



Metadata Management in the FAO Statistics Division (ESS)

Overview of the
FAOSTAT / CountrySTAT approach
by
Julia Stone

Metadata Management in ESS

1. Introduction
2. FAOSTAT metadata collection
3. CountrySTAT metadata collection
4. Improvements in the use of metadata
5. FAOSTAT/CountrySTAT metadata structure
6. Metadata components
7. New metadata characteristics

CountrySTAT

CountrySTAT, which is part of the FAOSTAT family, has been redeveloped in 2012 to benefit from the same technology recently implemented for the new FAOSTAT system.

The CountrySTAT system has been migrated to the FENIX platform which is an internet-based system that provides modularity allowing it to meet different user requirements and adapt to support global, regional and national level projects and initiatives.

One of the key benefits is that the system is now based on open source technology that allows significant improvements in the way we manage the metadata collected from countries.

Metadata collection

- **FAOSTAT**: sends regular questionnaires to the countries to collect data and metadata. The statistical metadata requirements are:
 - Focal point of National Institution who is responsible for data collection
 - The geographical coverage
 - The source of data
 - Frequency of collection
 - Dissemination methods
 - Timeliness of dissemination

- **CountrySTAT**: metadata is entered through online forms so that each statistical table must be accompanied by two sets of metadata.

Operational Metadata

- Title of table
- Description of labels of variables and description of values
- Measurement unit
- Matrix code
- Time reference
- Footnotes highlighting additional information on the data
- Source of data (Responsible National Institution)
- Explanation of symbols in tables
- Information about copyright
- Date of update
- Contacts for additional information

Methodological Metadata

- Focal Point of National Institution
- Funding
- Statistical population, geographical coverage, observation unit, classifications applied
- Description of methods used in collecting, revising, calculating and estimating the statistics
- Description of data quality, including information on error sources and accuracy of the statistics
- Comparability with alternative sources, etc.

The operational metadata are the attributes of the table specifying the information needed by the software to make queries possible

The methodological metadata are fundamental information required by users to understand the methodology adopted by the institution

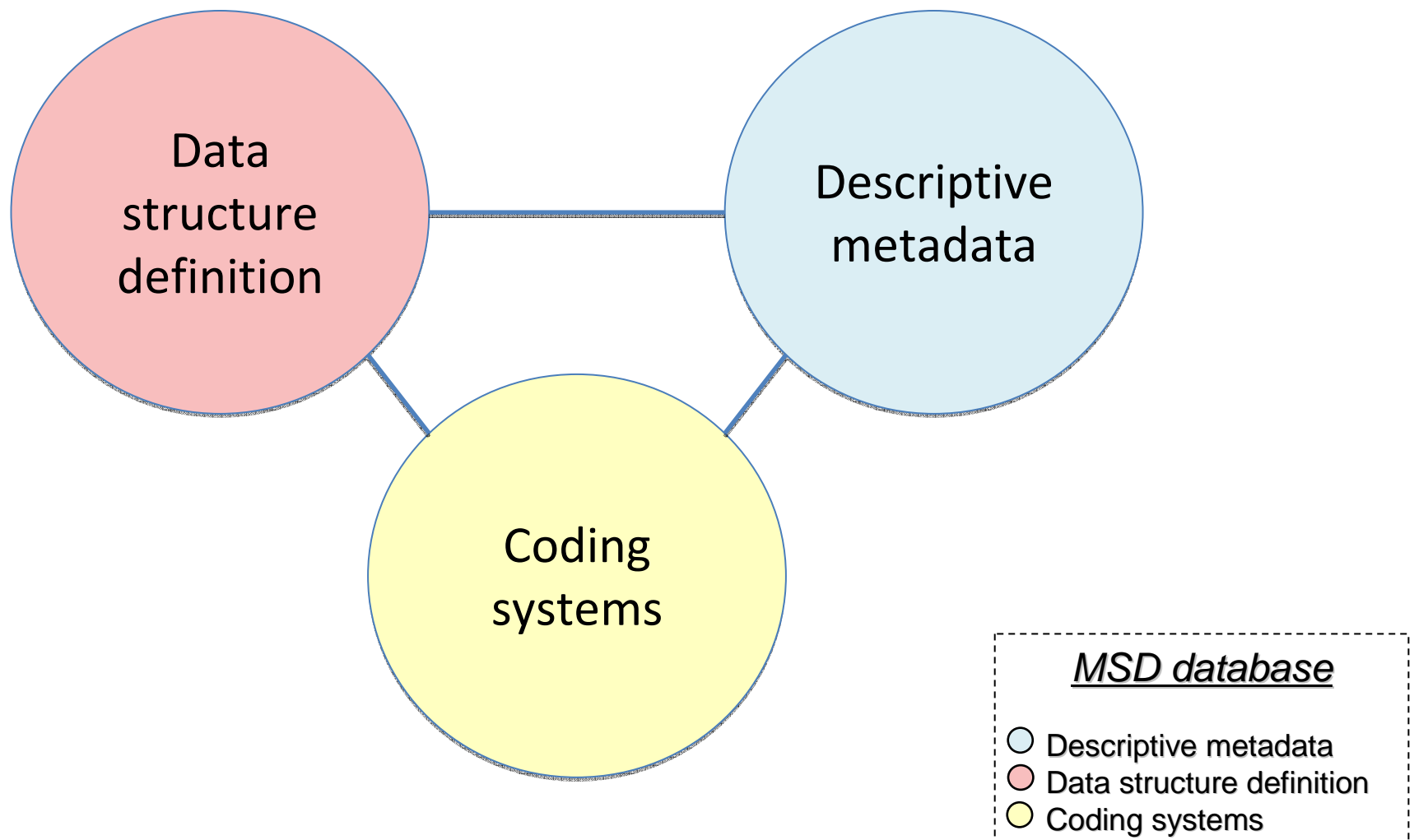
Improvements in the use of Metadata

FENIX metadata services provide support to a number of activities, such as:

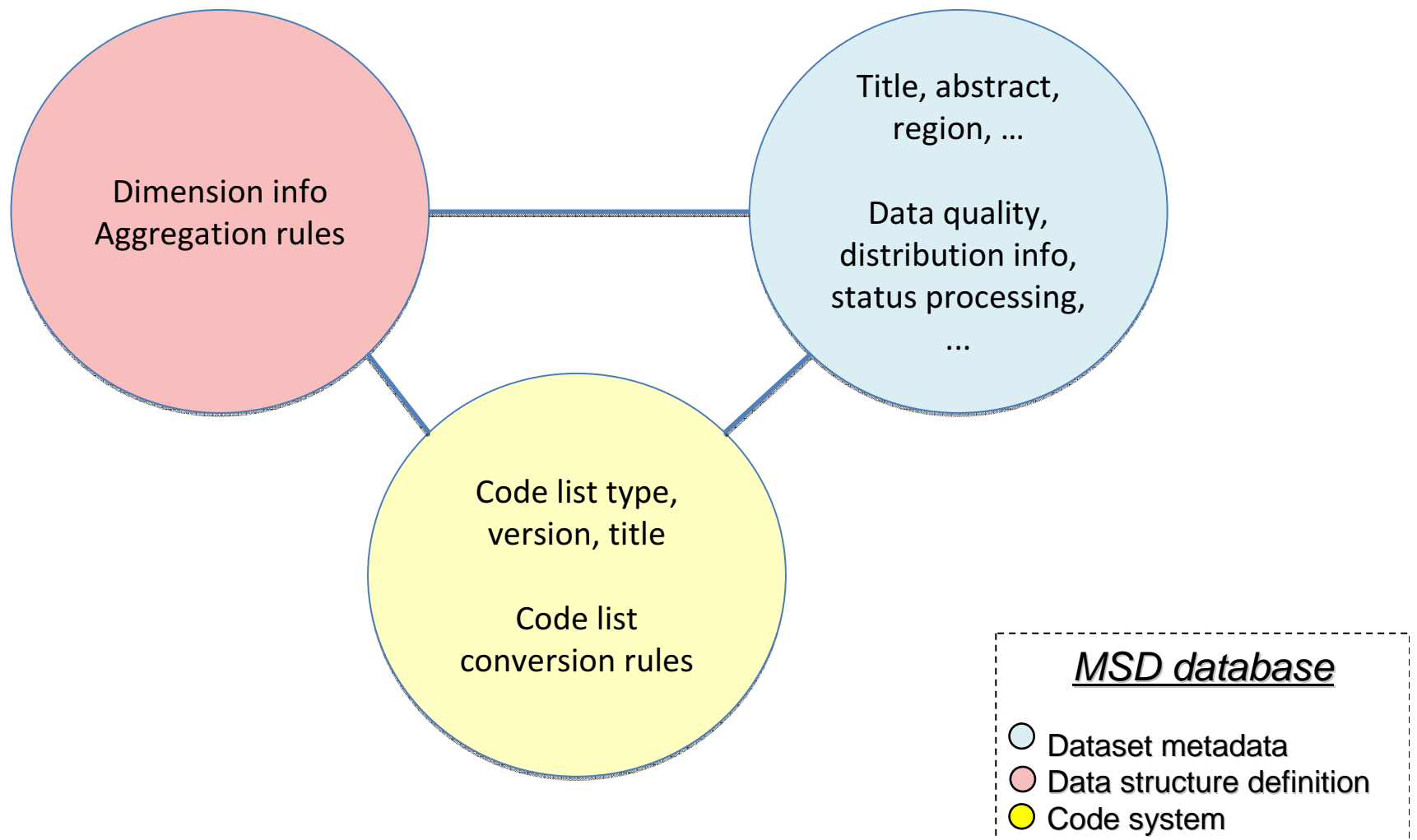
- Distributed data discovery services inclusive of full-text & semantic cross domain search
- Support to functionalities like OLAP, charts and maps
- Definition and enforcement of data policy and governance rules
- Support to data analysis and data mining (e.g. by making explicit the relationships between datasets)
- Compilation of metadata information by different stakeholders through online data management tools

Metadata structure

The Metadata Structure Definition (MSD) consists of three main areas:



Metadata structure



Metadata components - Business logic

- **Manage relationships** between different coding systems (manually and/or by using import services)
- Coding systems may **establish temporary joins** to show data from different domains
- Transcoding (**translation of codes** in another coding system)
- Establish **rules to perform aggregations**
- **Manage the units of measurement** through specific algorithms

New metadata characteristics

- Includes many SDMX elements to facilitate compatibility with SDMX-based applications
- The structure is compliant with the ISO 19115/139 standard specifications (code lists might be different)
- Easily extendable to include/modify elements or groups of elements

Thank you!

Focal point for contact

Yasin Tabit Ahmed
tabitays@yahoo.com

Concepts, definitions and classifications

Other [teff_commodity%20description.doc](#)

Concepts and definitions adopted at international level are those of FAOSTAT except for the classification of one product: Teff. Teff crop is one of the major staple food crops in Ethiopia and it cannot be classified under any other cereal crop. It should stand alone as one of major cereal grain crop when it comes to Ethiopia. The price of teff has been collected and reported in three varieties, (i.e. white, mixed and brown) since they have significantly different prices in the market.

Geographical coverage

Country

Administrative Lev 1

- Region

Surveys

Annual Agricultural Sample Survey

The CSA has been collecting data on cropland area, production and productivity for about three decades using probability based sample surveys from agricultural households under the umbrella of Rural Integrated Household Survey Program (RIHSP).

The agriculture statistics data collection on area, production, yield and farm management practice is conducted twice annually in order to capture both seasons agricultural practice of Ethiopia. These are: main season (Meher) and short rain season (Belg) survey.

Sample Design

- The primary agricultural statistics data are collected from two sources: small holding farmers and commercial farm enterprises.
- Farm enterprises are used as sampling units for commercial farm category and agricultural households (holders) are the ultimate sampling units in the case of small holding farmers.
- A stratified two-stage cluster sample design has been used to select samples of agricultural households. The agro-ecological zones and administrative divisions of the country is used to form the strata.
- Within each stratum the enumeration areas (EAs) are used as primary sampling units (PSUs) and agricultural households are secondary sampling units.
- The AgSS data is collected from more than 65,000 households in 2,280 sample areas and from over 1,700 large and medium scale commercial farms.
- The survey findings reports are produced for each administrative zones of the regional states and also aggregated at regional and country levels.

Data Collection methods

In Ethiopia, farmers are not able to provide their land size and their product in standard units and it is also difficult to convert from local units to standard units since these local units size are very much vary from area to another.

Therefore, the CSA uses objective methods to collect data on cropland acreage by applying rope and compass method and hand held GPS for area measurement and crop cutting exercises on sample plots of each crop is also done for crop yield estimation.

Direct interview method is used for collection of information on crop yield conditions, farm inputs, practices, land use, product utilization, livestock characteristics, etc.

Contents of AgSS

The annual agricultural sample survey is comprised of several questionnaires/modules administered to households at several times:

- Crop production forecast survey (cropland area, condition of crop yield, etc.) September - October.
- Agricultural practices and general information (fertilizer, pesticides, irrigation, soil conservation, etc. September -October.
- Cropland area measurement. October - November.
- Livestock survey (number, type, age, purpose, milk, egg, honey, animal, veterinary information). November
- Crop cutting exercise (yield). October- January.
- Utilization of crop and livestock product (percent used for consumption, sale, stock, etc.): January- February.
- Commercial farm survey (area, production, farm management practices, diary and fattening): February - March).
- Short rain season (Belg) agriculture survey (area, production, yield, farm input utilization, practice, etc.) April-June.

Statistical Reports of Annual Agricultural Sample Surveys

Based on annual agricultural sample survey, the nine statistical reports have been produced and disseminated annually. They are:

- Crop Production Forecast: November-December
- Area and production: April
- Farm management practices: May
- Land utilization: May
- Crop and livestock product utilization: May
- Livestock and Livestock characteristics: January-February
- Large and Medium scale commercial farms: June-July
- Belg season, area and production: August
- Belg season, farm management practices: August.

These survey findings reports are published in hard (bulletins) and electronic copies.

An example is the "[Agricultural Sample Survey 2007-2008 \(2000 E.C.\)](#)"

Reports are also available on the CSA website (<http://www.csa.gov.et>).