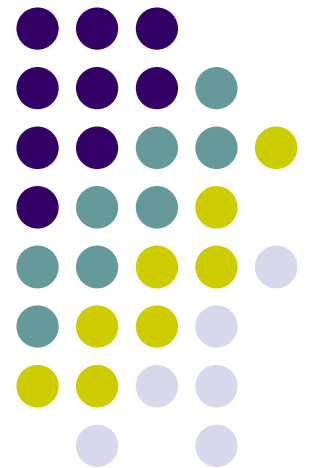


CountryData

Sharing Best Practice in the Dissemination of
National Development Indicators

Introduction to SDMX





What is SDMX?

- **Statistical Data and Metadata eXchange**
- An initiative sponsored by seven international organizations
 - Bank for International Settlements
 - European Central Bank
 - Eurostat
 - International Monetary Fund
 - Organization for Economic Cooperation and Development
 - United Nations
 - World Bank



Goals of SDMX

- “The rationale of SDMX is standardisation for statistical data and metadata access and exchange.”
- “...focus on business practices in the field of statistical information that would allow more efficient processes for exchange and sharing of data and metadata within the current scope of our collective activities.”



Goals of SDMX (2)

- The goal is to create common standards for data exchange and for statistical metadata to gain efficiency and avoid duplication of effort in our own work and possibly for the work of others in the field of statistical information.

Statistical Data Exchange



- Complex and expensive
- Setting up bilateral exchange requires a substantial effort for both counterparties, duplicated for every additional party.
- And duplicated by independent parties all over the world



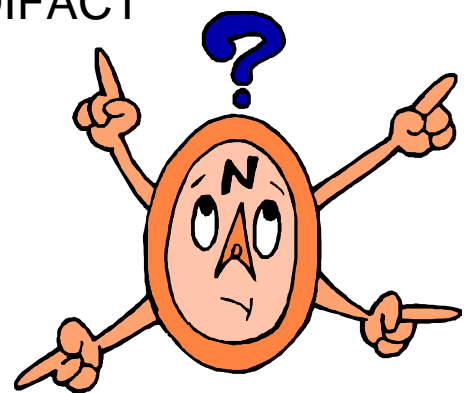
Statistical Data Exchange

Different **formats** of data and metadata

Different **places** to store data and metadata

EDIFACT

XML



paper form

Different **media**

Email

Web-form

dial-up

Structured Files

removable media

file upload

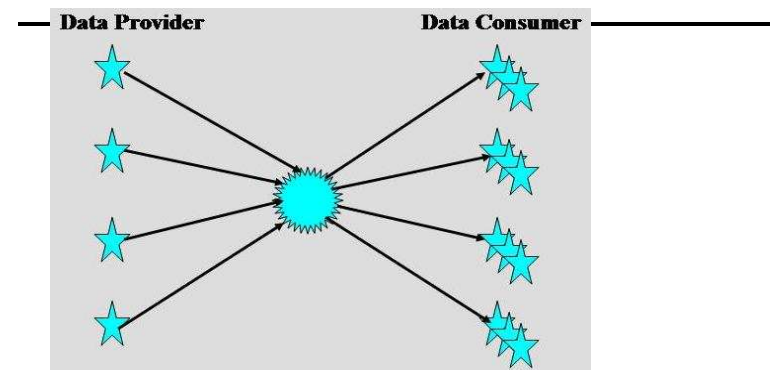
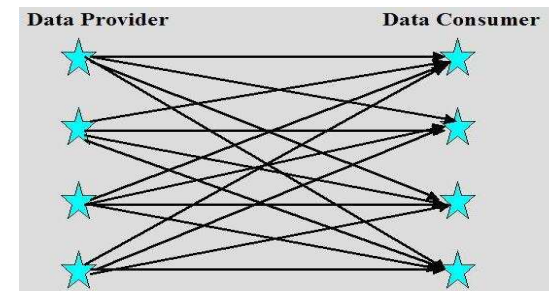
Paper





Patterns of Data Exchange

- **Bilateral**
 - individual provider-to-consumer links
 - could be batch or web dissemination
 - variety of format standards
- **Gateway**
 - many bilateral-to-central authority then on to others (bilaterally)
 - still many bilateral formats
 - typical in data bank exercises
- **Data sharing/multilateral**
 - statistical hubs through common standards and formats
 - data/metadata useable without prior agreement
 - dissemination could replace reporting

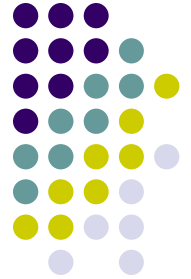


Bilateral Exchange of Statistics



- Reporting burden on NSOs
- Collection and pre-processing burden on international institutions
- Multitude of approaches and technologies
- Every data exchange link requires a significant effort to establish.
- Established links often cannot be reused with other parties.

Multi-lateral Data Exchange



- Requires that a common standard be shared by all parties
- Only one exchange mechanism needs to be set up
- Reduces burden and duplication

SDMX as the Infrastructure



- Standards for:
 - Structuring of statistical data
 - Packaging of statistical data as either XML or EDIFACT
 - Registry of data and metadata
 - Querying the registry
- Specifications, guidelines, tools, and manuals are freely provided to support implementation



Statistical Data Structures

- SDMX specifies how statistical data can be structured.
 - SDMX does not detail specific structures or codes, it only provides a framework for developing those.
- **Data Structure Definitions (DSDs)** provide characteristics of the data to be exchanged.
- A DSD **must** be developed before any SDMX exchange can take place.



Packaging Statistical Data

- Once a DSD has been designed, statistical data can be sent in messages structured as specified.
- SDMX provides several message formats, adapted for different scenarios.
- Validation is supported for most message formats.
- Tools are available for transformation between message types.



SDMX Registry

- Supports Pull data exchange
- Simplifies discovery by providing a central reference
- DSDs and other structural metadata, data sets, and reference metadata can be registered.
- Applications can query the registry for available data and be notified of changes.



Metadata in SDMX

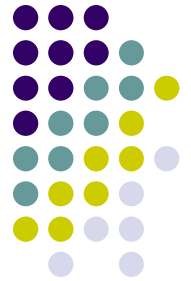
- Can be stored or exchanged separately from the object it describes, but be linked to it
- Can be indexed and searched
- Reported according to a defined structure
- **Metadata Structure Definition (MSD)** defines what objects metadata can describe; how these objects are identified; and the content of metadata.



SDMX and XML

- SDMX uses the e**X**tensible **M**arkup Language for its messages (GESMES/TS is also supported).
- XML is currently the de-facto standard for data exchange, particularly over the Internet.
- Usage of XML simplifies application of many Web-based technologies, such as Web Services.

Implementing SDMX Exchange



- Existing databases need not be modified.
- Software is needed (custom or off-the-shelf) to map database structures and codes to the DSD. The software retrieves data from the database and formats according to the DSD.
 - DevInfo 7 Mapping Tool supports mapping between a DevInfo database and MDG DSD.
 - SDMX Reference Infrastructure allows mapping between any database and DSD.

What Is Needed To Implement SDMX?

