and metadata models

- Interoperability is highly dependent on data and metadata modelling decisions and practices
 - ✓ Producers and users of data must have a common understanding of how it is structured in order to effectively exchange it across systems.
 - ✓ They must also share a common understanding of how the various components of a dataset relate to each other and to the components of other datasets.
- Data and metadata modelling decisions can ensuring that systems are designed with interoperability in mind from the outset



Data and metadata models

- There is no single “right” way of representing information
 - Some data structures are better suited for operational processes (e.g., capturing data from a survey or maintaining a civil registration database)
 - Others are being better suited for data sharing and dissemination (e.g., for the creation of data visualizations)

Can we ensure that all SDG data across the National Statistical System are mapped to a common data and metadata structure?



SDMX - SDG Data Structure Definition

- **Generic** template for the integration of SDG indicator data
 - Highly **reusable** and conducive to **data sharing**
 - **Focused on simplicity**, so data is easily understood by a wide range of users and applications
 - **Self-contained** and **stable** over time
 - Incorporates **standard definitions and classifications**
 - **Extensible** to include national disaggregations
 - Data platform and **technology-independent**
- **Already tested** in by various countries



Data and metadata models

- ✓ SDMX modelling guidelines and DSD guidelines provide a step-by-step introduction to data modelling
- ✓ The guidelines contain numerous links to further, more detailed guidelines and templates that can be used in the modelling process.



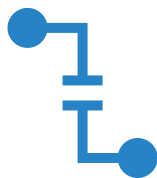
Standard metadata schemas

- A metadata schema specifies the metadata elements that should accompany a dataset within a domain of application.
- For instance: title, description, date of issue or modification, subject, publisher, license, keywords, etc.



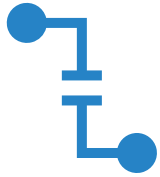
Caution!

- A common mistake in modelling datasets for sharing and dissemination is to try to replicate internal data structures from operational database systems.
- The focus should be on producing simple, self-contained datasets that are easy to understand and manipulate by users and client applications.



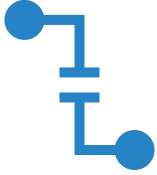
Open data formats and standard interfaces

- Data needs to be easily available and accessible to a variety of user groups.
- Interoperability is not only about standardized data production, but also about standardized “data logistics” (Walsh and Pollock)
- There is need for common “pathways” to get data from providers to users in a fast, convenient, effective, and efficient manner.



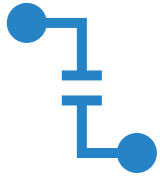
Open data formats and standard interfaces

- A first step is to make the open data available through bulk downloads in open data formats (such as CSV, JSON, XML, and GeoJSON, etc)
- Data interoperability is greatly enhanced when electronic data files are made available using openly documented, non-proprietary formats.
 - ✓ Human-editable and machine-usable
 - ✓ Agnostic to language, technology and infrastructure.



Open data formats and standard interfaces

- Application Programming Interfaces (APIs)
 - ✓ Highly-reusable pieces of software that enable multiple applications to interact with an information system.
 - ✓ Provide machine-to-machine access to data services
 - ✓ Enable users to focus on the data rather than spend their time collecting it. APIs
 - ✓ Allow to automate data flows that involve repetitive and frequent data sharing and exchange operations, avoiding costly and error-prone manual intervention.
 - ✓ Provide the building blocks for users to easily pull the data elements they need to build their applications.



Open data formats and standard interfaces

- System interfaces should **prioritize interoperability and flexibility over specificity and optimization.**
- A balance must be struck between specific user group needs and broader usability.
- Over-customization of an interface can inhibit its accessibility, usability and interoperability with other systems.



United Nations

DESA
Statistics Division



Thank you.

Questions or inquiries?
Contact us at unsd-dfid@un.org