TWINNING CONTRACT

JO 21 ENI ST 01 22

Strengthening the capacity of Jordan's Department of Statistics in terms of compilation, analysis and reporting of statistical data in line with International and European best practices

MISSION REPORT

on

Component 1
Roadmap for the development of an integrated administrative data system in Jordan with pilots on Statistical Business registers (SBR) and population statistics

Activity: 1.5.2:
Metadata awareness and Structural metadata

Mission carried out by
Mr. Søren Kristensen
Mr. Kim Rene Duncan-Bendix
Mr. Kim Huuhko

Amman, Jordan
18-21 November 2024

Version: Final

Strengthening the capacity of Jordan's Department of Statistics

Authors' names, addresses, e-mails

Mr. Søren Kristensen
Chief advisor, Quality Unit, Department of Methodology and Analysis
Statistics Denmark
Sejrøgade 11
DK-2100 Copenhagen Ø
Denmark
Email: skr@dst.dk

Kim Rene Duncan-Bendix
System Analyst, Metadata-team, IT Department
Statistics Denmark
Sejrøgade 11
DK-2100 Copenhagen Ø
Denmark
Email: kdb@dst.dk

Mr. Kim Huuhko
Senior Specialist
Statistics Finland
Työpajankatu 13
00580 Helsinki
Finland
Email: kim.huuhko@stat.fi

Table of contents

Executive Summary 5
1. General comments 6
2. Assessment and results 6
3. Conclusions and recommendations 6
Annex 1. Terms of Reference 7
Annex 2: Programme for the mission 12
Annex 3. Persons met 13

List of Abbreviations

- BC – Beneficiary Country
- DDI – The Data Documentation Initiative
- DoS – Department of Statistics
- GSIM – Generic Statistical Information Model
- GSBPM - Generic Statistical Business Process Model
- MS – Member State
- NDC – Interactive National Data Center in Jordan
- PL – Project Leader
- RTA – Resident Twinning Advisor
- STE – Short-term Expert

Executive Summary

The mission provided an in-depth exploration of structural metadata, focusing on its conceptual framework, implementation, and application. Key objectives included understanding the international conceptual model Generic Statistical Information Model (GSIM) developed by UNECE, its connection to the implementation standard Data Documentation Initiative (DDI), and the practical use of metadata in statistical production and dissemination.

Experts introduced the foundational principles of GSIM, highlighting its role in creating a common language for metadata management within and across National Statistical Institutes (NSIs). They emphasized the flexibility of GSIM as a conceptual model, allowing organizations to adapt and focus on core objects relevant to their metadata strategy. This approach was illustrated with examples from Statistics Denmark, where only a subset of GSIM objects is implemented in practice.

The mission also covered the advantages of separating metadata from data, allowing reuse across multiple data collections and dissemination, thereby improving efficiency and consistency. Particular attention was given to the DDI Lifecycle standard. Practical implementations of DDI, such as the Colectica metadata editor used in Statistics Denmark, were discussed alongside similar systems employed by Statistics Finland.

Over subsequent days, the mission delved into the internal and external uses of metadata. Internally, tools like SMOF-BI (a Power BI application integrating metadata for error detection) demonstrated how metadata enhances statistical production by improving data quality checks. Externally, metadata's role in fostering transparency and usability for the public was emphasized, with examples from the National Data Center in Jordan and the Finnish StatFin Databank.

Another critical aspect was the delineation of roles and responsibilities among quality, IT, and statistical staff. Collaborative efforts between these groups, underpinned by management support, were identified as essential for successful metadata initiatives. The mission also stressed the importance of embedding metadata throughout the statistical production process, as exemplified by Statistics Finland's transition to a process-oriented production system based on the UNECE's Generic Statistical Business Process Model (GSBPM).

Finally, the mission addressed specific metadata documentation challenges within the Department of Statistics (DoS). Presentations focused on practical steps for organizing and storing structural metadata, such as classifications, and strategies for reusing metadata effectively. The discussions underscored the importance of clear ownership and comprehensive documentation to maintain high-quality metadata systems.

The mission concluded with actionable insights for developing metadata repositories, leveraging standards like DDI, and fostering cross-functional collaboration to enhance metadata management and utility.

1. General comments

This mission report was prepared within the Twinning Project "*Strengthening the capacity of Jordan's Department of Statistics in terms of compilation, analysis and reporting of statistical data in line with International and European best practices".* This Mission related to the following Mandatory Results (MR) and indicators:

MR 1.5: Implement training programmes and develop training materials both within DoS and with partner institutions on the use of administrative records for statistical purposes, based on pilot projects above.

- **Indicator 1.5.A:** Detailed documentation on statistical standards, classifications, identifiers etc. developed.
- **Indicator 1.5.B:** Comprehensive training programs and workshops provided for DoS staff and partner institutions
- **Indicator 1.5.C:** DoS leadership role in ensuring proper statistical standards applied across the Jordanian statistical system reinforced.

The purpose of this activity is to introduce international standards for structural metadata, raise awareness of role and benefit of complete and comprehensive structural metadata in the production process of official statistics and to demonstrate how structural metadata can support a metadata driven production.

The consultants would like to express their sincere thanks to all officials and individuals met for the kind support and valuable information which they received during the online sessions which highly facilitated their work. The views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, Statistics Denmark or, Statistics Finland.

2. Assessment and results

The mission gave a comprehensive overview of the work with structural metadadata. The experts shared their knowledge and experience in relation to the conceptual framework for metadata (GSIM)[1], the implementation of GSIM via standards such as the DDI[2], and the implementation and usage of metadata when publishing data.

The the external consusltants currently working on developing an Interactive National Data Centre (NDC) in Jordan gave a presentation on how data is to be presented to the public and what metadata would be necessary to allow the public to get a full understanding of data.

Figure 1 provides an overview of the topics that were dealt with during the mission and also describes how the relate to each other. The figure was presented as part of the summary of the mission.

[1] https://unece.org/statistics/modernstats/gsim
[2] https://ddialliance.org/Specification/DDI-Lifecycle/

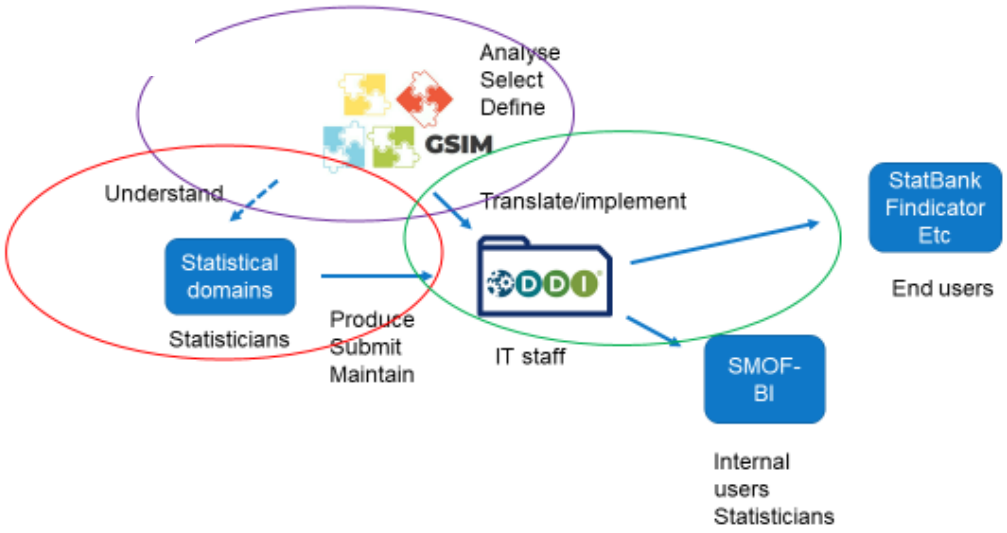

Figure 1: *Interrelations between topics discussed at the mission.*

On the first day focus was on GSIM, the information model that has been developed to encompass all relevant objects in the production of statistics. Mr Kristensen gave an introduction to the purpose and object of an information model in general and also specifically in relation to statistics. GSIM is a very complex model and aspires to provide a complete and comprehensive overview of elements in the statistical production from the business level to individual datapoints. Understanding and familiarizing oneself with the model is very useful when working with metadata. It supports the creation of a common language both internally in a statistical institute and internationally. It also underpins the important notion that the many metadata objects are and should be understood as interconnected. It is important to stress though that the model need not be picked as a whole for the metadata work. GSIM is a conceptual model. When it comes to the implementation of the model in a metadata repository it is advicable to focus on a few key objects. Statistics Denmark's metadata model uses only a fraction of the objects described.

Mr. Kim Duncan-Bendix then gave a presentation of the standards which are used to implement the conceptual GSIM into a metadata structure or repository. DDI is the standard used in Statistics Denmark. Specifically we use Colectica which is a DDI editor. Other standards are available as are other editors. Further information on DDI can be found at the homepage of the DDI alliance. Mr Duncan-Bendix described how the work with metadata has undergone an evolution from being attached to data to being separated. This has the great advantage that metadata can be reused for other data – for instance in newer cycles or other data collections. This has the added advantage that certain metadata will only have to be entered once. Apart from being more efficient it also ensures a higher degree of consistency. The most recent DDI standard – DDI lifecycle – accommodates recurring data collection and is therefore more appriotate for statistical production than its predecessor – DDI codebook.

The first day thus covered the conceptual model, the implementation standard as well as the implementation of the metadatamodel as illustrated by the blue arrow between GSIM and DDI in figure 1.

On the second day the focus turned to the use of metadata internally and externally. This was also the focus of the presentation by the Jordanian Interactive National Data Centre team on the previous day. Mr. Kim Huuhko gave a presentation on how structural metadata is used by Statistics Finland in the Statfin databank and also in data views created with PxGraf and made available to the users on the Statistics Finlands homepage. Statistics Finland has also a strategic goal to harmonise all of its information processes and base all information and service production on common data and metadata resources. These systems converse with each other through API's using structural metadata as a common vocabulary.

Following that, Mr. Kim Duncan-Bendix reported on how in Statistics Denmark metadata is used by internal user in the error detection and editing system known as SMOF-Bi. Its a power bi application which combines data with metadata so that statistical producers more easily can detect outliers and errors in data. The metadata that feeds into the power bi system is sourced from the metadata bank which is the metadata repository in Statistics Denmark. Structural metadata, such as codelists and classifications, are stored in the metadata bank and reused in SMOF bi. This gives the aforementioned advantages of efficiency, transparancy, and consistency. SMOF bi is a very useful system for the statistical producers and it has given them an extra incentive to use and update the metadatabank.

In this way the second day, was dedicated to the use of metadata in communication and production. In figure 1 this is represented by the two arrows going from the DDI – the metadata repository – to SMOF_BI and the Statbank.

On the third day the focus was on staff and roles on request of the DoS. Mr. Søren Kristensen took departure in Figure 1 to delineate the tasks and domains assigned to different groups of staff. The three circles represent different groups of staff. The purple circle is the quality division, the green IT staff, and the red statistical producers. Each have their own area of responsibility and the tasks that go with it.

The quality division is responsible for the conceptual dimension of the metadata work and for identifying how GSIM – or similar models – should be defined in the particalur context of the NSI. It is also typically the quality division which takes the initiative in regard to developments within metadata. Quality staff needs to be familiar with GSIM and also have an understanding of the implementation standards (such as DDI), as well as being familiar with the statistical production process. Not every employee in the quality division needs to be an expert on GSIM, DDI and statistics as long as the quality division collectively possesses the competences.

IT staff is responsible for the development and maintance of the IT infrastructure supporting the storage of metadata (the repository) and the access to metadata from various platforms. Both in Statistics Denmark and Statistics Finland APIs are used to establish the connection to the metadata repository from user platforms. Staff working in IT will need a thorough understanding of GSIM and also som understanding of statistical production,

Statisticians or producers of statistics are the experts on the statistical domains. They are responsible for producing, submitting and maintaining metadata. They need some understanding of GSIM but only to the level where they are familiar with the terminology and the relations between objects.

It is important to stress that in order to create good metadata solutions cooperation is necessary beteween these three groups of staff. They need to understand each others domains to a certain extent and to offer guidance to the other groups.

In order for this work to be successfull it is also important that it is backed and prioritised by management. NSIs are typically concerned with data and data quality and unless metadata is made a specific priority there is a risk it will not be given adequate attention.

Following this Mr. Kim Huuhko gave a presentation focussing on data and metadata repositories. Statistics Finland is undergoing a transformation from a silo based production system to a functional system based on GSBPM. At the same time Statistics Finland has taken the opportunity to embed harmonized and standardized metadata in all the processes from data collection to the dissemination of data. Previously there has been a tendency that metadata descriptions has been created and used independently by different statistical experts. In many cases it has also led to a situations where this metadata has been out of sync from the actual data its suppose to describe.

Mr. Kim Huuhko presented some of the core principles of the metadata system of Statistics Finland. They are very similar to the principles of the metadata system in Statistics Denmark. Amongst them are the principles that information is stored only once, that concepts are defined and that ownership and responsibilities are documented. Mr. Huuhko went on to present the tools that Statistics Finland use for metadata and documentation. One is the Metsy metadata system which is roughly built on the same principles as the metadata bank in Statistics Denmark. The quality assessments are stored in the documentation system Lara, whereas in Denmark they are stored in the metadata bank.

Statistics Finland make extensive use of the Px tools, that is for instance the PxWeb and the PxGraf. A recent important development is the PxWebApi 2.0, which has been developed in a cooperation between Statistics Sweden and Statistics Norway. Statistics Finland considers using this and other APIs as a primary tool of publishing in the future.

Finally Mr. Kim Huuhko argued that the important but very difficult goal of harmonizing metadata among all different statistical domains can be achieved through many different sets of IT solutions or organizational structures. What is important though, is that the critical contents are validated in every possible phase by strict interconnected systems as well as skillful content owners who work as the final gatekeepers of quality.

On the fourth day, focus was on how to work specifically with structural metadata from DoS. Here the attention was the role of the statisticians and the question of how to best document and store metadata. In figure 1 this is illustrated by the blue arrow going from the statistical domains to the repository as represented by DDI.

Mr. Kim Duncan-Bendix gave a presentation on structural metadata and how to work with it in the context of DDI lifecycle. The focus of the representation was how structural metadata should be stored in one location and then reused. Ms Nielsen, the RTA, gave a presentation on

the work with structural metadata in DoS taking departure in some concrete classifications relating to geography and marital status. The presentation focused on how what fields are necessary and how to organise codes and classifications.

3. Conclusions and recommendations

Based on the input and the discussions following the presentations the experts presented the following recommendations on the final day of the mission. The recommendations are grouped into three categories; conceptual, organisational and process oriented recommendations.

Conceptual:

- DoS should familiarize itself with GSIM, but should also only select a few items to work with initially. This could for instance be code lists and variables. DoS should at the same time start considering which items could be included in later phases. Overall it is not recommended that the GSIM in its entirety is implemented. It is a conceptual model, which main purpose is to create on overview of information objects related to the production of statistics.
- DoS should familiarize itself with DDI or similar architecture. This is necessary in order to implement the objects of GSIM in the production of statistics.

Organizational:

- Ensure that management makes it a priority to work on metadata. If metadata work is not given sufficient priority and backing, it is likely that it will not take off.
- Clarify roles and responsibilities (IT, quality division, statistical domains, others). Metadata work demands cooperation between different groups of staff and it is important that they know their task and roles.
- Engage the statistical domains in the work on classifications and code lists.

Process (think big start small):

- Start with official code lists and classifications as phase 1
- Get inspiration from SDMX code lists
- Start with a few classifications. Choose important ones, e.g. ISIC, Geography, Education
- Build metadata repository to store and versioned code lists and classifications (initially)
  - Harmonize
  - Update in one place and one place only
- Consider how to make metadata available to end users via PxWeb or similar platforms.
- Use API to access metadata.

Annex 1. Terms of Reference

Terms of Reference

EU Twinning Project JO 21 ENI ST 01 22

Component 1:
Roadmap for the development of an integrated administrative data system in Jordan with pilots on Statistical Business registers (SBR) and population statistics

Activity 1.5.2:
Data and metadata structure II – Structural metadata

Dates: 18 - 21 November 2024

Content

List of abbreviations 12
1. Objective and Mandatory Results for the component 13
1.1 Objective of the component from the Fiche 13
1.2 Mandatory results and indicators for achievement for each sub-component 13
2. Purpose of the activity 14
3. Expected output of the activity 14
4. Participants 14
4.1 MS Short Term Experts (STE's) 14
4.3 Participants for opening and closing sessions: 16
4.4 The Twinning Team 16
5. Current status 17

List of abbreviations

BC Beneficiary Country
DDI The Data Documentation Initiative
DoS Department of Statistics
ESS European Statistical System
FAIR Findable, Accessible, Interoperable and Reusable.
GSIM Generic Statistical Information Model
GSBPM Generic Statistical Business Process Model
MS Member State
RTA Resident Twinning Advisor
SDMX Statistical Data and Metadata eXchange
STE Short Term Expert
ToR Term of References

1. Objective and Mandatory Results for the component

1.1 Objective of the component from the Fiche

To prepare a roadmap for the development of an integrated administrative data system for Jordan, and conduct pilot projects on creating an SBR and strengthening population statistics.

1.2 Mandatory results and indicators for achievement for each sub-component

Component 1 is sub-divided in five sub-components each with a Mandatory Results (MR) and two to four indicators of achievements associated with the sub-component (Table 1)

Table 1: *Mandatory results and indicators for achievement (IA) for each sub-components within Component 1: an integrated administrative data system for Jordan*

MR from the Twinning Fiche	**Indicator**
MR 1.1: Compile an inventory of administrative data on business and households and an indicative roadmap for inclusion in an integrated system	• **Indicator 1.1.A:** Inventory of administrative data variables and detailed supporting metadata prepared; • **Indicator 1.1.B:** Tentative roadmap prepared for inclusion of data in integrated system;
MR 1.2: Pilot project to develop strategy for integrating administrative data sources for the purposes of creating an SBR	• **Indicator 1.2.A:** Administrative data sources identified and assessed and plan developed for integrating these with Census of Establishments (CoE) information in an SBR; • **Indicator 1.2.B:** Documentation prepared on database structures and compliance with statistical standards, classifications (e.g. ISIC, Rev 4) etc. and use of common identifiers etc.; • **Indicator 1.2.C:** Explore how SBS can benefit other statistical domains in the DoS;
MR 1.3: Undertake pilot project on how administrative records can be used to strengthen population statistics and inform framing of the 2025 CoP questionnaire	• **Indicator 1.3.A:** Inventory of data sources prepared and assessed and action plan for incorporation in DoS statistics developed; • **Indicator 1.3.B:** Methodology developed for incorporating administrative data • **Indicator 1.3.C:** Documentation prepared on statistical standards, classifications, identifiers, mapping etc.; • **Indicator 1.3.D:** Review of how administrative data can assist in developing the COP 2025 questionnaires
MR 1.4: Develop strategy for ensuring flows of data between the DoS and counterpart institutions are established on an ongoing basis for pilot projects above	• **Indicator 1.4.A:** Review of technical infrastructure for data transfers and action plan prepared based on 1.1 and 1.2 above; • **Indicator 1.4.B:** MoUs agreed between DoS and partner institutions; • **Indicator 1.4.C:** Agreement on statistical standards, classifications, identifiers etc. between DoS and partner institutions; • **Indicator 1.4.D:** Review of data flows within the DoS;
MR 1.5: Implement training programs and develop training materials both within DoS and with partner institutions on the use of administrative records for statistical purposes, based on pilot projects above	• **Indicator 1.5.A:** Detailed documentation on statistical standards, classifications, identifiers etc. developed; • **Indicator 1.5.B:** Comprehensive training programs and workshops provided for DoS staff and partner institutions; • **Indicator 1.5.C:** DoS leadership role in ensuring proper statistical standards applied across the Jordanian statistical system reinforced;

MR 1.6: A governance roadmap for decisions makers data access outlined	• **Indicator 1.6.A:** Best international practices for NDC's outlined • **Indicator 1.6.B:** Stakeholder awareness raised and needs from stakeholder mapped; • **Indicator 1.6.C:** Organizational structure and required skills for staffing the National Data Center outlined; • **Indicator 1.6.D:** Requirements and standards for data and metadata layer outlined;

2. Purpose of the activity

The purpose of this activity is to introduce international standards for structural metadata, raise awareness of role and benefit of complete and comprehensive structural metadata in the production process of official statistics and to demonstrate how structural metadata can support a metadata driven production.

During the Mission the following topics will be introduced, demonstrated and discussed

- Role of metadata in the modernization process of official statistics;
- International standards for structural metadata introduced:
  - The conceptual model GSIM (Generic Statistical Information Model);
  - The implementation standard DDI (The Data Documentation Initiative);
- Examples of metadata driven production in MS
- Metadata repository in MS
  - IT infrastructure
  - Establishment and maintenance of metadata
  - Organizational structure to coherent and updated metadata
- Metadata management and harmonization among different statistical domains
- Data management (database based)
- Reporting and data exchange
- Dissemination via GUI and APIs

3. Expected output of the activity

- Activity report;
- Metadata awareness raised;
- Common understanding of metadata driven production and international standards;
- Process and requirement for creating and maintaining a repository of metadata outlined both in terms of IT requirement and organizational requirements;

4. Participants

4.1 MS Short Term Experts (STE's)

- **Mr. Søren Kristensen**, Quality Unit, Department of Methodology and Analysis, Statistics Denmark. Mr. Søren Kristensen has a solid knowledge on structural metadata both at a conceptual level and practical level. Mr. Søren Kristensen has more than 20 years of experiences in statistical production and dissemination.

- **Mr. Kim Rene Duncan-Bendix**, System Analyst, Metadata-team, IT Department, Statistics Denmark. Mr. Kim Rene Duncan-Bendix has been working with metadata since 2009 and has a substantial theoretical and practical experiences in metadata based on solid international cooperation. Mr. Kim Rene Duncan-Bendix has been one of the key persons in the practical implementation of the metadata model in Statistics Denmark and has been the main developer for system supporting metadata-driven production e.g. for supporting the visualization of collected data, validation and publication.

- **Mr. Kim Huuhko**, Senior Specialist, Statistics Finland. Mr. Kim Huuhko is specialized in dissemination databases and database-related web services as well as in statistical graphs, presentations, and statistical literacy in general. In addition, Kim Huuhko is also an expert in issues regarding semantic interoperability, linked open (statistical) data, and the automatization of statistical processes.

COMPONENT LEADER

Data Management Directorate

- Mr. Jaffar Ababneh, Director of Data Management Directorate
  Jafaar.Ababneh@DOS.GOV.JO
- Mr. Mohammad Al-Omari, Administrative Data Division, Data Management Directorate
  Mohammad.Omari@DOS.GOV.JO
- Ms. Safa Abu Aiteh, SBR Division, Data Management Directorate
  safaa.aboeta@dos.gov.jo
- Ms. Lina Al-Jazzazy, Administrative Data Division, Data Management Directorate
  Lina.AlJazzazy@DOS.GOV.JO

Directorate of Electronic Transformation and Information Technology

- Ms. Manal Al-Khofash, Directorate of Electronic Transformation and Information Technology
  Manal.Khuffash@DOS.GOV.JO
- Mr. Abdullah Al-Sous, Directorate of Electronic Transformation and Information Technology
  Abdullah.AlSous@DOS.GOV.JO
- Ms. Tamara Al-Wadi, Directorate of Electronic Transformation and Information Technology
  Tamara.Wadi@DOS.GOV.JO

Directorate of Methodologies and Studies

- Mr. Thafer Al-Shawawrah, Quality Division
  dhafer.alshawawreh@dos.gov.jo
- Ms. Roqayah Al-Sanabrah, Quality Division
  roqayah.alsanabra@dos.gov.jo

National Data Center team

- Mr. Nabil Abu Sel, Director of the Interactive National Data Center
  Nabil.AbuSall@DOS.GOV.JO
- Mr. Mohammad Al-Hiari, Interactive National Data Center
  Mohammad.Alhiary@DOS.GOV.JO

- Mr. Mohammad Al-Janada, Interactive National Data Center
  Mohammad.Al-Janada@DOS.GOV.JO

4.3 Participants for opening and closing sessions:

- HE General Director Dr. Haidar Fraihat
  dg.of.dos@DOS.GOV.JO
- Dr. Ali Al-Shibli, Operational Assistant to the Director General
  Ali.Shebli@DOS.GOV.JO
- Dr. Tayseer Muqdadi, Technical Assistant to the Director General
  Tayseer.Megdady@DOS.GOV.JO
- Dr. Fozan Alhurout, Administrative and Financial Affairs Assistant to the Director General
  Fozan.Hrout@DOS.GOV.JO
- Ms. Ahlam Al Rosan, Director of Electronic of Transformation and Information Technology Ahlam.AlRosan@DOS.GOV.JO

4.4 The Twinning Team

- Eng. Mr. Tamer AlRosan, Head of plant Statistics Division Jordan (RTA Counterpart),
  Tamer.AlRosan@DOS.GOV.JO
- Dr. Charlotte Nielsen (RTA), cln@dst.dk
- Ms. Zaina Amireh (Language Assistant), zainaamireh3@gmail.com
- Ms. Thekra Altorah (RTA Assistant), thekra.twinning.rtaa@gmail.com

5. Overall agenda

- **Day 1:**
  - BC: Current status for Structural metadata in DoS
  - International standards for structural metadata:
    - The conceptual model GSIM (Generic Statistical Information Model) – Søren Kristensen;
    - International implementation standards (DDI) – Kim Duncan
- **Day 2:**
  - MS: Role of metadata in the modernization process of official statistics;
    - In Finland - Kim Huuhko
    - In Denmark
  - MS: Examples of metadata driven production in MS
    - In Finland - Kim Huuhko
    - In Denmark (SMOF-BI and SMOF-DAF) – Kim Duncan

- **Day 3:**
  - MS: Metadata repository in MS
    - In Finland - Kim Huuhko
    - In Denmark – Søren Kristensen/Kim Duncan
  - MS: IT infrastructure
    - In Finland - Kim Huuhko
    - In Denmark –Kim Duncan
    -
  - Establishment and maintenance of metadata
    - In Finland - Kim Huuhko
    - In Denmark – Søren Kristensen/Kim Duncan
  - Metadata management and harmonization among different statistical domains
    - In Finland - Kim Huuhko
    - In Denmark – Søren Kristensen/Kim Duncan

  - BC and MS: Discussion and outline for implementation of structural metadata in DoS

- **Day 4:**
  - BC and MS: Follow up and conclusion

6. Background information

Need for disaggregated statistics and diverse data services:

Official statistical producers operate in a rapidly changing landscape, where the pace of change accelerates annually. Increasingly complex societal issues demand more timely, disaggregated statistics and diverse data services. Adoption of new data sources like administrative and big data poses challenges in analysis methods, data access, ethics, and privacy. Amidst competition from other data providers, statistical organizations must enhance product communication and brand advocacy for trustworthiness. To tackle these challenges, statistical organizations must invest in modernization, staff capabilities, and technology. Despite limited resources, efficiency improvements are crucial to ensure adaptability and resilience in the dynamic data ecosystem.

Jordan Economic Modernization Vision 2030:

Recently the Jordan Economic Modernization Vision 2030 was launched and "Smart Jordan" was identified as one of the eight Growth Drivers to implement the Economic Modernization Vision. The 'Smart Jordan Driver' includes seven sectors where data is one of them. This indicates the national interest to ensure constant and reliable data sources, and robust statistical systems that contribute to timely and informed policy making. One of the measures that will be taken is to establish an Interactive National Statistical Center (NSC) that will provide data to all users groups according to their need.

Interactive National Statistical Center:

The NDC will be built around the following four pillars

I. **A Data Management Center (DMC)** that will support all internal operations and production of statistics in accordance with best international practices with a high level of security to protect data

II. **Governance tools** e.g. such management, organizational structure, security and confidentiality policies etc.

III. **Data dissemination ecosystem platform (DDC)**

IV. **A platform for uploading external data** to the DDC e.g. administrative data owner in Jordan, NSI's from other countries as well as International organizations

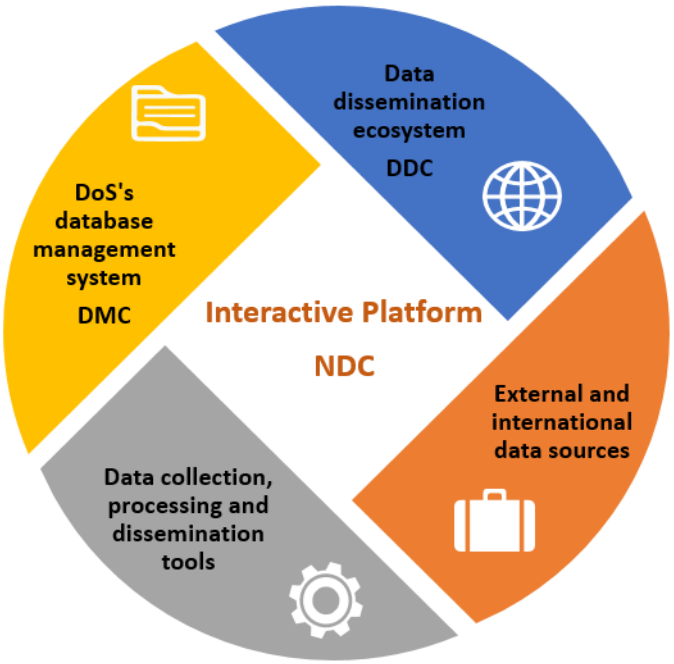

Figure 1: *Illustration of four pillars of the the future National Data Center (NDC).*

Metadata:
Standardized structural metadata is a central element ensuring an integrated approach to facilitating statistical data and metadata exchange, production and dissemination of data and statistics.

But what is metadata? Metadata is, literally, "data about data". There are two types of statistical metadata: Structural metadata and Reference metadata.

- **Structural metadata** is used to identify statistical data, for example titles, variable names and units of measurement (eg Jordanian dinars).

- **Reference metadata** describes the statistical concepts and methods used to collect and generate data and provides information about data and statistical quality. They thus help users with the interpretation of the outputs provided by the National Statistical Institute (NSI). Unlike structural metadata, reference metadata can be decoupled from the physical data.

Metadata is thus absolutely essential when NSI's disseminates statistics or makes data available to users. Basically, you can say that data has no value without metadata. This is illustrated in Figure 2, which illustrates statistical data without metadata, data with structural metadata, and data with both types of metadata. Figure 3 illustrated the principle of classifications that can support interactive data presentation where user can drill down and extract data or visualize data at a more and more granular level.

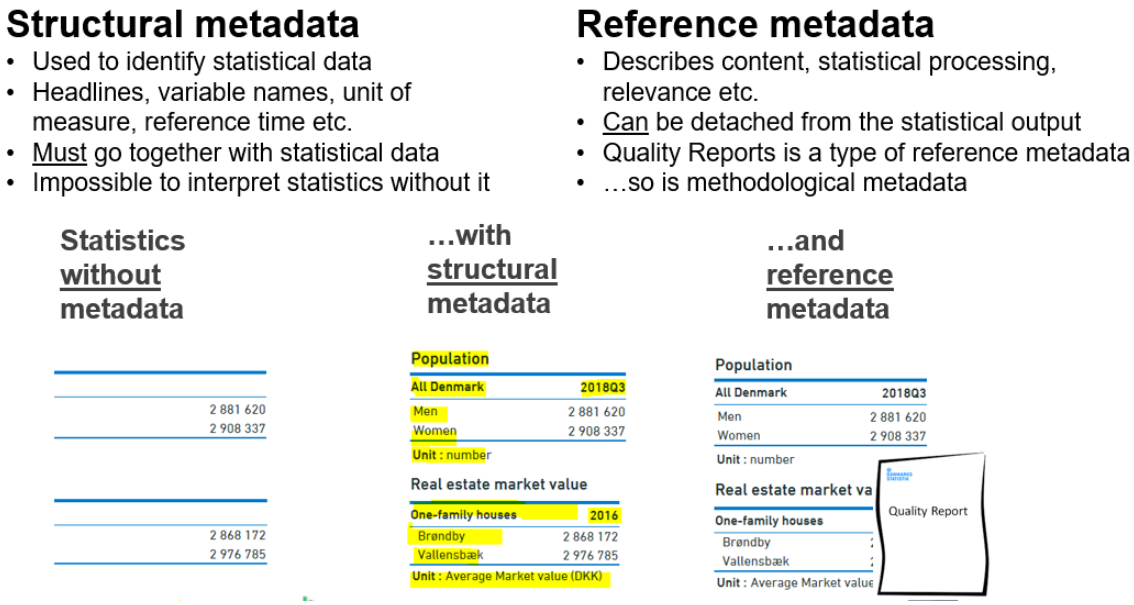

Figure 2: *Illustration of the value of metadata in the dissemination. The tables are meaningless unless they are provided with structural metadata.*

Figure 3: *Illustration of a hierarchical geographical classification that can e.g. support interactive data presentation where user can drill down and extract data or visualize data at a more and more granular level.*

International standards for metadata:

- The Generic Statistical Information Model (GSIM) (version 2.0) defines a number of information objects that are relevant for the production of statistics. The information objects can thus be used to describe input and output in the design, development and production of statistics.

  GSIM contains more than 130 information objects organized into five main groups according to their purpose. The objects include, for example, variables, code lists and classifications, i.e. objects that are either input to or output from the individual process steps in the Generic Statistical Business Process Model (GSBPM). Examples of process steps where metadata is absolutely essential are, for example, *"Review and validate data", "Aggregate data"* and" *Produce dissemination products"*.

  GSIM provides a common international language to describe the production of statistics. GSIM has been developed under the auspices of the High-Level Group for the Modernization of Official Statistics, which belongs to The United Nations Economic Commission for Europe (UNECE).

- Among a number of technical standards for how GSIM can be implemented technically, there are the Statistical Data and Metadata eXchange (SDMX) and the Data Documentation Initiative (DDI). They complement GSIM, which does not prescribe how the objects are to be implemented technically.

  DDI was primarily developed in the research world, while SDMX was developed in collaboration between the EU, OECD, the UN and other international organisations, and aims to standardize and modernize the processes for the exchange of statistical data and metadata between international organizations and their member states.

  DDI supports the FAIR principles, which seek to ensure that data and metadata are Findable, Accessible, Interoperable and Reusable.

The technical implementation standards are compatible with GSIM right down to the object level, and DDI refers in their guides to GSIM for each individual information object.

- For the classifications, the Neuchâtel Terminology Model is also used to define the objects and the relationships between them. The model has been developed by the Neuchatel Group, which is an informal collaboration between the statistical offices in Denmark, Norway, Sweden and Switzerland.

- The Single Integrated Metadata Structure (SIMS) is used as a starting point for the statistical documentation in all EU Member states and has also been adobted in Jordan. SIMS consists of a number of fields that primarily describe various properties of the statistics in free text.

Status for structural metadata in DoS:

Currently, there is no central repository or central managements of structural metadata in DoS. Each statistical Unit maintain their own structural meta

Annex 2: Programme for the mission

Day 1 (9:30 – 15:00):

- BC: Current status for Structural metadata in DoS
- MS: International standards for structural metadata:
  - The conceptual model GSIM - Søren from DK
  - International implementation standards (DDI) - Kim from DK

Day 2 (9:30 – 15:00): (Each MS have a presentation):

- MS: Role of metadata in the modernization process of official statistics;
- MS: Examples of metadata driven production in MS (SMOF-BI)

Day 3 (9:30 – 15:00): (Each MS have a presentation and joint work):

- MS: Metadata repository in MS
  - IT infrastructure
  - Establishment and maintenance of metadata
  - Metadata management and harmonization among different statistical domains
  - Organizational structure
  - BC and MS: Discussion and outline for implementation of structural metadata in DoS

Day 4 (9:30 – 15:00): (Joint work):

- BC and MS: Follow up and conclusion
- DDI examples of breaking down dataset, and looked at DoS data and classifications.

Abbreviations:
***MS** = EU Member State (Denmark, Germany, Italy, Lithuania, Finland);*
***DoS** = Department of Statistics, Jordan*

Annex 3. Persons met

Data Management Directorate

- Mr. Mohammad Al-Omari, Administrative Data Division, Data Management Directorate
  Mohammad.Omari@DOS.GOV.JO
- Ms. Safa Abu Aiteh, SBR Division, Data Management Directorate
  safaa.aboeta@dos.gov.jo
- Ms. Lina Al-Jazzazy, Administrative Data Division, Data Management Directorate
  Lina.AlJazzazy@DOS.GOV.JO

Directorate of Electronic Transformation and Information Technology

- Ms. Manal Al-Khofash, Directorate of Electronic Transformation and Information Technology
  Manal.Khuffash@DOS.GOV.JO
- Mr. Abdullah Al-Sous, Directorate of Electronic Transformation and Information Technology
  Abdullah.AlSous@DOS.GOV.JO
- Ms. Tamara Al-Wadi, Directorate of Electronic Transformation and Information Technology
  Tamara.Wadi@DOS.GOV.JO

Directorate of Methodologies and Studies

- Mr. Thafer Al-Shawawrah, Quality Division
  dhafer.alshawawreh@dos.gov.jo
- Ms. Roqayah Al-Sanabrah, Quality Division
  roqayah.alsanabra@dos.gov.jo

National Data Center team

- Mr. Nabil Abu Sel, Director of the Interactive National Data Center
  Nabil.AbuSall@DOS.GOV.JO
- Mr. Mohammad Al-Hiari, Interactive National Data Center
  Mohammad.Alhiary@DOS.GOV.JO
- Mr. Mohammad Al-Janada, Interactive National Data Center
  Mohammad.Al-Janada@DOS.GOV.JO