DOI: 10.25140/2411-5215-2022-1(29)-39-48
UDC 303.064
JEL Classification: D83

Tetyana Lumpova
PhD in Economics, teacher
Professional Pre-Higher Education Institution "Optical-Mechanical College"
of Taras Shevchenko National University of Kyiv (Kyiv, Ukraine)
E-mail: taivlu@meta.ua. **ORCID:** https://orcid.org/0000-0002-7576-4289
Scopus Author ID: 0000-0002-7576-4289

USING THE EXPERIENCE OF INTERNATIONAL STATISTICS TO CREATE A REPOSITORY OF APPLIED METADATA FOR INFORMATION SYSTEMS IN ECONOMIC AREA

Introduction of process approach in organizations engaged in processing statistical and economic information for analytical research, transition in information systems (IS) to service-oriented architecture, integration of data from different sources, use of data warehouses necessitated the development of new approaches to creating and modernizing IS. At the same time, the requirements for the meta-information component of the IS in terms of meeting the requirements of internal and external users regarding the possibility of self-assessment of data quality are increasing. All these challenges require strengthening the meta-information component of the IS for the use of metadata by different user groups. The article considers the key points that need to be taken into account when adapting to the existing IS existing international experience in the modernization of economic IS direction. This information will be useful when creating a metadata repository, data warehouse, information portal on statistical and economic research.

Key words: *data quality; data warehouse; Economic Indicators Classifier; meta-information; metadata repository.*
Bibl.: 9.

Problem setting. Modern changes in information technology (IT) have led to the need to develop new approaches to the creation and modernization of information systems (IS) economic direction, in particular:

– process approach implementation to production in organizations engaged in the processing of statistical and economic information for analytical research,

– transition in IS to service-oriented architecture (COA),

– data integration from different sources, which include indicators derived from "big data",

– use of data warehouses (DW).

At the same time, the requirements for the meta-information component of the IP in terms of meeting the requirements of internal and external users regarding the possibility of self-assessment of data quality are increasing. This concerns especially of relevance, completeness, integrity and consistency. Use external data from open data sources, such as social media (Twitter, Facebook, etc.) and search networks (Internet search, web browsing) as an auxiliary source of information for in-depth analytical research in combination with traditional sources information (statistical and administrative data, data of own production systems) increases the importance of these requirements. Participation in international initiatives for the exchange of economic information, methodologies and technologies for its processing encourages the improvement of corporate standards and, if necessary, their adaptation to existing international ones. All these challenges, in turn, require strengthening the meta-information component of the IS, in particular, the creation or modernization (if any) of the metadata repository (MDR). MDR should be a powerful tool for maintaining and using metadata by different user groups, as well as to provide automated centralized management of the production process. To solve this problem, it is necessary to clearly define the requirements for the functionality of MDR in terms of supporting meta-information of IS in the new environment of rapid changes in the IT field and the constant growth of information flow.

Analysis of the recent researches and publications. Regarding the adaptation of IS to the new situation in the field of IT, it is advisable to refer to international experience, in particular, summarized in recent years by the United Nations Economic Commission for Europe (EEC) at its plenary sessions, meetings, conferences and workshops. This is the best experi-

ence of countries in modernizing their IS in the field of statistics and economic research. Analysis and use of this experience will improve the existing information model of the organization, adapting the best practical solutions to their own IS. First of all, this applies to the use of standard models proposed by the EEC: Common Statistical Production Architecture (CSPA), focused on SOA, typical model of statistical organization (Generic Activity Model for Statistical Organization - GAMSO), standard models of statistical description production processes (The Generic Statistical Business Process Model - GSBPM) and statistical information (Generic Statistical Information Model - GSIM). The EEC website at https://unece.org/statistics/standards-and-metadata provides detailed descriptions of these models, the concept of using applied (statistical) metadata systems in applying the process approach to the organization of production, outlines the structure of the metainformation system and the basic principles of its support, place and role of metadata in the IS of statistical organization, formulated organizational and physical context of their use, examples of practical use of metadata for various statistical surveys in national statistical organizations.

The topic of creating and maintaining MDR in IS is quite scientifically developed. More than three decades ago, Bo Sundgren in [1] formulated the conceptual framework for the creation of MDR as a metadata database as an effective basis for the development of the entire IS. These ideas have not lost their relevance at this time and have been supplemented by the important theoretical material accumulated over the years on the maintenance of metainformation systems, which has been largely implemented in statistical organizations.

In terms of practical use of metadata, the experience of the Swedish Statistical Office in creating a production environment with a high level of standardization and solving the problem of implementing SOA for the national version of GSBPM, which provides a basis for standardizing organizational processes and activities, uses standard methods and tools. scheme for the production of statistical products [2]. For the implementation / modernization of MDR it is advisable to consider the proposed M. Scanu and S. Casagrande in [3] implementation of metadata support system, linked from the point of view of joint use of GSBPM and GSIM, where input information of the production process (PP) performances include: 1) microdata characterizing the units of observation; 2) tables of macrodata as aggregate economic indicators; 3) hypercubes as measurable / spatial data structures. Regarding the creation of RMD for the microdatabase, important practical aspects are considered in [4], which describes the project, which is based on three fundamental pillars: development of structural metadata to improve data management, development of reliable methodologies for data processing at different stages of the statistical process and development of computer tools to ensure easy and efficient access to all information. In [5] M. Erickson considers applied aspects of the implementation of search functions for statistical data, the principles of increasing detail of metadata as a means of providing a basis for assessing the quality of data obtained, as well as providing metadata that allows research and harmonization of data. In [6; 7] the issues of economic data exchange, their integration and reuse, the main issues that need to be considered when creating / modernizing MDR, in particular to ensure the assessment of data quality are disclosed.

Defining of not solved before parts of general problem. It should be noted that the definition of prerequisites for the formation of MDR and approaches to the development of its architecture in terms of centralized retrospective preservation of applied metadata for the subject area for use in the production of statistical and economic data generated from various information sources, serving the needs of external and internal users did not have sufficient coverage in domestic scientific sources.

Target setting. The aim of the study is to determine the basic principles of the generalized information model of MDR to solve the complex problem of supporting the automated process of processing consolidated economic information using SOA, applying the process scheme of production and to ensure multifunctional use of economic data storage.

Statement of the main material. The subject of the article is the MDR for economic IS, which is used for various aspects / types of production activities of the organization: conducting its own statistical observations, surveys, marketing research; analysis of information from external sources (eg official statistics, administrative data, partner data, results of "big data" processing); calculation of statistical and analytical indicators; formation of analytical reports and other documents.

Such IS can be focused on the operating environment (in particular, OLTP-systems, technically configured and optimized to perform the operation of entering and modifying large amounts of information in the database), and / or DW, including analytical systems (in particular, OLAP-systems, technically configured and optimized for data retrieval and retrieval operations in DW). For IS that contain both types of systems (OLTP and OLAP), it is advisable to create a single MDR as a separate functional unit that accumulates all applied metadata and provides tools for working with them.

The main purpose of the MDR is the need to obtain a tool to support the database, which accumulates information from different sources, as well as support for metadata-driven IS and provide different groups of users with requested information on economic data, according to user access. Metadata-driven IS can be defined as a system where information on processing constraints (eg period) is provided to perform actions through special external metadata descriptions, and the infrastructure of such IS should provide the user with clear and convenient access to information for analysis. It needs to be scalable and extensible in terms of data volumes, integration needs and naturally growing analytical challenges, based on resource management, access and privacy policies.

An important task for the metadata of such IS is to ensure the quality of data, which according to [8] is to address the following key issues:

1) meeting the needs of internal and external users to use the results of production activities - existing products in the IS (analytical tables, reports, etc.);

2) ensuring product quality in accordance with descriptions and specified quality requirements, which include end-to-end harmonization of methods and concepts for all products;

3) introduction of systematic quality control of products, processes, solutions according to formal descriptions and provision of analytical material to improve their quality indicators;

4) providing support in MDR easily accessible metadata for all information objects (IO) and processes. The use of such metadata will allow staff to increase the internal efficiency of the organization.

Let's consider how these issues are solved when using MDR and, accordingly, what requirements must be met by MDR.

The task of creating a single MDR for the accumulation of information from different sources in a single data warehouse requires coordination and centralized monitoring of all actions with metadata (production, transformation and distribution), as well as unification and standardization of application metadata in the organization using conceptual and technical standards. Conceptual standards define the format of description and harmonization rules, technical - structure (list of attributes, rules for matching applied metadata of different sources) and storage format (length of code and text, code format, etc.). Unification of applied metadata should be performed during RMD formation. This implies that for all sources of these metadata, redundant information should be removed, comparability between applied metadata from different sources, as well as between different periods of their receipt, must be ensured. Defining procedures for accessing, searching and obtaining applied metadata from MDR should be the result of analysis of metadata needs of different user groups: economists, analysts, statisticians, IT professionals, external users, etc.

Carrying out unification and standardization will allow to form a meaningful information model that will adequately reflect the functionality of processes and the hierarchy of data and metadata of all IO. In this regard, it is advisable to consider the experience presented in [2] for the creation of RMD for centralized management of metadata in the IS, based on a model where information is stored in a coordinated and structured DW of all products. All components of such IS are integrated and cooperated in an efficient, well-documented and standardized way, which clearly defines the functionality of each component. The link between activities and data is maintained by a common communication platform, where standardized tools and services necessary for the implementation of activities are available through a central MDR, which contains the information needed to describe data in DW, monitors and launches processes which are constantly analyzed and evaluated using process metrics created and used to improve them.

In the process scheme of production organization, the basis for automation of metadata-controlled IP should provide a formalized description of this scheme in the form of a technological plan (TP), which determines the sequence, iterativeness and timing of certain production actions, which are described as PP. The allocation of PP groups in the TP, which are performed as certain production services (for example, the formation of a sample), provides an opportunity to create / adapt the main software applications as services according to the requirements of the SOA. The TP defines the points of data quality analysis and PP performance on the possibility of continuing data processing, which records quality assessments for internal users (for example, the timeliness of a particular aircraft). Unification of the TP description for different activities of the organization will require the creation of a reference book / classifier of PP, which determines the resources and results for each aircraft, which in turn will require the creation of a relevant reference IO. Here, strong methodological support can be obtained from the principles underlying the formation of GSIM [9].

Considering MDR as a repository of meta-information and a component of IS that serves both users and owners of information, providing a deeper understanding of sources of information, the nature of economic indicators and the process of their formation, usually determine the following main functions:

1. Ensuring the performance of two basic metadata functions: the first is a unique and formal definition of the content and relationships between objects and processes in the IS; second - determination of all relevant technical parameters for software applications. In terms of the use of MDR for the entire IS, this means ensuring process management and data management, which, in turn, accordingly defines the two blocks of IO applied meta-information. The first block is a description of the technological process in the form of TP, the core of which is the classifier of PP, combined with reference books / lists of descriptions of resources for PP and the results of PP. The second is a description of economic information, the core of which is the Economic Indicators Classifier (EIC), which covers: a) accounting indicators that come from primary data sources; b) statistical indicators, which are formed from registration indicators and / or features of statistical tools (for example, the number of women / men in the questionnaire is determined as a result on the summation of "gender"); c) analytical indicators, which are calculated from statistical indicators (for example, national accounts); d) other types of indicators (for example, administrative data obtained from external sources). These two blocks are related at the level of description of resources and results, the composition of which is determined by the EIC and related IOs. Given the importance of the EIC, to which the descriptions of data collection and results generation tools will be linked, the information in it should be systematized and structured.

2. Providing information on data in the DW or database (hereinafter - for generalization will be defined as DW), including a description of their composition, structure, presentation format, method of access and authority required for access, location, data semantics, source, owner and etc. The following rule must be followed: no data in the DW should appear

or disappear without proper fixation in the metadata storage and support system. MDR is a place where logically organized metadata is stored in a way that allows it to be retrieved on request, edited and managed.

3. Providing descriptive capabilities in the SOA communication process for both OLTP and OLAP systems.

4. Providing users with information on the context of data in the DW to better understand the sources of information, the nature of economic indicators provided as a result of the implementation of certain PP, as well as the actions performed by the PP.

5. Support for convenient search and navigation tools in MDR.

6. aggregation at least at the level of descriptions of all components of the IS required for the consistent use of IO.

7. Determining the information flows of the process scheme of information processing through the analysis of TP, resources and results of each PP with details to the level of EIC.

8. Support for data lifecycle stage descriptions in a format that allows the user with the appropriate authority to see the stage in real time to determine the ability to perform specific actions with data (for example, the use of data that has not yet been checked is not allowed).

9. Centralized storage ensuring in a single repository of all documents and metadata reports for all users and all needs, *manage and support different types of documents* that will be provided to the user for viewing from the appropriate catalog / list of documents, and create copies of documents.

10. When integrating data from different sources, it is necessary to *import metadata from these sources*. This also applies to the procedure of integration through extraction, transformation and loading (ETL). It should be used for "dirty" / raw metadata by copying their complete sets from source systems, transforming it into the desired format, if necessary, cleaning the metadata to improve its quality to further upload the resulting metadata to the MDR, where it will be stored and from where provided to the user.

11. *Export metadata and documents for different user needs in different formats*. Pre-export software applications and services should be able to view and edit a set of metadata and/or document.

12. Management and support of applied metadata, data quality indicators and PP performance, as well as providing the ability to view applied metadata through their "history".

13. Management and support of sets of conceptual definitions for different types of documents, which involves the formation of a list of metadata for conceptual definitions that can be provided for a particular type of document, as well as full and partial inheritance of metadata for future versions of documents. This means that each element of applied metadata must be able to be attached and used to documents related to different activities, and can be attached to a specific type of document to further generate copies of this type of document during the reference period as a certain period of time. applied metadata is true. When preparing a new document for a certain period and type of activity, it should be possible to automatically attach the required metadata element, and it should be possible to delete and renew the established link, for example, when creating a document for a new reference period.

14. Formation and maintenance of the life cycle status of applied metadata (prepared, published, eliminated / blocked). Published metadata can be updated, which means that it is valid for the period of time when it was updated. For the new reference period, the updated metadata can be made public and the old copy can be "closed" (locked for changes). Eliminated metadata should remain in the MBR, but as "closed" and, accordingly, should be provided a period of their relevance. Regarding "closed" metadata, the following should be noted:

– "closed" metadata cannot be changed and inherited (ie is the end of a chain of descriptions);

– the "closed" context remains in the MDR as long as it is attached to the valid documents for the old reporting periods.

Thus, metadata should be stored with the period of their relevance.

The implementation of certain functions will allow the use of MDR as a repository and source of metadata for all services and other software applications of IS support DW, as well as a comprehensive system of documentation of applied metadata and the results of the PP. Such a MDR can be the foundation of a platform that will remain usable in the long run and allow for further development. In the production process MDR must provide solutions to three important tasks: 1) control of the automated process of data processing from collection to dissemination; 2) providing internal and external users with the requested economic information in accordance with the granted access powers and all necessary information on the context of economic data, as well as their quality; 3) ensuring the coordinated accumulation of economic information in the DW. Solving the latter problem will be especially critical when replacing data collection sources (for example, when replacing own statistical observation data with administrative data), as it will be important to maintain a retrospective relationship of data and provide justification for the quality of such information.

Defining the basic functionality of MDR is important for the organization of reuse of services in the implementation of COA, as it provides a basis for organizing technical infrastructure and developing new or upgrading / linking existing services involved in the organization's IS. Services can cover the following tasks: design of collection tools, data exchange between organizational units and external partners, metadata management, maintenance of the register of data providers, data collection, storage of primary data, TP process management, search, research and analysis of data, processing data according to a certain methodology, dissemination and storage of disseminated data, IT security, identity and access management, quality assessment, support of IS catalogs, tracking user needs in data and metadata.

To form high-quality time series and compare information for different periods requires a consistent sequence of versions of documents, classifiers, conceptual definitions, etc . This can be achieved by providing a "history" of economic information in the DW. From the point of view of semantic compatibility, this approach will provide clarity of the exact meaning of information for exchange with any other software applications that were not developed specifically for this purpose, which, in turn, will allow systems to combine information with other information resources and process it in a meaningful way, and will allow to provide services to the user through a multilingual external interface. In addition, the MBR is the basis for technical compatibility in addressing the harmonization of computer systems and services provided by them, in particular, it includes open interfaces, services to connect to services for providing, exchanging and processing information, data integration and middleware tools access to data and meta-information, security services and other aspects.

In order to implement the definition of functionality above, it is necessary to carry out such preliminary preparatory actions, without which it is impossible to ensure high efficiency of MDR.

1. Creation of metadata catalogs, where information about all versions of all used metadata should be stored in a permanent place: classifiers, reference books, standards, etc. Such location fixation is important for determining information retrieval paths for both search and view tools and metadata and data.

2. Creation according to similar principles of data catalogs, which represent systematized information, which may include own primary data, macrodata obtained from these data, micro- and macrodata obtained from external sources, data obtained by processing a combination of these sources. These should be links to relevant metadata to ensure semantic compatibility. In addition, data catalogs should define the tools for accessing (relevant services) to this data and thus provide the opportunity to ensure technical compatibility.

3. In [9] it is recommended to group the metadata by IO, which will be integrated into the MDR. In order to organize the automated operation of IS on TP, it is necessary to create an appropriate platform that would initiate the launch of applications and services by meta-descriptions, which fix the parameters of their launch and the conditions of their activation. Accordingly, in the preparatory stage, it is necessary to prepare a scheme of binding IO to the relevant PP and the relevant meta-descriptions.

4. Organizing the existing documents on the organization and implementation of the production process, creating an electronic archive of these documents and creating an appropriate catalog to concentrate in one place all the necessary documentation.

5. Formation of a thesaurus as a terminological dictionary tied to certain IOs, which should be used to search for documents and necessary meta-information, including by catalogs.

6. Identify the characteristics that determine the direction of search in the IS (indicators, specific activities, types of observation units, etc.), and develop tools for thematic searches such as a list of keywords / indexes in combination with thematic headings and thesaurus terms.

7. Develop a description of the connection between frequently requested data and metadata to create a "semantic network" of IOs of different types with the definition of conceptual relationships between them, as well as a tool to access the information found. When presenting information in a web application, such a "semantic network" should provide explanatory information about the data selected by the user.

8. Defining administrative tools and access rights for different categories of users (for example, wide access for external users via the Internet to a certain range of MDR resources, narrow access for external users to RMD through internal search engine, access of internal users with different powers to act with MDR information).

In the process of analysis, systematization and streamlining of knowledge on IS metadata, the requirements for existing IS applications can be clarified, for example, the description of the data collection tools may provide such additional information necessary for the implementation of the PP as a list of reference books and classifiers, the relevance of which should be ensured at the beginning of the PP.

With regard to methodological documents and the descriptive part, it is advisable to create a catalog of references to the storage of electronic versions of relevant documents (for example, on the methodology of quality assessment, or descriptive part of algorithms for calculating economic indicators). All these documents should be systematized, collected in one place, creating an electronic library - archive of methodological documents, tools, descriptions of results, etc., to typologize documents and normative reference information, as well as create a catalog to this archive, supplementing it with a list of categories for easy navigation and search.

There should be a reference book for TP with a list of IO resources and results, and instances of these IOs (specific tools and products / information) need to be allocated to separate lists / reference books with linkage to the relevant methodological support. For example, an IO can be defined as a "database table", have a corresponding list of attributes, and a specific PP is associated with an instance of this IO, for which the location in the IS is determined, in particular, specified physical location in a particular database.

Providing the user with a "history" of data and relevant metadata lies in the development of the MDR architecture, taking into account the retrospective preservation of data and metadata, as well as the need to analyze their life cycle status. This important point in the creation of IP will require, in addition to the inclusion in many IOs of the period of their validity / relevance and current status of the life cycle, also in the software applications appropriate analysis of this information and choice of algorithm. Therefore, this requires careful planning of the development / modernization of MDR with the gradual complication of the tasks of retrospective presentation of data while maintaining the current functionality.

Conclusions and propositions. In IS, metadata as structured data that describes the characteristics of IS and helps to identify, locate, evaluate and manage these objects, not only simplify management, query, ensure full use and understanding of data, but also allow you to generate user-friendly information and provide the user with a tool that promotes better analysis. Metadata is used as an add-on bridge between PP to ensure continuous quality control of the latter. The use of MBR as the only storage location of reference metadata for all users and all needs, respectively, for all documents and reports provided, provides the following important advantages:

– fix this metadata once in one place for all users;

– simplification of the creation of bilingual / multilingual descriptions (for example, Ukrainian and English) provided that the translation of metadata is provided in the MDR;

– minimization of errors due to a single entry of metadata and updating them in one place, including prevention of duplication of information;

– ensuring the availability of metadata ready for use in the context of the information needs of internal and external users;

– reducing the burden on those responsible for performing certain production processes by determining the internal and external requirements for metadata, data quality indicators and the quality of PP performance during the technological process;

– prevention of potential inconsistencies in the external presentation (for example, on the website);

– providing the history of any metadata for all types of production activities of the organization, while preserving the history of changes in metadata;

– support for communication between metadata;

– facilitating the preparation and storage of metadata documents;

– saving time and money in the implementation of support and maintenance of metadata in general for the organization.

MDR, formed as a single meta-information base based on a single system of classification and coding of information, together with the introduction of standardization and unification of description and implementation of the process scheme of production organization, becomes a tool to ensure the integrity of IS. This approach will help in the future to create new functionality, subsystems and other components without disrupting the automated IS management system, will also allow the use of typical, standard and unified elements and design solutions that optimize IS operation technology, unify methods and techniques used user when working with IS. In addition, this approach will establish the security of information processing at the level of MDR, which will ensure the protection of information flows in the process of processing, storage and exchange of information through the distribution of access and authority in the system.

References

1. *Guidelines for the Modeling of Statistical Data and Metadata. Conference of European statisticians methodological material*. (1995). Geneva: United Nations. http://www.unece.org/fileadmin/DAM/stats/publications/metadatamodeling.pdf.

2. Stefan Berg, Klas Blomqvist, Eva Holm, Lars-Göran Lundell, Henrik Lundström, Thomas Nyberg, & Jens Olofsson. (2015). Case Study: Use of GSIM – Statistics Sweden's model for a Central Metadata Repository. *United Nations Economic Commission for Europe Conference of European Statisticians Workshop of the Modernisation Committee on Standards: International Collaboration for Standards-Based Modernisation* (Geneva, Switzerland, 5–7 May, 2015).

3. Mauro Scanu, Cecilia Casagrande. (2016). GSIM and the Sistema Unitario dei Metadati: State of application of the standard. *United Nations Economic Commission for Europe Conference of European Statisticians Workshop on Implementing Standards for Statistical Modernisation* (Geneva, Switzerland, 21-23 September 2016). http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2016/mtg4/Paper_17_Italy_-_The_Generic_Statistical_Information_Model__GSIM__and_the_Sistema_Unitario.pdf.

4. A modernization project in the INE of: Microdata repository and structural metadata. Prepared by Ana I.Sanchez-Luengo, María del Mar Ballesteros, Aranzazu Baltanas Dept. Methodology and Development of Statistical Production Statistics Spain (INE) Paseo de la Castellana, 183 28046 Madrid (Spain). *United Nations Economic Commission for Europe Conference of European Statisticians Workshop of the Modernisation Committee on Standards: ModernStats World Workshop 2018* (Geneva, Switzerland, 11-13 April 2018). https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2018/mtg1/Spain_SANCHEZ-LUENGO_Paper.pdf.

5. Magnus Eriksson The Register Utilisation Tool: A Practical implementation of GSIM as support in register-based research. *United Nations Economic Commission for Europe Conference of European Statisticians Workshop on Implementing Standards for Statistical Modernisation* (Geneva, Switzerland, 21-23 September 2016). http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2016/mtg4/Paper_15_Sweden.pdf.

6. In-depth review of exchange of economic data and data sharing. Note by Statistics Finland. *United Nations Economic Commission for Europe Conference of European Statisticians Sixty-fifth plenary session* (Geneva, Switzerland, 19-21 June 2017). https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2017/CES_10-In-depth_review_of_Data_Exchange.pdf.

7. In-depth review of data integration. Note by the High-Level Group for the Modernisation of Official Statistics. *United Nations Economic Commission for Europe Conference of European Statisticians Sixty-fifth plenary session* (Geneva, Switzerland, 19-21 June 2017). https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2017/CES_8-In-depth_review_of_data_integration.pdf.

8. Mogens Grosen Nielsen and Martin Lundø. Implementation of quality declarations at Statistics Denmark using DDI, SDMX, GSBPM and other standards. Working paper. *United Nations Economic Commission for Europe Conference of European Statisticians Workshop on the Modernisation of Statistical Production* (Meeting, 15-17 April 2015).

9. Generic Statistical Information Model (GSIM): Specification (Version 1.1, December 2013). https://statswiki.unece.org/display/gsim/GSIM+Specification.

Список використаних джерел

1. Guidelines for the Modeling of Statistical Data and Metadata. Conference of European statisticians methodological material. – Geneva: United Nations, 1995. – 31 p. – Aceeseed mode: http://www.unece.org/fileadmin/DAM/stats/publications/metadatamodeling.pdf.

2. Case Study: Use of GSIM – Statistics Sweden's model for a Central Metadata Repository / Stefan Berg, Klas Blomqvist, Eva Holm, Lars-Göran Lundell, Henrik Lundström, Thomas Nyberg and Jens Olofsson // United Nations Economic Commission for Europe Conference of European Statisticians Workshop of the Modernisation Committee on Standards: International Collaboration for Standards-Based Modernisation (Geneva, Switzerland, 5–7 May, 2015). – Geneva, 2015. – 18 p.

3. Mauro Scanu, Cecilia Casagrande. GSIM and the Sistema Unitario dei Metadati: State of application of the standard // United Nations Economic Commission for Europe Conference of European Statisticians Workshop on Implementing Standards for Statistical Modernisation (Geneva, Switzerland, 21-23 September 2016). – Geneva, 2016. – 12 p. – Aceeseed mode: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2016/mtg4/Paper_17_Italy_-_The_Generic_Statistical_Information_Model__GSIM__and_the_Sistema_Unitario.pdf.

4. A modernization project in the INE of: Microdata repository and structural metadata. Prepared by Ana I.Sanchez-Luengo, María del Mar Ballesteros, Aranzazu Baltanas Dept. Methodology and Development of Statistical Production Statistics Spain (INE) Paseo de la Castellana, 183 28046 Madrid (Spain). United Nations Economic Commission for Europe Conference of European Statisticians Workshop of the Modernisation Committee on Standards: ModernStats World Workshop 2018 (Geneva, Switzerland,11-13 April 2018). – Geneva, 2018. – 5 p. Aceeseed mode: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2018/mtg1/Spain_SANCHEZ-LUENGO_Paper.pdf.

5. Magnus Eriksson The Register Utilisation Tool: A Practical implementation of GSIM as support in register-based research / United Nations Economic Commission for Europe Conference of European Statisticians Workshop on Implementing Standards for Statistical Modernisation. Geneva, Switzerland, 21-23 September 2016). – Geneva, 2016. – 11 p. – Aceeseed mode: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2016/mtg4/Paper_15_Sweden.pdf.

6. In-depth review of exchange of economic data and data sharing. Note by Statistics Finland / United Nations Economic Commission for Europe Conference of European Statisticians Sixty-fifth plenary session (Geneva, Switzerland, 19-21 June 2017). – Geneva, 2017. – 23 p. – Accessed mode: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2017/CES_10-In-depth_review_of_Data_Exchange.pdf.

7. In-depth review of data integration. Note by the High-Level Group for the Modernisation of Official Statistics / United Nations Economic Commission for Europe Conference of European Statisticians Sixty-fifth plenary session (Geneva, Switzerland, 19-21 June 2017). – Geneva, 2017. – 23 p. – Aceeseed mode: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2017/CES_8-In-depth_review_of_data_integration.pdf.

8. Mogens Grosen Nielsen and Martin Lundø. Implementation of quality declarations at Statistics Denmark using DDI, SDMX, GSBPM and other standards. Working paper. / United Nations Economic Commission for Europe Conference of European Statisticians Workshop on the Modernisation of Statistical Production (Meeting, 15-17 April 2015). - 6 p.

9. Generic Statistical Information Model (GSIM): Specification (Version 1.1, December 2013) – 236 p. – Aceeseed mode: https://statswiki.unece.org/display/gsim/GSIM+Specification.

Отримано 28.01.2022

УДК 303.064

Тетяна Лумпова
кандидат економічних наук, викладач
ВСП «Опитико-механічний фаховий коледж КНУ ім. Т. Г. Шевченка» (Київ, Україна)
E-mail: taivlu@meta.ua. **ORCID:** https://orcid.org/0000-0002-7576-4289
Scopus Author ID: 0000-0002-7576-4289

ВИКОРИСТАННЯ ДОСВІДУ МІЖНАРОДНОЇ СТАТИСТИКИ ЩОДО СТВОРЕННЯ РЕПОЗИТАРІЮ ПРИКЛАДНИХ МЕТАДАНИХ ДЛЯ ІНФОРМАЦІЙНИХ СИСТЕМ ЕКОНОМІЧНОГО СПРЯМУВАННЯ

Впровадження процесного підходу у виробництво в організаціях, які займаються обробленням статистичної та економічної інформації для аналітичних досліджень, перехід в інформаційних системах до сервісорієнтованої архітектури, інтеграція даних із різних джерел, до яких залучаються показники, отримані з "великих даних", використання сховищ даних спричинило потребу в розробленні нових підходів до створення та модернізації інформаційних систем. Поряд з цим підвищуються вимоги до метаінформаційної складової ІС з точки зору задоволення вимог внутрішніх та зовнішніх користувачів щодо можливості самостійного оцінювання якості даних, насамперед релевантності, повноти, цілісності та несуперечливості. Всі ці виклики вимагають підсилення метаінформаційної складової ІС, зокрема, створення або модернізації (за наявності) репозитарію метаданих для підтримки та використання метаданих різними групами користувачів, а також для забезпечення автоматизованого централізованого керування виробничим процесом.

Не менш важливою позитивною тенденцією розвитку інформаційних технологій в організаціях, які займаються обробленням економічної інформації з різних джерел, є використання процесного підходу до реалізації виробничого процесу. Застосування процесного підходу передбачає уніфікацію технологічного процесу оброблення інформації, розкладання його на складові, виконання яких не прив'язано до конкретного обстеження. При наявності репозитарію метаданих стає можливим фіксація важливих проміжних результатів для подальшого використання в інших дослідженнях. Для такого використання необхідним є збереження контексту отримання даних (період обстеження, вибірка, характеристики якості та ін.).

Для формування / модернізації інформаційних систем на засадах використання репозитарію метаданих доцільно звернутися до міжнародного досвіду, зокрема до досвіду, узагальненого протягом останніх років Європейською економічною комісією Організації Об'єднаних Націй на своїх пленарних сесіях, нарадах, конференціях, робочих засіданнях щодо модернізації ІС в галузі статистики та економічних досліджень. У статті розглядаються ключові моменти, які потрібно враховувати при адаптації цього досвіду до власних інформаційних систем. Ця інформація буде корисною при створенні репозитарію метаданих, сховища даних, інформаційного порталу щодо статистичних та економічних досліджень.

Ключові слова: *класифікатор економічних показників; метаінформація; репозитарій метаданих; сховища даних; якість даних.*

Бібл.: 9.

Lumpova, T. (2022). Using the experience of international statistics to create a repository of applied metadata for information systems in economic area. *Problems and prospects of economics and management*, (1(29)), 39-48.