

Metadata and classification system development  
Bosnia and Herzegovina

*Project BA 08-IB-FI-03*

## MISSION REPORT


on

### **3.2.2 Activity Metadata system I**

Mission carried out by  
Mogens Grosen Nielsen, Statistics Denmark  
and  
Carlo Vaccari, ISTAT

Banja Luka, 29 June 2012

Version: *Draft*

		 <b>STATISTICS DENMARK</b>
		Statistics Denmark

**Author's name, address, e-mail**

*Mogens Grosen Nielsen  
Sejrøgade 11  
DK-2100 Copenhagen Ø  
Denmark  
Tel. +45 3917 3894  
mgn@dst.dk*

**List of Abbreviations**

BiH	Bosnia and Herzegovina		
BHAS	Agency for Statistics of Bosnia and Herzegovina		
FIS	Federal Institute of Statistics of the Federation of ' Bosnia and Herzegovina		
RSIS	Republic Institute for Statistics of the Republika Srpska		
GSBPM	Generic Statistical Business Process Model		
DDI	Data documentation initiative		

## Executive summary

This mission report was prepared within the state and entity statistical institutions in Bosnia and Herzegovina. It was the second mission to be devoted to topic 2: “Metadata and classification system” within Component 3 “Strengthening the Institutional Capacity of BiH Statistical System” of the project concerning Metadata and Classification System Development for the BiH statistical system.

### 1. Issues and recommendations

#### 1. Standards and tools

The use of standards and tools (DDI, Neuchatel, Metaplus etc) were discussed. Solutions in Sweden, Portugal, Denmark and Australia were presented and discussed. It was apparent that all standards in more or less complicated ways can handle information on statistical units, questions, variables, concepts and classifications. It is recommended to investigate the DDI-standard. A pilot study on using DDI on SBS should be conducted.

*Status:* The tool Colectica and repository were installed. A small test-survey has been put into Colectica. Training on Colectica and DDI was done as part of the mission.

*Recommendation:* 1) Continue with work on testing DDI with SBS data 2) include a description of possible standards and a recommendation in the section on standards in the metadata-strategy.

#### 2. Generic Statistical business process model (GSBPM)

There are many differences in business processes depending on which people are involved. This makes it difficult to build common metadata-system and other common systems. It is recommended to adopt GSBPM as a common generic business process model. As a first step the GSBPM model should be translated. On the next mission a strategy for implementing GSBPM should be prepared.

*Status:* The GSBPM-model was introduced and discussed.

*Recommendation:* The consultants prepare a first draft version for the strategy (2-4 pages) with a) introduction and benefits b) suggested objectives and result c) organisation and d) proposal of main steps in the implementation. The following steps were agreed upon at the mission:

- Establish Business Process and Metadata Group (see suggested organisation in issue 9)
- Translate and review the full document by The Business Process and Metadata Group
- Disseminate knowledge on GSBPM
- GSBPM must be used where possible. This requirement should be included in all relevant strategic documents
- Official adoption
- Documentation of surveys using GSBPM (as-is)
- Decision on future work (to-be)

The translation and review of GSBPM should be finalised before next mission. It is recommended that the next mission includes meeting with directors with presentation and discussion of GSBPM and Metadata Strategy.

The following initiatives should be considered for the dissemination of knowledge of the GSBPM:

- Present GSBPM together with metadata strategy for staff at institutions.
- External presenter from METIS/UNECE
- Share translation: how? Intranet? Web?

### 3. The current situation on metadata

Metadata today are stored locally with no common structure. A template for describing business processes were tested for the purpose of getting an overview of metadata. SBS were used pilot. It is recommended to use a simple template with the purpose of getting an overview of metadata, data and application. It is recommended to collect the following information for every survey: a) data and application and b) metadata (Methodology and quality metadata, Concepts, variables, Variables and Classifications/codelists). The collection for each survey must show the situation in BHAS, FIS and RSIS.

*Status:* The work on producing the list was started

*Recommendation:* The work should be continued.

### 4. Location and responsibility of the common metadatabase

According to the contract: "Metadata and classifications have to be stored and updated in one place: the Agency." RSIS stressed that the database should not be centralized, since not all metadata would be shared among the three parties. The Agency and FIS stressed the importance of having metadata stored only one place. There was no agreement on this. It is recommended to build a centralised common database in order to move towards a situation where metadata can be established as a common resource. It should be possible to update metadata by and from all three parties.

In order to make sure that all three parties work in the same direction is recommended that each part is chairing the process of data and processes for 1 year and after one year the responsibility moves to the next part. If one of the parties has their own surveys they can use the common metadata system but it is not mandatory. It is recommended that the same standards are used as much as possible.

*Status:* same

*Recommendation:* same

### 5. Differences in classifications and code-lists in BHAS, FIS and RSIS

Regarding classifications there are only few differences. Regarding code-lists there are more differences. It is recommended to set up a working-group that first describe the current situation and afterwards do the harmonization.

*Status:* a plan for description and harmonizing code list was prepared.

*Recommendation:* The results of the work should be part of description of current situation in the metadata strategy (summary and annex). The plan should be coordinated with other metadata-activities by the Business Process and Metadata Group.

## 6. Differences in legislation

Due to differences in legislation in BHAS, FIS and RSIS, there are differences in variables in administrative sources. It is recommended that statistical organizations are involved in definition of administrative files including discussion of metadata in order to fulfill user needs from statistics. It should also be noted that using administrative sources requires work on defining statistical concepts and do necessary harmonization.

*Status:* same

*Recommendation:* same

## 7. Old and new coding systems

What to do with historical surveys. It must possible to include historical metadata in the coming metadata-system. It is recommended to asses the needs for historical metadata and hereafter decide on the extent of historically metadata to include.

*Status:* same

*Recommendation:* same

## 8 How do implementing metadata standards affect our work?

There is both an organizational and a technical challenge. The organizational challenge should be handles in the coming organization and the implementation of GSBPM. Regarding the technical challenge: In the short run: metadata should be extracted from existing systems and put into a common metadata-system. In the long run it is recommended to establish a library with concepts, variables, statistical unit-types, questions etc. and use these as building blocks when you create a survey.

*Status:* same

*Recommendation:* same

## 9. Human resource and organization

There is a lack of resources in all three organizations. It is recommended to ensure knowledge is shared among IT-staff. A list of activities should be drafted and discussed and put in the strategy. The top-management must approve the strategy and ensure the necessary resources are available. As part of the establishing a metadata-system subject-matter and methodological staff is important players as they define requirements.

*Status:* same

*Recommendation:* An organization diagram including short decriptions on business process and metadata was prepared. The Business Process and Metadata group are (among other things) responsible for management and operation of the metadata-support process. Diagram and description of the organization should be included in the mission-report.

## 10. Inclusion of methodological and quality indicators in the metadata system

Today the quality report is produced and used after end of production. It is recommended to include a top-level methodological document and quality-declarations in the coming metadata-system. The work on methodology and quality indicators should be an integrated part of needs design and build processes. The reason for this is that quality and methodology information is valuable both in relation to discussion with external and internal user. For the technical solution it is recommended to publish / present metadata in an integrated way using ideas of presentation from Portugal ([www.ine.pt](http://www.ine.pt)) and others.

*Status:* The technical solution was discussed

*Recommendation:* 1) include quality elements as custom field in Colectica. (assumption: DDI and Colectica is used as standard and tool) 2) use element form quality report / guide-lines as agreed by all institutions as a part of the work on long-term strategy.

11. The reports from SIDA were evaluated. The recommendations have been taken into account during the mission including setting up an outline of the coming metadata-strategy. SIDA suggested an organisation with content- and metadata-staff forming groups. It is recommended to further discuss the recommendations from SIDA. Regarding organisation it is recommended to establish a metadata-process as a supporting process to the statistical processes in GSBPM. Discussions on how to do this will continue on the next mission.

*Status:* Same

*Recommendation:* See recommendation 9 on human resource and organization.

## 12 User needs

Often metadata systems are only seen as a technical system used by IT-staff. The development and discussion of metadata today has a much broader perspective. Metadata-systems should support all users in their work-processes. It is similar to how the catalog system in the library in the old days helped to inform about content of books and where and how to find a book.

All users of statistical metadata should have the following benefits:

- Better search, retrieval, exchange of data and metadata
- Common terminology, names and descriptions for standard metadata
- Improved efficiency through central metadata repositories
- Better quality statistical information
- Improved interpretability of statistics

Subject matter statisticians should have the following benefits:

- Consistent store of concept, variables and classifications that can be used in new surveys
- Access to related surveys (resource for new employees and for statisticians in other fields)
- standard processes with common method for the creation and use of metadata
- who is making use of which metadata; content information sharing between subject matter areas

*Status:* See above

*Recommendation:* 1) Prepare a list of data-request: variables etc (Business process and Metadata group) 2) Meeting with internal and external users on next mission 3) Prepare and present examples on the use of the coming metadata-system. See issue 16.

## 13 Metadata-strategy - functions and priorities

An introductory meeting with RSIS about user needs was conducted. The meeting showed that it is of great importance to include users in the work on metadata, since they can give input to what kind of practical problems the metadata-system should help to solve.

The following general functions of the coming metadata system were discussed:

1. Documenting data collection, storage, evaluation and dissemination. First step: documentation and harmonization of existing surveys with suggested metadata-element

2. Unifying statistical terminology to improve communication between managers, designers, subject-matter statisticians, methodologists, respondents and users of statistical information systems. First step: codelist harmonization
3. Disseminating statistical information to end users. End users need reliable metadata for searching, navigation, and interpretation of data.
4. Managing communication with end-users of statistical outputs and gathering of user feedback.
5. Planning, designing, implementing and evaluating statistical production processes.
6. Managing statistical data sources and cooperation with respondents.

The functions were selected from [“Part A. Statistical Metadata in a Corporate Context”](#) from UNECE.

*Status:* see above

*Recommendation:*

- Suggested prioritization: The functions above are placed in prioritized order.
- Create and discuss practical examples with all user groups (internal and external)
- Coordinate with Dissemination Units.

#### 14 Metadata Strategy - Timetable

The suggested timetable in the metadata-strategy was discussed.

*Status:* See above

*Recommendation:* use an iterative approach for design, implementation, maintenance and evaluation.

<b>Time</b>	<b>Activity</b>
June 2012 – December 2012	Develop Strategy and approve 1. version of strategy
January 2013 – September 2013	1. iteration – design, implementation, use, maintenance and evaluation
September 2013 - ?	2. iteration – design, implementation, use, maintenance and evaluation

#### 15. Metadata Strategy: IT-solution – applications and databases

*Status:* Colectica tested

Follow up:

- Ask Algenta for possible reduction on prices (Statistics Denmark)
- Ask management and project-manager for Twinnig project on possible finansing. (BHAS)
- Test other tools (Ddieditor ets). Does it give same functionality as Colectica etc. (ISTAT, Statistics Denmark and Business Process and Metadata Group)

#### 16. Metadata Strategy: IT-solution – dataflow

Processes (not all) in the GSBPM use and produce metadata.

*Status:* discussed

*Recommendation:* The coming discussion on GSBPM should include discussion on the use and production of metadata. Metadata input and output from GSBPM processes should be described.

#### 17. Metadata Strategy: Guidelines for production and use of metadata

The guideline should contain descriptions on how to use and produce metadata in the production-process.

*Status:* has been discussed

*Recommendation:* The guidelines should be added to descriptions of business processes. Summary of and reference to the GSBPM-document should be included in the metadata strategy. The topic should be discussed at next mission.

#### 17. Should the strategy for GSPBM and the strategy for Metadata be in the same document?

*Status:* Discussed at mission

*Recommendation:* It is recommended to place strategies in two separate documents. The preparation and implementation of the strategies should be done by the Business Process and Metadata Group.

#### 18. Practical examples / training

All user groups need information on the content and use of a coming metadata-system.

*Status:* has been discussed with users at RSIS.

*Recommendation*

- Training for users (ISTAT and Statistics Denmark)
- Examples from other NSI's (ISTAT and Statistics Denmark)
- The document ["Part A. Statistical Metadata in a Corporate Context"](#) from UNECE should be read and discussed at the Institutions.

#### 18. How to design future missions

*Status:* The contract has three missions on creating a classification server. Since the classification metadata is already present in DDI we should consider changing the content of these missions.

*Recommendation*

- Find out about possible change of the contract. Is it possible? Who should be involved? etc.
- If the contract is changed the work on next mission should prepare a prioritized list of functions to be implemented. This list should be used as input for decisions on the content and form of the rest of the missions.



## 2. Plan for work until next mission

1. Continue work on DDI / Colectica (BiH) and test other software-tools (BiH and ISTAT and Statistics Denmark)
2. Harmonization of code-list (BiH)
3. Translation and review of full GSBPM document (BiH)
4. Overview of metadata as part of current situation (BiH)
5. Prepare meeting with internal and external user (BiH)
6. Request from external users (BiH)
7. Examples on the use of metadata in surveys (functionality, interface etc) (ISTAT and Statistics Denmark)
8. Prices on Colectica (ISTAT and Statistics Denmark)
9. GSBPM strategy (2-4 pages) (ISTAT)

## 3. Topics for the next mission:

1. Follow up on work between missions (see chapter 2)
  - Testing DDI with SBS data
  - Test of other tools (Ddieditor etc)
  - Current situation (metadata etc)
  - Examples from other NSI
2. Discussion and finalise draft on GSBPM – strategy
3. Metadata-strategy
  - Discussion of elements in Strategy
  - Meeting with users (internal / external)
  - Meeting with directors
4. Discussion, training etc
5. Detailed road-map for rest of activities in component 3.2.

Final topic-list will be included in the Terms of Reference.

## 4. Signatures

For the approval of the contents of this summarized report, representatives from BHAS, FIS and RSIS as well as the experts and the RTA signs here:

Date: 29/6-2012

---

Component leader, BHAS

---

Component leader, FIS

---

Component leader, RSIS

---

RTA

---

Expert, Statistics Denmark

---

Expert, ISTAT