ESA/STAT/AC.322
UNCEEA/11/4b

DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS
STATISTICS DIVISION
UNITED NATIONS

Eleventh Meeting of the UN Committee of Experts on
Environmental-Economic Accounting
New York, 22-24 June 2016

Cover Note for Session 4.b -

Breakout session 1 on Statistical Production Processes and Harmonized Tools

Session Organizer: South Africa

(for discussion)

Cover Note: Statistical production process and harmonised tools

How do we introduce the use of metadata driven statistical production architecture (CSPA and GSBPM) in support of the SEEA/SNA implementation?

Introduction

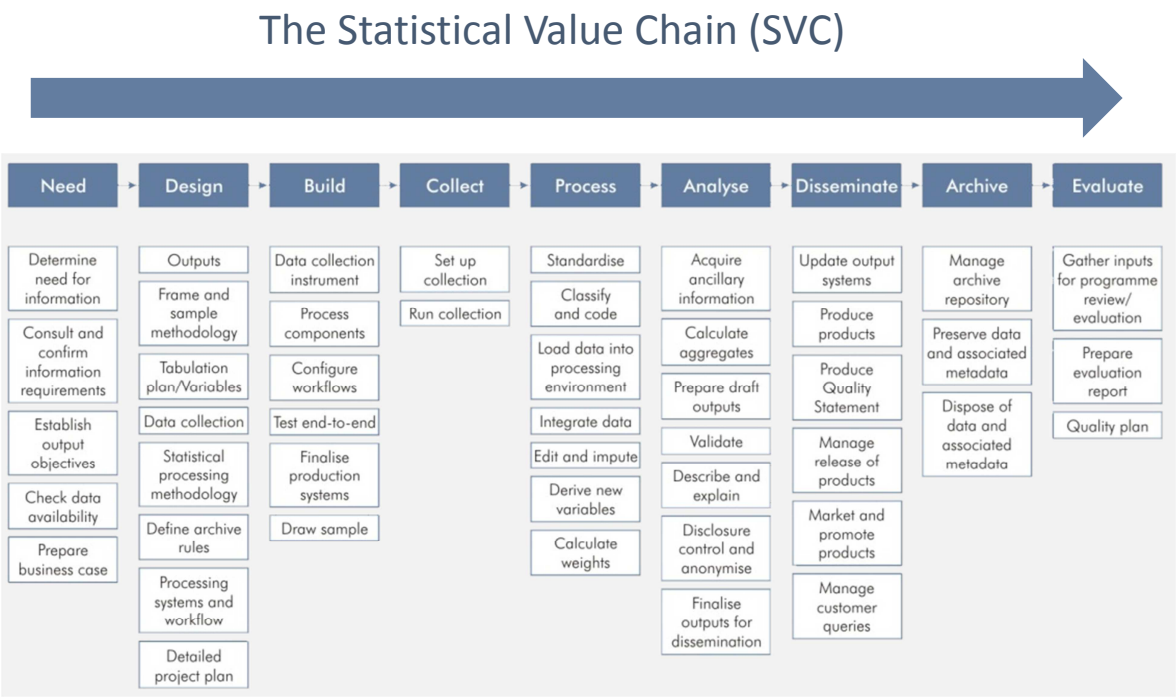
The Generic Statistical Business Process Model (GSBPM) was developed by the UNECE (Current version was released in December 2013). The GSBPM was originally developed based on the Generic Business Process Model from Statistics New Zealand, supplemented by input from other statistical organisations with experience of statistical process modelling.

In the context of statistical modernisation, the aim is to align the enterprise architectures of different organisations, creating an "industry architecture" for the whole "official statistics industry". This approach is intended to facilitate collaboration, sharing and joint development of the components and services that are needed for the different parts of the statistical business process (defined in relation to the GSBPM). The result is the Common Statistical Production Architecture (CSPA), first released at the end of 2013.

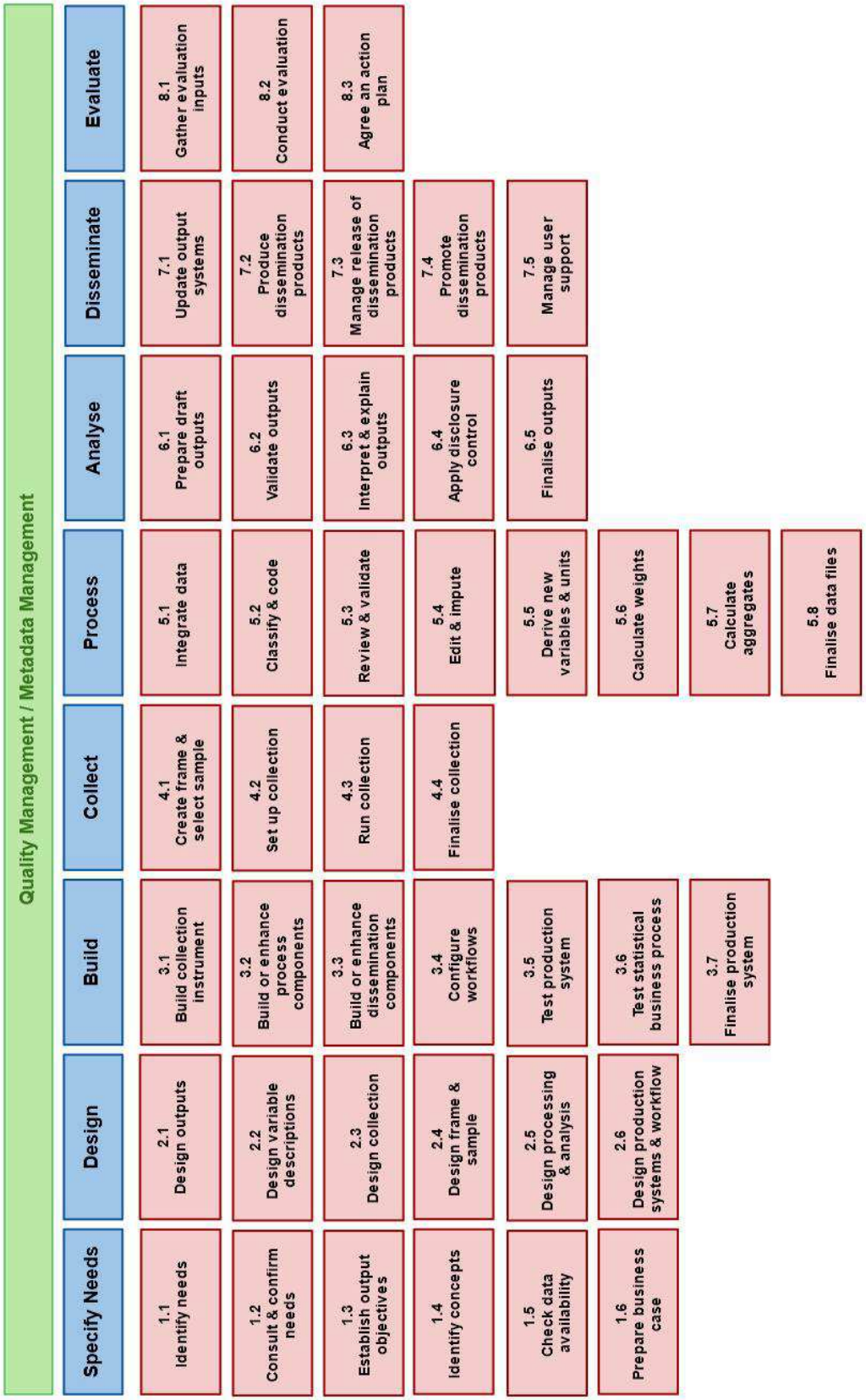
The GSBPM describes and defines the set of business processes needed to produce official statistics. It provides a standard framework and harmonised terminology to help statistical organisations to modernise their statistical production processes, as well as to share methods and components. The GSBPM can also be used for integrating data and metadata standards, as a template for process documentation, for harmonizing statistical computing infrastructures, and to provide a framework for process quality assessment and improvement.

Figure 1 provides an overview of the value chain for statistical operations as is applied within Statistics South Africa.

Figure 1: Statistics South Africa generalised statistical value chain (SVC)



Levels 1 and 2 of the Generic Statistical Business Process Model



Stats SA experience

Stats SA tried to implement an end to end business model driven by metadata, data repositories and triggers for events. The primary goal was to improve the way the Organisation organises itself and deliver statistics that are of quality, however that is defined. Although many gains were made in terms of the development of metadata, standardisation of questionnaires and improvements in the processes, the biggest improvement was gaining corporate agreement on the SVC. But the project didn't deliver as was planned after 4 years. This is primarily due to the following factors:

1. The Organisation tried to go too big too soon
2. The Organisation wasn't mature enough to work in an environment with management and responsibilities for different phases of a value chain (structure by theme vs structure by process).

It involves a change in the working culture of staff, which didn't come easy and the required change management wasn't done well enough. Although 'the DMID project was formally stopped, it lives on in the form of a Business Modernisation division within Stats SA that focuses on improvements such as hand-held devices, electronic questionnaires etc. But the Organisation haven't found a "home" for dealing with modern issues such as big data in a consistent manner across the Organisation.

Going forward.

It would be important to share country experiences on the introduction of CSPA with GSBPM etc. related to statistical value chains, corporate support services and a culture of innovation and modernization in the national statistical office and system

What kind of tools need to be developed (metadata driven statistical production processes, use of standardised questionnaires, approaches for then use of administrative data, etc)?

Introduction

Do new tools need to be developed for SEEA implementation or can the SEEA implementation process make use of existing tools (or a possible combination of both)?

The International Household Survey Network (IHSN) Central Survey catalog provides a searchable list of surveys and censuses conducted in low- and middle-income countries. This catalog is maintained in collaboration with the World Bank and a large number of national and international agencies. This catalog provides rich metadata including, when available, the survey questionnaire(s), manuals and report(s), and list of related citations. It does not provide access to microdata, but when available, provides a link to external catalogs where the data can be obtained.

IHSN provide guidelines and best practices related to all stages of survey implementation, which includes:

- designing survey programs.
- creating survey budgets.
- implementing surveys.

- integrating surveys.
- archiving and dissemination of microdata. Although archiving and documentation of survey data are part of the GSBPM, IHSN provide guidelines in a separate section as this is an area to which the IHSN has devoted much attention and resources.

Possible software tools as described by the IHSN includes:

- DDI Metadata Editor (Nesstar Publisher 4.0.9): The IHSN Metadata Editor, also known as the Nesstar Publisher, is a specialized XML editor compliant with the Data Documentation Initiative (DDI) 2.n and the Dublin Core metadata standards;
- Microdata Cataloging Tool (NADA):NADA is a web-based cataloging system that serves as a portal for researchers to browse, search, compare, apply for access, and download relevant census or survey information. It was originally developed to support the establishment of national survey data archives; and
- Statistical Disclosure Control (SDCMicro): SDCMicro is free, R-based open-source package for the generation of protected microdata for researchers and public use. Data from statistical agencies and other institutions are mostly confidential. This package can be used for the generation of anonymized (micro) data, i.e.

Stats SA experience

Stats SA developed and uses the South African Quality Assessment Framework (SASQAF) in assessing both internal and external statistical data (and administrative data) as part of the National Statistical System (NSS). Data is assessed on the following criteria: relevance, accuracy, timeliness, accessibility, interpretability, methodological soundness and integrity. Stats SA successfully uses Nesstar as a software tool for data dissemination.

Adherence to the different SASQAF dimensions are used as the guiding principle internally when new surveys are developed, data collected and finally disseminated. This is supplemented by organisational structures such as questionnaire clearance committees and others.

Going forward

For the SEEA implementation, a similar survey catalog could be developed for sharing country general business survey and census forms as well as specialized forms used as collection instruments for different modules (e.g. water, energy) as well as sharing experience in extending the coverage of business registers variables for sampling purposes. In addition, experience should be shared in the use of administrative data sources.

Similarly, experience should be shared on specific software tools.

Could software tools be provided to countries (e.g Eurostat's Physical Energy Flow Accounts builder) to compile simplified accounts?

Introduction

Eurostat provides an IT-tool ('PEFA (Physical Energy Flow Accounts)-builder') facilitating the compilation of PEFA tables A (physical supply table for energy flows), B (physical use table

for energy flows), and D (key energy indicators) starting from energy statistics (IEA/Eurostat annual energy questionnaires).

This approach has benefits as well as limitations. The benefits could include:

- Standardised accounting structure for countries with limited national environmental databases;
- Increased use of global databases for country specific account compilation (in the absence of a national database for a specific resource); and
- Assistance to countries in the start-up process of SEAA implementation and development (comparison of software tool results versus actual EEA account compilation).

The following limitations must be kept in mind when using software tools for EEA compilation:

- Institutional knowledge of the SEEA and EEA compilation – software tool replaces country specific SEEA implementation and EEA compilation institutional knowledge?;
- Institutional understanding of the source data used for EEA compilation (methodology used, data limitations, etc);
- Data interrogation of the source data by the EEA account compiler – will they really know the quality (and limitations of) of the source data used?; and
- Country trust in the global database used.

Stats SA experience

South Africa's Energy Accounts are based on the Energy Balances that is provided by the Department of Energy (DoE). The DoE has an obligation to the International Energy Agency (IEA) in terms of the methodology and classifications.

Going forward

A technical committee could be created under the UNCEEA that could become the focal point in advancing the use of common statistical production architecture and sharing of data collection and software tools to support the implementation of the SEEA in countries. A common international repository should be created for this purpose.

Sources:

http://www.ihsn.org/HOME/survey-catalogs
http://www.ihsn.org/HOME/guidelines
http://www.ihsn.org/HOME/software
http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model
http://ec.europa.eu/eurostat/web/environment/physical-energy-flow-accounts