



INTERNATIONAL COLLABORATION USE CASE: THE OECD'S STATISTICAL INFORMATION SYSTEM COLLABORATION COMMUNITY

HLG, 25-27 November 2013
By Jonathan Challener



Community members

OECD

AUSTRALIAN BUREAU OF STATISTICS

EUROPEAN COMMISSION

INTERNATIONAL MONETARY FUND

ITALIAN NATIONAL INSTITUTE OF STATISTICS

NATIONAL BANK OF BELGIUM*

STATISTICS ESTONIA

STATISTICS NEW ZEALAND

UNESCO INSTITUTE FOR STATISTICS

UNIVERSITY OF MANCHESTER



* Pending signature of MoU



Interested organisations

INSEE

Statistics Slovenia

Central Statistics Office Ireland

Tunisian Statistics Institute

Asian Development Bank

UNESCWA

"Rich in functionalities and easy to use"

"Built for ENDUSERS"

"Very good process automation options"

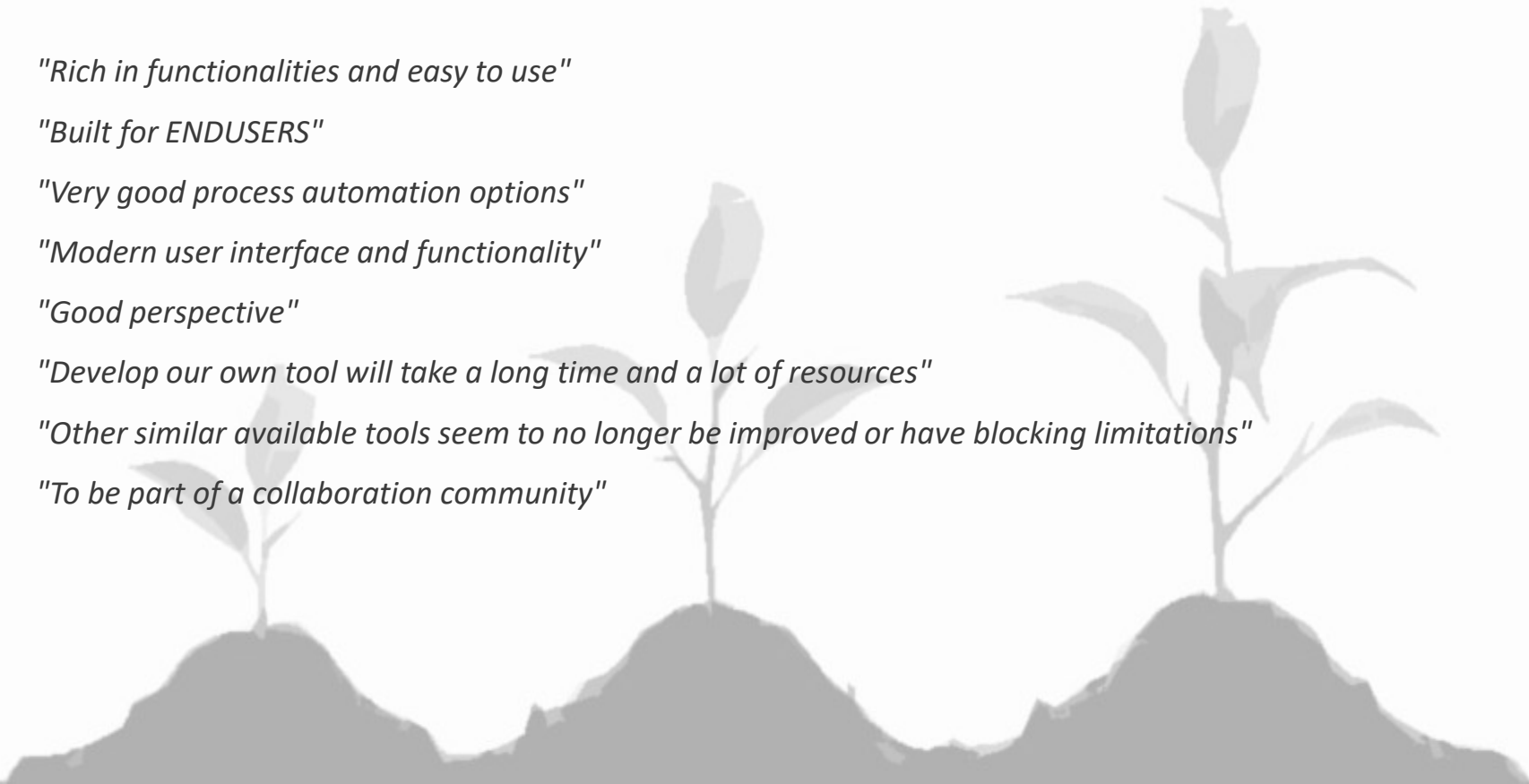
"Modern user interface and functionality"

"Good perspective"

"Develop our own tool will take a long time and a lot of resources"

"Other similar available tools seem to no longer be improved or have blocking limitations"

"To be part of a collaboration community"





Community vision

Provide an international collaboration framework for a

*more open,
more innovative,
and more industrialised data dissemination*

to collectively develop software, leverage innovations, mutualise costs, and promote standardisation.



Community Objectives

Collectively produce and develop software, by leveraging on the OECD.Stat platform and related components, and in so doing build a robust, component-based and scalable architecture.

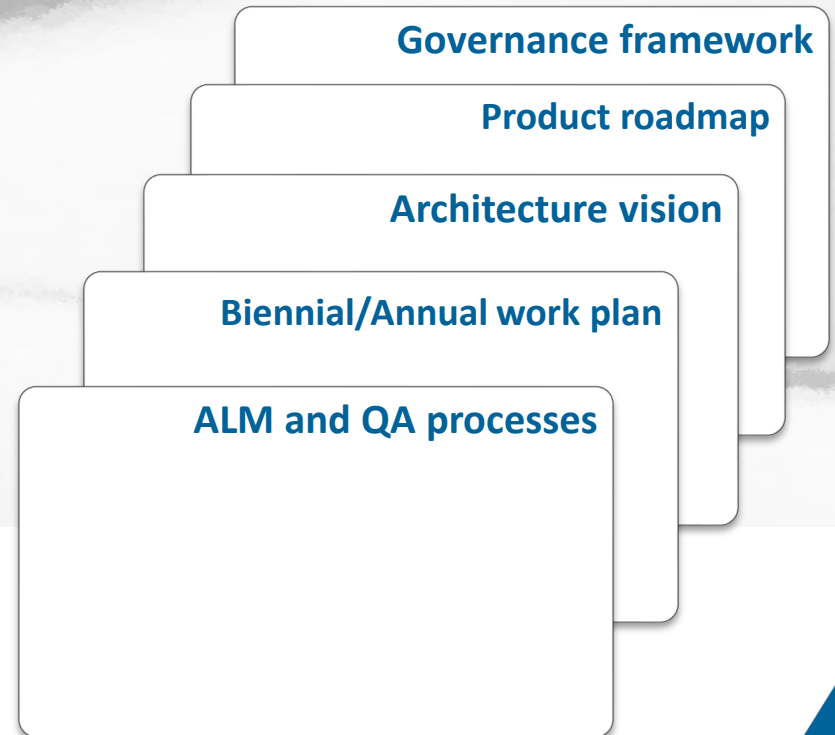
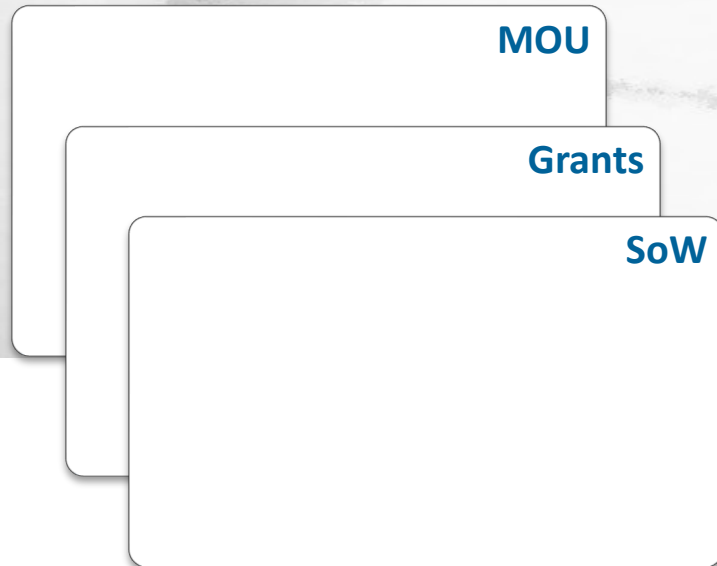
Share experiences, knowledge and best practices through multilateral collaboration and building of a collective capacity, concerning the Community.

Contribute to International Collaboration, by accelerating the implementation of standards and contributing to the international 'Plug and Play' architecture vision.



Community framework

Strategic 5 year directions





Central to collaboration

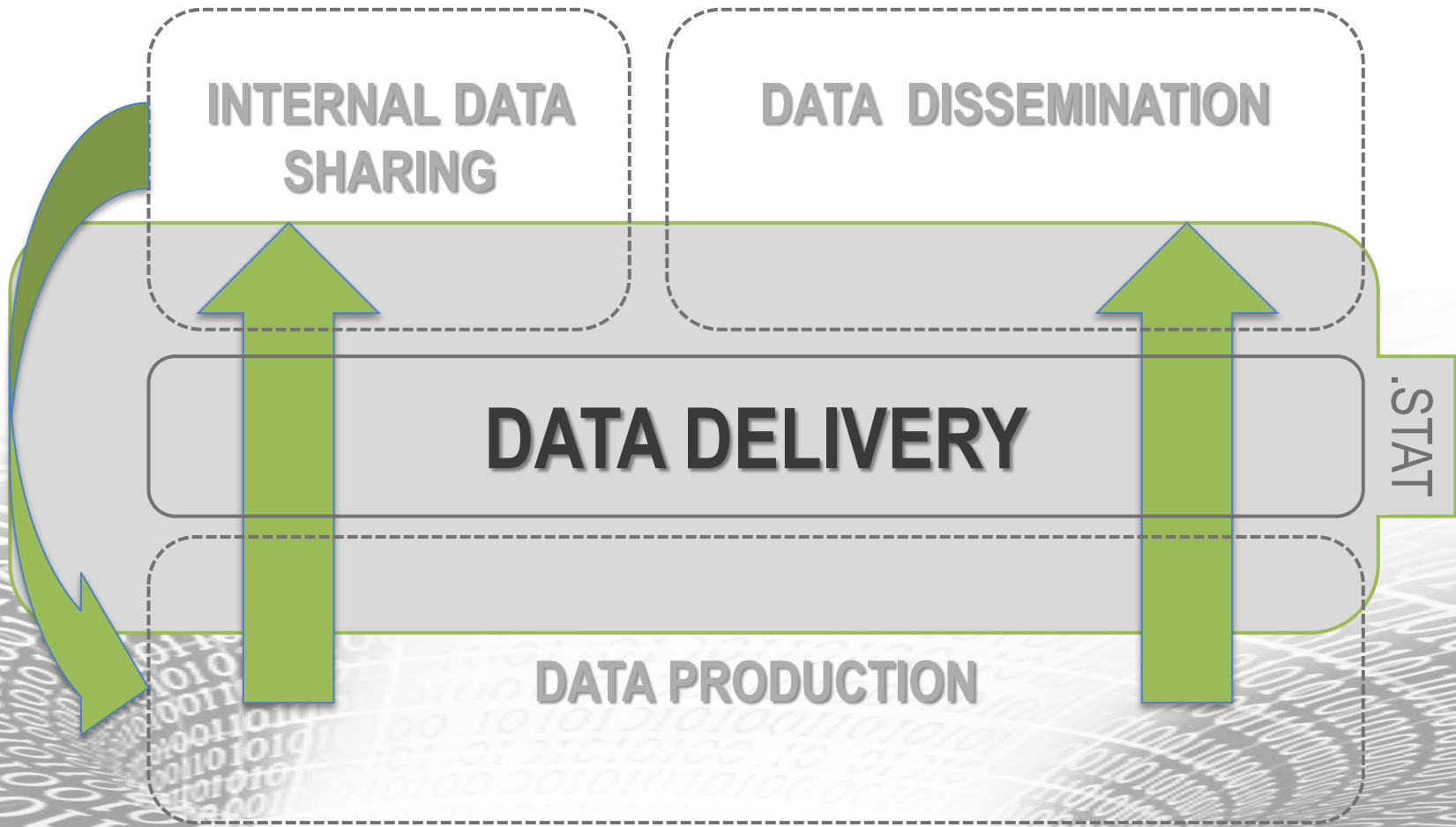
.Stat is the **central repository** ("warehouse")
of validated statistics and related metadata

.Stat is the **central hub** connecting data production, sharing &
dissemination processes

It is the corporate source of data for
data sharing and dissemination purposes



Central hub





Key functions

Internal Data Sharing enabling data consolidation moving away from a fragmented ‘silo’ environment towards a coherent and strategic approach to data.

Machine to Machine Data Exchange making statistical data and metadata “open” via machine-readable formats, in particular the Statistical Data and Metadata standard eXchange format (SDMX).

Data Exploration through a feature-rich data browser providing users the ability to explore and customise data and metadata, visualise it with dynamic charts and download it for further analysis offline.

Streamlined Data Dissemination Processes enabling to structure statistical data dissemination processes and integrate a wide range of data production tools.

International Standards Setting providing a means for disseminating data in the SDMX and more recently the first working API using the common statistical JSON format as defined by the SDMX Technical Working Group

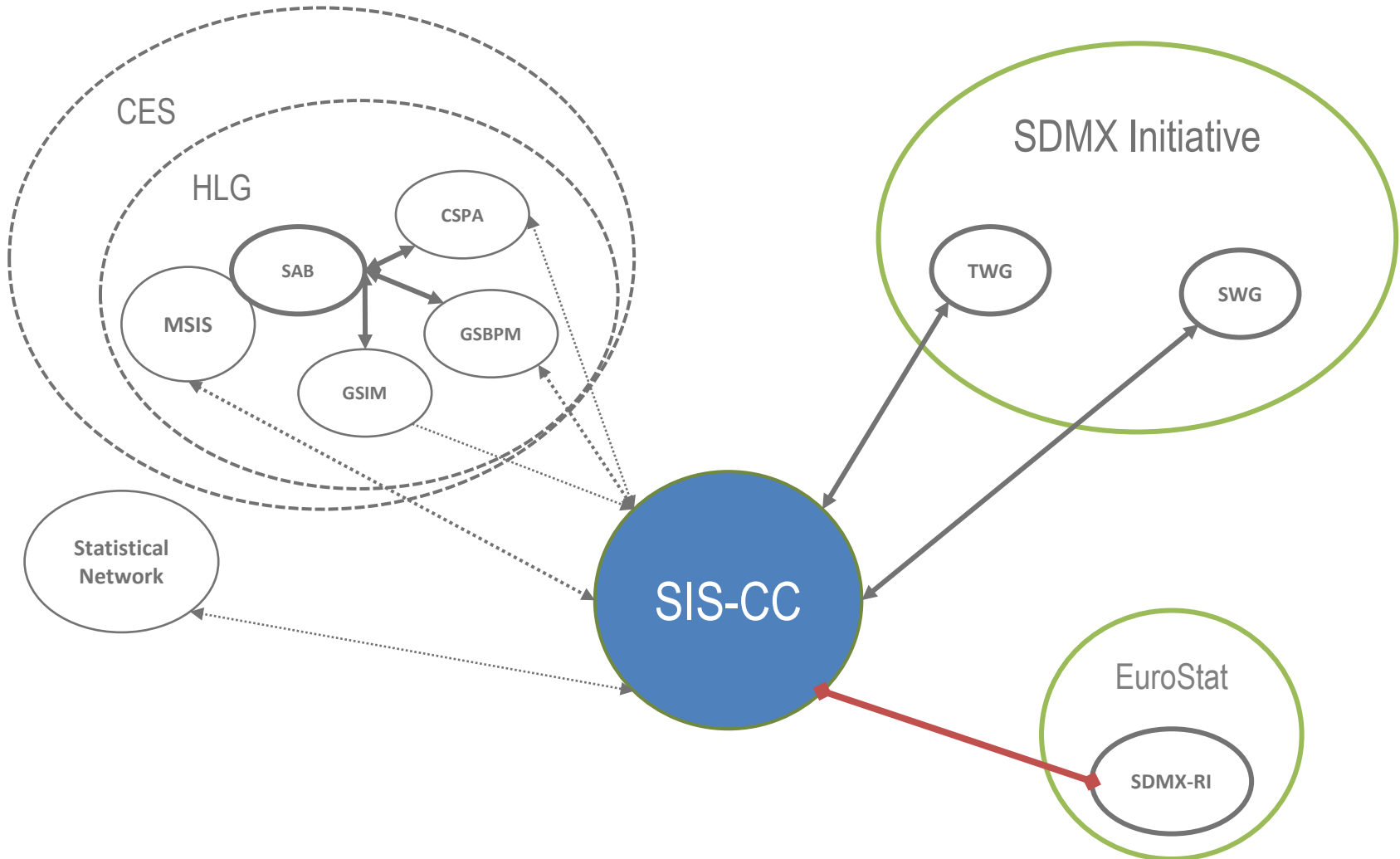


International context



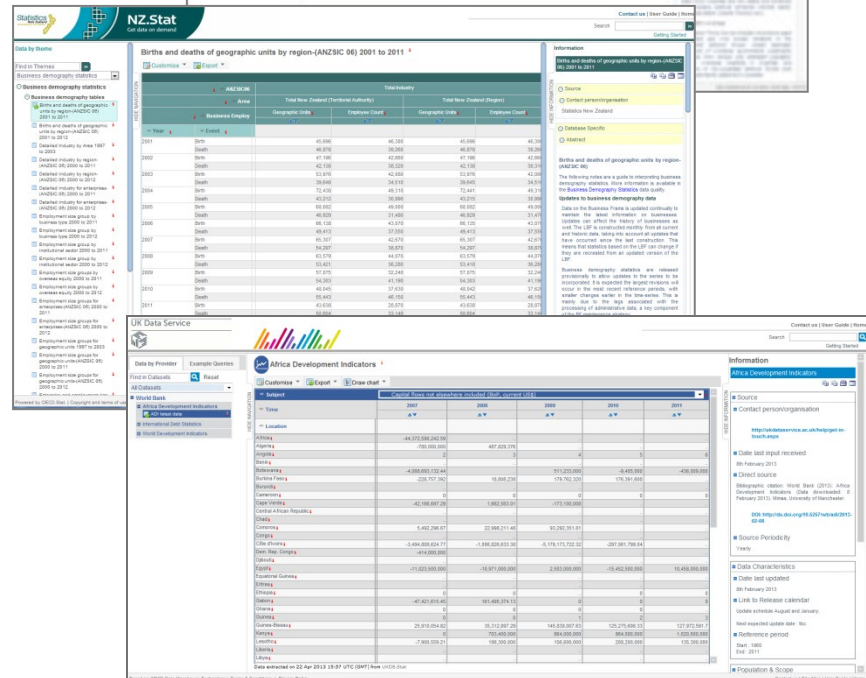
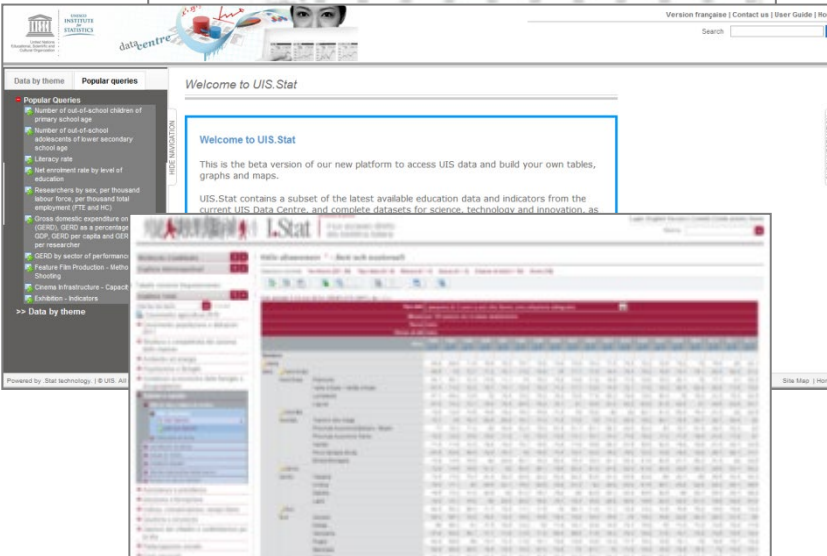
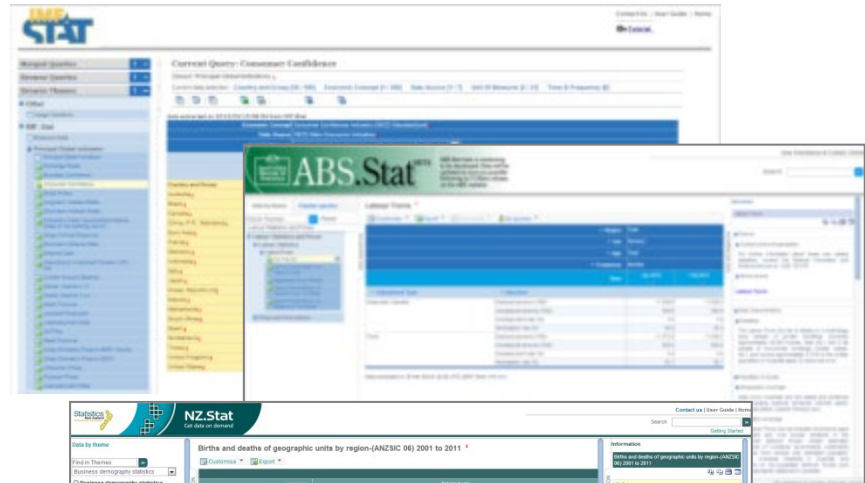


Part of the Overarching International Collaboration Framework





International platform for dissemination





GSBPM v4.0 mapping

Quality Management / Metadata Management								
1	2	3	4	5	6	7	8	9
Specify Needs	Design	Build	Collect	Process	Analyse	Disseminate	Archive	Evaluate
1.1 Determine needs for information	2.1 Design outputs	3.1 Build data collection instrument	4.1 Select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Define archive rules	9.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Manage archive repository	9.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design data collection methodology		4.3 Run collection	5.3 Review, Validate & edit	6.3 Scrutinize & explain	7.3 Manage release of dissemination products		8.3 Preserve data and associated metadata
1.4 Identify concepts	2.4 Design frame & sample methodology	3.3 Configure workflows	4.4 Finalize collection	5.4 Impute	6.4 Apply disclosure control	7.4 Promote dissemination products	8.4 Dispose of data & associated metadata	
1.5 Check data availability	2.5 Design statistical processing methodology	3.4 Test production system		5.5 Derive new variables & statistical units	6.5 Finalize outputs			
1.6 Prepare business case	2.6 Design production systems & workflow	3.5 Test statistical business process		5.6 Calculate weights		7.5 Manage user support		
		3.6 Finalize production system		5.7 Calculate aggregates				
				5.8 Finalize data files				

Stat contributes to
Planned additions



GSBPM v5.0 mapping

Quality Management / Metadata Management

Specify Needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Build data collection instrument	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design data collection	3.3 Build or enhance dissemination components	4.3 Run collection	5.3 Review & validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit & Impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production system		5.5 Derive new variables and units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production system		5.7 Calculate aggregates			
				5.8 Finalise data files			

Archive to be incorporated into the over-arching process of data and metadata management



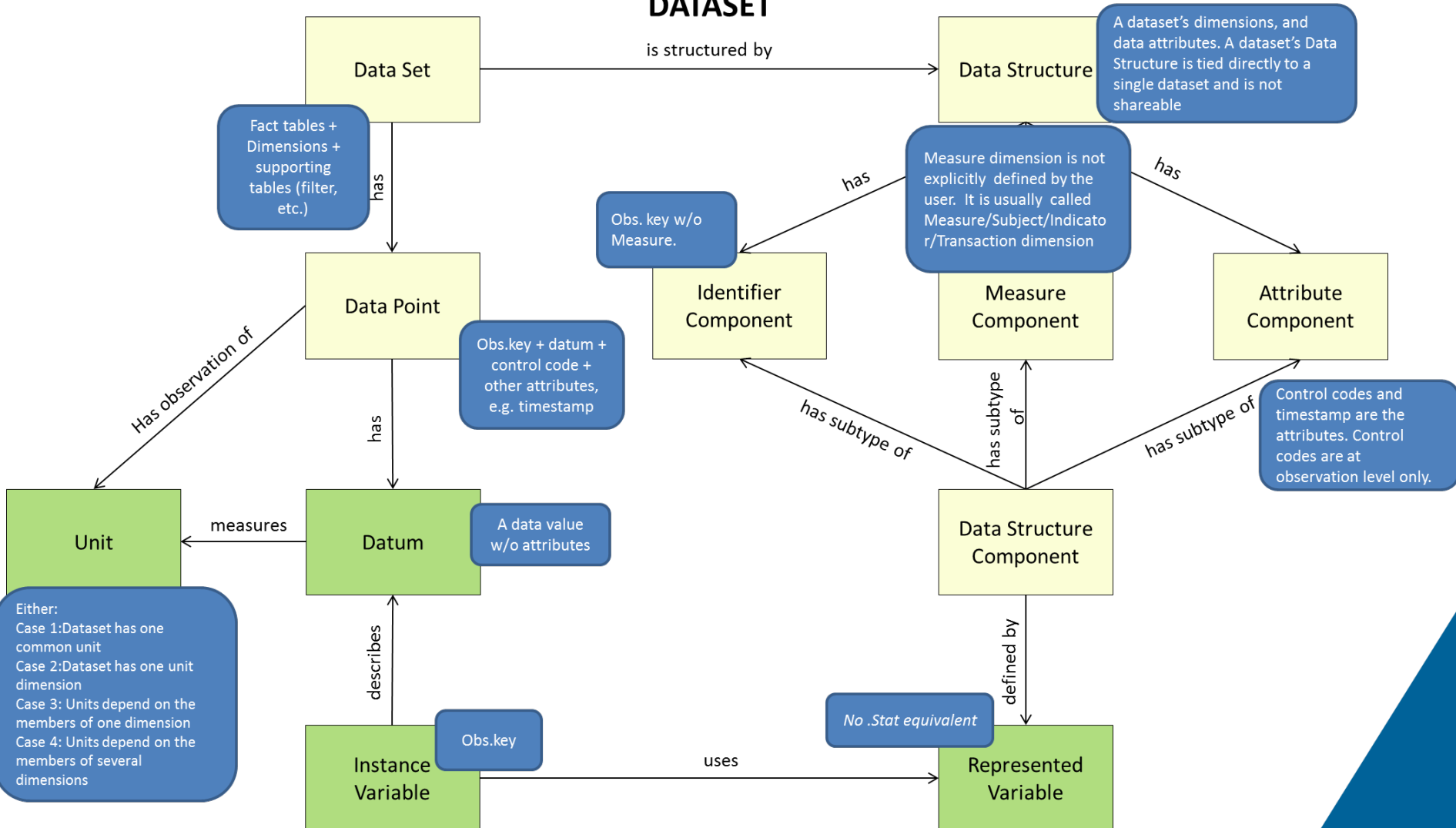
GSIM v1.0 mapping





GSIM v1.0 mapping

DATASET





GSIM v1.0 mapping findings

Forces an analysis of the system which can lead to identifying improvements

Gaps/equivalences can be seen when compared to other systems/standards/methodologies/processes

“GSIM-lite” could speed-up mapping and make for greater understanding

Include an example business process and system mapping in documentation

GSIM mapping registry would be compelling to see examples of how other systems/processes have been mapped as a starting point



For more information contact: David.BARRACLOUGH@oecd.org



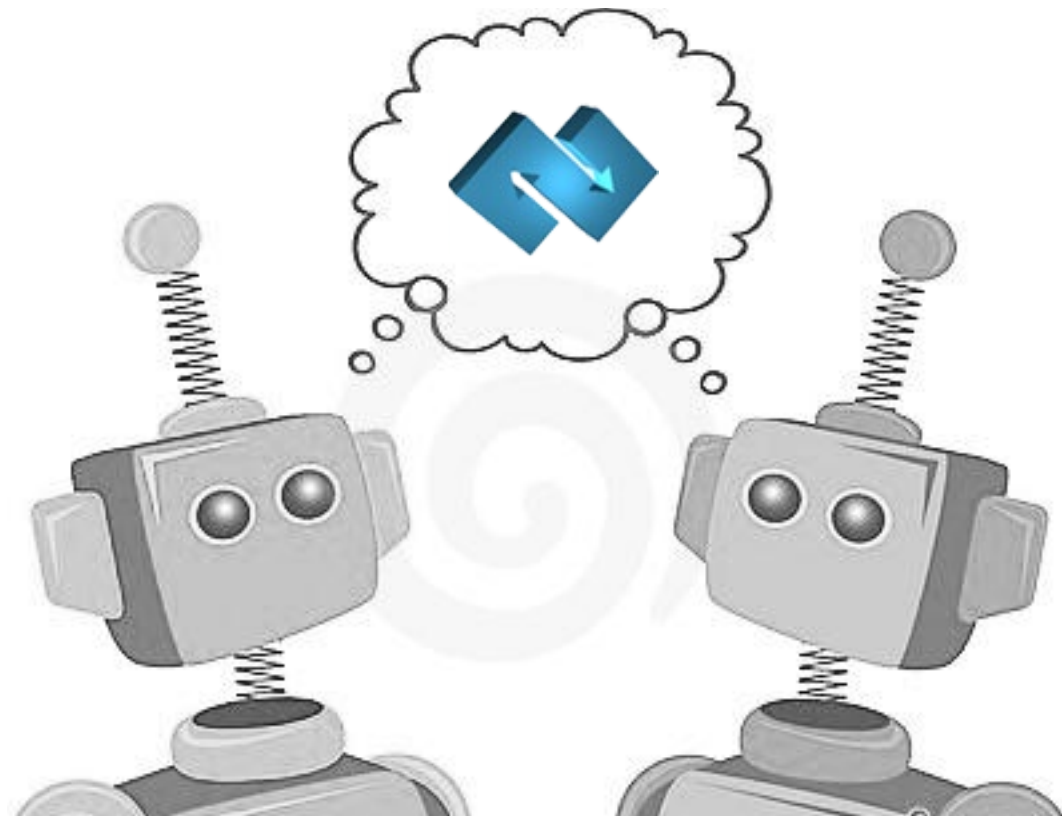
Standard setting and adoption

SDMX Reference Infrastructure integration mapping of multiple data sources to a local source and supporting 2.1 standard

SDMX-JSON API providing an easy to use format for developers to reuse statistical data

Globally agreed DSDs future support through SDMX artefacts

SDMX consumption providing support for both SDMX import and export





CSPA “Plug-and-play”



Decision Principles aligned to the SIS-CC strategic directions

Design Principles at the core of OECD.Stat and SIS-CC work

Requirements aligned to GSBPM, work underway to map to GSIM, and current architecture evolutions will strengthen alignment to CSPA



Plug-and-play “disseminate” candidate?



OECD Stat





Thank you

Questions?

Email:

jonathan.challener@oecd.org

or

dotstatcollaborationsupport@oecd.org