Lumpova, Tetyana; Kasianchuk, Ihor

Article

Finding a conceptual approach to developing an architecture of general-purpose services for economic researches

Technology audit and production reserves

Provided in Cooperation with:
ZBW OAS

**Tetyana Lumpova,
Ihor Kasianchuk**

FINDING A CONCEPTUAL APPROACH TO DEVELOPING AN ARCHITECTURE OF GENERAL-PURPOSE SERVICES FOR ECONOMIC RESEARCHES

The object of research is the architecture of general-purpose services for economic researches. The excessive popularity of service-oriented architecture implementations has led to consideration of the application of this approach in the production of organizations engaged in processing statistical and economic information for analytical research. This is due to the possibility of reusing services, each of which is designed to perform specific statistical functions, and the availability of design solutions for compiling these functions into a program for implementing a statistical task solution, using orchestration as a basis for dynamic service calls.

The paper proposes an approach to the development of external statistical services as a set of software products for solving analytical tasks by groups of independent researchers. Since such researchers do not have sufficient financial support, the use of external statistical services will allow them to conduct an in-depth analysis of their research results. The proposed approach assumes that such a service should have its typical data storage structure with fixed types and attributes, to which the researcher imports their own data, selects a calculation execution scheme, initializes their execution, and receives results. Taking into account this storage scheme, the authors have identified the main groups of metadata necessary to describe the implementation of the statistical process. A statistical business process typically consists of a set of procedures that implement certain statistical operations and data transformations, and the provided set of metadata allows for the realization of the process goal and activity.

Issues that require further in-depth research for the practical implementation of the proposed idea have been identified. The use of service-oriented architecture for building statistical processes allows for standardizing the approach by unifying the functions of individual services. The availability of possibilities for practical implementation of software products in a service-oriented architecture using a cloud environment in modern IT industry serves as a guide for further research direction.

Keywords: *service-oriented architecture, statistical functions, processing information, statistical and economic information.*

1. Introduction

In today's world, statistics cannot be overlooked when it comes to revealing the essence, patterns, and trends in the development of economic processes at the level of both the country and a specific industry or territory. In organizations involved in analytical research in the economic sphere, the production process or, as it is often called, the business process, almost always includes statistical calculations, which can be considered as statistical business functions. In this regard, the question arises about creating publicly available tools for performing such functions by compiling them into a specific statistical program as a plan/algorithm for solving a particular statistical problem (for example, forming a table of price volume dynamics and changes in tariffs for paid services by region). To implement such statistical programs,

it is advisable to pay attention to service-oriented architecture (SOA) for creating software services that implement the execution of statistical business functions, from which a statistical process can be composed using metadata. The peculiarity of SOA is the ability to present the architecture as a set of services and processes that can be combined and form a scheme of calls in accordance with the requirements of the business process, using planning and task flow formation tools for execution. The transition to such technology will require the restructuring of the business process organization by unifying specific business functions from the point of view of their ability to be used as components that ensure standardization of the production process.

Analyzing recent studies and publications [1–6], it can be concluded that the process approach to processing economic information is becoming increasingly popular in various

fields, including statistics. Its implementation allows for standardization of information processing procedures and provides higher quality results. In particular, the use of CSPA, GAMSO, GSBPM, GSIM, LIM, and GSDEMs allows statistical organizations to meet the requirements of the modern market, provide more accurate and understandable statistical information, as well as reduce the time and costs of its collection and processing. However, it should be noted that the use of CSPA requires sufficient resources and technical knowledge. Organizations may consider CSPA for their own needs without being tied to the statistical industry, but it is important to take into account the specificities of their activities and resource capabilities. Overall, the implementation of a process approach to processing economic information, including using CSPA and other proven approaches, can help organizations achieve better results and respond to the challenges of the modern market.

The European Statistical System (ESS) proposes a series of services within the SERV project [2] that are compatible with the Common Statistical Production Architecture (CSPA) to support the integration of statistical services into the production processes of national, ESS and ECE level statistical data. These services include:
– Structural validation of statistical data files according to the Statistical Data and Metadata eXchange (SDMX) information model for a specific data flow.
– Harmonized transformation interface for several implementations of file transformation services (converters).
– A set of methods for time series disaggregation and comparative analysis, which represent a set of plugins for the independent platform JDemetra+.

It should be noted that these services are provided to registered users and are focused on ECC standards, including CSPA, which has become an unofficial corporate standard that statistical organizations adjust to. From a methodological perspective, this is undoubtedly a powerful toolkit.

Thus, *the aim of this research* is to define a conceptual approach to creating statistical tools for independent researchers who can use them to process data from their own research.

2. Materials and Methods

The object of research is the architecture of general-purpose services for economic researches.

When accumulating data from various sources in a data warehouse, problems arise not only of a methodological nature (for example, harmonization of metadata, classifiers, directories), which are not the subject of consideration, but also problems of extracting data from the warehouse according to the researcher's needs. When developing a data warehouse with the corresponding support tools, appropriate visualization tools are usually provided to generate queries. However, even with such an approach, problems remain regarding the preparation of analytical information with the performance of corresponding statistical calculations. Typically, information obtained from a query is imported into databases of specialized statistical programs (e. g., SPSS). For small research groups, data storage is usually carried out in poorly structured databases that still have insufficient support from information technology (IT) professionals. It is precisely this group of researchers who require simple standard data processing tools.

For processing the results of economic and marketing research, mostly the same procedures of control and data processing are used as in statistics. These procedures include:
– «Primary» control, where the compliance of data with the permissible value range and the consistency of economic indicators within the document describing the phenomenon for a particular research unit (such as a company or territory) is checked.
– Control of data consistency over observation periods, where the current data is compared with the data from the previous period, and deviations are identified, the magnitude of which may question the reliability of the data.
– Formation of dynamic series.
– Calculation of derived data based on available information (e. g., calculation of averages).
– Execution of data grouping and aggregation, where new, more generalized groups are formed based on the existing breakdown by certain groups, and additional summaries are calculated.
– Execution of regulated requests for obtaining horizontal and vertical cuts of data in the information database.

These are only the most general approaches to processing information for preparing data for analytical conclusions.

All the procedures mentioned above can be implemented through standard services without direct binding to a specific DB, providing means for formalizing actions. These descriptions should include data source descriptions – names of fields and tables, descriptions of actions, and descriptions of result representations, which may also include names of fields and tables where results will be written, if necessary. Based on these descriptions, an SQL query should be generated and its execution initiated.

3. Results and Discussion

In [3], it was determined that a statistical service should perform one or several actions to solve statistical tasks through a clearly defined interface, with specific results for a given set of input resources. The statistical service may have different levels of detail, divided by complexity into atomic and aggregated. An atomic statistical service concentrates a small part of functionality, accepting input parameters and data from the outside and providing a result. An aggregated statistical service may contain a broader functionality, consisting of atomic statistical services, and even cover a certain production process. To create an appropriate service as a reusable software application, the problem question is to obtain resources from an information repository.

The data in an information repository is displayed as hierarchical structures that have clear logic. To support the procedure for creating descriptions, which should be used by statistical services, it is necessary to identify groups of metadata for which it is advisable to create a repository. Creating a metadata repository will achieve independence from specific data and software, providing advantages from the reuse of data. For analytical research, data warehouses are typically used, which are subject-oriented, integrated, time-variant, and a non-changing set of data intended to support decision-making. The concept of a multidimensional information space or hypercube, where numerical indicators that are analyzed are stored in cells, underlies the information repository. The dimensions (axes) of the hypercube are represented by a set of values for the feature being analyzed. Thus, an indicator is a cell in the hypercube, the value of

which is unambiguously determined by a fixed set of dimensions characterizing a particular fact. Note that relational databases are mostly used to accumulate economic data at the present stage, where multidimensionality is supported by using certain schemes for organizing relational tables.

It should be noted that relational databases are mostly used for accumulating economic data at the present stage, where multidimensionality is supported by using certain schemes for organizing relational tables. The most common and tested schemes are the «star» and «snowflake» schemes, with the «star» scheme considered more effective in terms of performance tuning and ensuring the most efficient servicing of queries in a data warehouse due to high search operation performance [7]. The essence of this scheme is that for storing the analysis attribute, it is separated from the actual data, forming an inverted organization of data storage where a reference is stored instead of the attribute. In further discussions, the author will refer to the use of the «star» scheme. A multidimensional conceptual view of data is a natural perspective of an economist on the research object. It is a multiple perspective consisting of several independent dimensions, through which certain sets of economic data can be analyzed, including simultaneous analysis across multiple dimensions in solving multidimensional analysis tasks. Each dimension contains directions for data consolidation, which consist of a series of sequential levels of generalization, where each higher level corresponds to a higher degree of data aggregation along that dimension.

An important aspect of implementation in application development can be identified as the need to develop statistical services as a set of complex software products for conducting statistical calculations in an SD environment. For small groups of independent researchers who do not have sufficient financial resources, the use of external statistical services will provide the opportunity for in-depth analysis. Such a service should have a typical data storage structure with fixed data types and attributes, to which the researcher exports their own data, selects the calculation execution scheme, initializes their execution, and receives the results.

Taking into account this storage scheme, let's determine the main groups of metadata necessary to describe the implementation of the statistical service process:

1. *Metadata that describe actions on data* are descriptions of data sources and transformations that need to be performed when loading data into a data warehouse. This includes descriptions of elementary steps in the workflow process, as well as constraints on action execution parameters (group «Actions»). Each workflow step is a specific task, the execution of which determines a certain transformation procedure, control of existing data, or creation/calculation of new data. Here, transformation refers to the implementation of correction of existing data using predefined algorithms.

2. *Metadata that describe entities of the subject area*, the projection of which is the objects of the data warehouse. These metadata determine the physical types of attributes and provide descriptions of the structures and attributes of dimensions and dimension types, names, logical and physical names of schemas when organizing the structure according to the «star» schema and accumulated tables, descriptions of their attributes and relationships with dimensions (group «Entities»). These metadata specify the steps of the processes in the «Actions» group.

3. *Metadata that describe user interaction with the data warehouse*, that is, describe the relationship to the data warehouse of different categories of people who work with the data warehouse in different stages of its life cycle. These metadata determine the user's access to information and other characteristics of database security (group «Users»).

4. *Metadata that describe the driving forces of data warehouse creation and development* are metadata that describe the goals and tasks that cause the creation and formation of a data warehouse. These are data that allow planning work with the data warehouse, managing customer information services of the data warehouse, controlling data quality, consistency of components, etc. (group «Plan»). These metadata determine the logic of process management and may therefore contain elementary and/or structured actions described in the «Actions» group. Essentially, the «Plan» group determines the task flow.

5. *Metadata that describe the temporal aspects of data warehouse operation* provide the ability to track the chain of events and time intervals that are important for data warehouse clients, are metadata that define the time characteristics of data and the order of events in the data warehouse (group «Time»). These metadata are important for decision-making and analysis based on the data warehouse.

6. *Metadata describe the data types of a database* that correspond to real data types and can define the data domain as a set of possible attribute values. They also allow for the introduction of new data types, the conversions of which can be described through a group of «Actions» (the «Types» group). Defining a custom data type makes it possible to establish additional checks at the domain level and simplifies the ability to set the data type for identical objects in the subject area. For example, it is possible to introduce a domain «Balance account» for 4-digit numbers and set restrictions on the possible values as «values from the range of 1000–9999».

A statistical business process typically consists of a set of procedures that implement certain statistical operations and data transformations. From the perspective of the metadata provided, the realization of the process goal and the activity itself are described by the «Actions» group, resources are described by the «Entities» group, the sequence of activities is described by «Plan» and «Time», while the «Users» and «Types» groups establish additional constraints on the execution of the business process. It should be noted that different statistical business processes can be formed from a set of procedures, meaning that these procedures can be considered as business functions, which can then be used to create various statistical business processes. It is advisable to create a knowledge base from these business functions, which contains the rules for calculation described through metadata and is tied to the structure of the database. For the end user, an economist, it is advisable to provide the knowledge base in the form of a catalog with a list of functions, each of which is accompanied by a description of the parameters and results, as well as a corresponding methodological description, which allows not only to determine the calculation method but also the prerequisites for its implementation and a description of the expected results and, possibly, recommendations for their further use.

Metadata communication in the production process can be described as follows. The «Plan» and «Time» groups define the sequence of actions defined in the «Actions» group, with information described in the «Entity» group, and taking into account the constraints on user access to the data set (SD) defined by the «Users» group. When

using services to perform specific operations defined in the «Actions» group, orchestration as the dynamic creation of a schema of service calls is formed by the remaining groups of metadata. For example, a specific «action» is an operation that uses input data (the «Entity» group) and parameters to perform data transformations described by the instance description in the «Action» group, and the output data result is also defined by the «Entity» group description.

Directly creating metadata descriptions will be a non-trivial task because end-users not connected to the IT field must be able to independently create a plan for implementing research result processing without delving into technical details and accordingly launch the execution of a defined workflow, using changes in data selection parameters. Another way to create an execution plan is through the automatic discovery of a sequence of services, each responsible for achieving a specific goal. The approach used in [4] involves defining the ontology of each service, based on which the process executor finds the optimal path to achieve the desired outcome.

The ontology of a web service is built upon three interconnected sub-ontologies, namely «profile», «process model», and «grounding» (Fig. 1). The profile ontology is used to express «what the service does» for advertising, query construction, and search purposes. The process model describes «how the service works» to enable activation, interaction, composition, monitoring, and recovery. The grounding represents the components of the process model, considering detailed specifications of message formats, protocols, etc. All these sub-ontologies are combined into the top-level concept «Service», which refers to the aforementioned three components (Service Profile, Service Model, Service Grounding) through properties such as «presents», «described by», and «supports» [5].
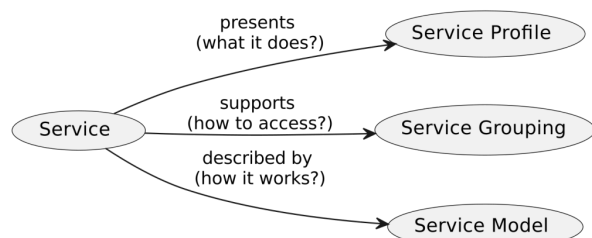

Fig. 1. Semantic Web Service Ontology Template

The most important type of information represented by the profile and playing a key role in the service discovery process is the service's functionality aspects. This includes the transformation of information, represented by the inputs and outputs of the service, and the change of states during execution, represented by the preconditions and expected effect. In this way, functional properties describe the inputs, outputs, preconditions, and effects of the service, commonly referred to as IOPE (inputs/outputs/preconditions/effect). On the other hand, non-functional properties describe semi-structured information intended for human users. Examples of non-functional properties include the service's name, service parameters, security requirements, service quality, geographical data, and more. These properties provide additional information beyond the functional aspects, catering to the needs and preferences of the service users.

Semantic annotation of services enables the organization of various steps involved in achieving a customized task, including service discovery, selection, composition, mediation, execution, monitoring, replacement, and compensation. These steps collectively form the execution process of semantic web services. The fundamental procedure is the discovery of a service, which involves finding a relevant service from a set of services available in a central or distributed service repository. Once multiple suitable services are discovered, the most appropriate service can be selected and invoked/referenced. The first step in the service discovery process is comparing service categories. The service category is stored in the Service Profile, providing information that aids in the comparison and selection of services. This value is compared to the service category in the user query. If a match is found between the categories, the web service is selected for participation in the subsequent discovery phase.

The degree of functionality match is calculated in the second stage of service functionality matching. The attributes «hasInput», «hasOutput», «hasPrecondition», and «hasResult» are compared with those present in the user-requested service. These attributes are defined in the Service Profile. In the third stage, the degree of Quality of Service (QoS) match is computed based on evaluating response time and system reliability in service discovery. After completing these three stages, a list of services that best fit the request is published.

Semantic approaches allow for the detection of similarity or closeness between descriptions, even in the presence of syntactic differences. The capabilities of a service are described using the Service Profile class, so the available and requested services are compared based on their profiles. The comparison algorithm matches when all the outputs of the requested service match the outputs of the available service, and all the inputs of the available service match the inputs of the request. This ensures that the available service satisfies all the requirements of the request. Existing semantic search approaches based on ontology typically rely on assessing conceptual similarity between ontology elements [6]. If services and requests are represented as sets of triplets, semantic similarity calculations between the triplets can be performed. The similarity score between a service and a request is a numerical value that expresses the degree of similarity between them.

Therefore, the operation of the service should be based on already formed metadata descriptions of resource usage, action implementation, and result formation. The process of metadata description formation will require IT specialists to work with the corresponding service and provide a detailed and qualified description of the task and the possibility of involving an economist if necessary. Assuming that the description structures are developed, the formation of the description and implementation of the production process will be divided into stages, for which services can be used to perform the following actions:

1. Provides a user interface for an average user to create a scenario for executing a working process based on the meta-descriptions.

2. Establishing a connection to the database.

3. Checking the availability of tables and attributes defined in the meta-description (the necessity of this step is debatable).

4. Getting current parameters (data selection period, data selection categories, etc.) from the user and verifying the logical consistency of these parameters with the data available in the database.

5. Generating and executing SQL queries and presenting the obtained results in accordance with the available

meta-descriptions (saving in the database, generating a text document, etc.).

Let's consider an example of using metadata groups to implement a business process for calculating the trend and seasonal fluctuations of product sales [8]. Let's define the sequence of steps for implementing the calculation algorithm in general terms as a basis for determining the data manipulation actions and forming the metadata group:

1. Check user permissions for data access and modification using the «Users» group description.

2. Obtain data from the database on sales of goods by quarters of the year, using the descriptions from the «Entities» and «Time» groups.

3. Verify the results from Step 1 for consistency of units of measurement and homogeneity of time intervals and completeness of ranges (for example, quarterly breakdowns for several years), using the description from the «Actions» group.

4. If the results of the verification are positive (negative results will be disregarded), let's perform the calculations in the following steps:

– calculate the trend (using descriptions from the «Entities» and «Actions» groups);

– calculate the seasonality indexes (using descriptions from the «Entities» and «Actions» groups);

– calculate the trend adjusted for seasonality (using descriptions from the «Entities» and «Actions» groups).

5. Save and/or provide the results to the user, using the descriptions from the «Entities» and «Actions» groups.

In general, this and similar calculation execution schemes in SOA are illustrated in Fig. 2, where a database with observation results requiring statistical processing is assumed to be present (denoted as «Statistics repository» in Fig. 2).

Working with the service is performed by the user through an interface defined in Fig. 2 as the «metadata terminal». Interaction with the interface is divided into two groups of operations: metadata definition and process execution.

The definition of metadata involves the construction of a workflow by an administrator (a non-IT specialist) according to the statistical program. It should be noted that the statistical program can contain both aggregated and atomic statistical services. Atomic services are formed first, and then complex services are formed from them. This information in the repository becomes a certain knowledge base that can be used in creating other statistical programs.

Using the user interface, the administrator describes metadata groups and the execution plan. After that, the system performs the following stages of metadata determination:

– process validation – verification of the conformity of input and output data in operations; correctness of transformations, including data type conversion; compliance of data descriptions with those defined in the CD;

– saving process metadata – transforming the metadata defined by the administrator into a formal process description that can be reproduced from the metadata repository during the execution phase;

– creating a default user meta-description with access rights – an account that has access to all statistical data for primary execution or process testing. It is assumed that the system allows creating descriptions for new users with different access rights both to data and to transformation results.
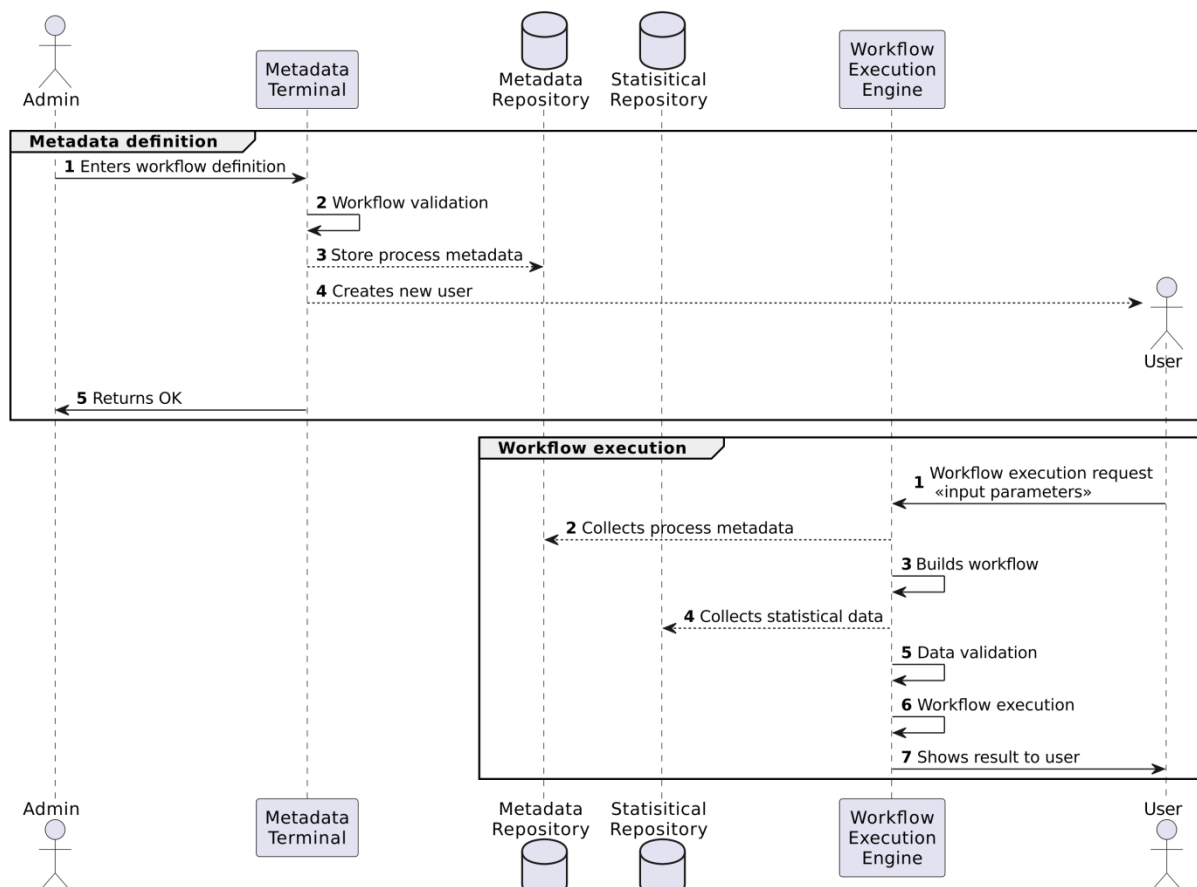

Fig. 2. Sequence of actions when working with a metadata-driven service

After creating the metadata descriptions, the user gains access to execute the process by joining the system with the corresponding account. The group of operations (represented in Fig. 2 as «Process Execution») allows actions to be carried out for processing the results of statistical observation for a given time frame and input data. The execution is initiated by the control program (labeled as «Process Execution Engine» in Fig. 2) and is carried out according to the following scheme:

– receiving a request for execution from the user, who was created by the administrator during the metadata construction phase;

– the process description is obtained from the metadata repository and reproduced as internal objects. In programming, this stage has a similar name called compilation, which is the transformation of the program code (metadata objects) into a sequence of instructions for execution. If the process is performed regularly and metadata changes infrequently, this allows for optimization of the stage through caching – executing pre-generated instructions without retrieving them again from the repository;

– validation of input data according to the received process description;

– execution of the process and returning the result to the user.

The calculation scheme is defined by the «Plan» group. Fig. 3 illustrates the execution of this sequence of actions.

The execution of the process and the return of the result to the user depend on the objects obtained from the metadata repository and the data from the statistics repository. The «User» and «Time» groups are independent from others, as they are directly associated with the execution plan. On the other hand, the «Plan» group is represented by a sequence of «Actions» that depend on «Entities» and «Data Types».

The «Data Types» group can include both standard types (e. g., integers, floats) and specific types (e. g., percentages, deciles), which are selected when creating an entity description. Entities are considered to have two main types at the logical level for describing the criteria used in the actions: a specific value (e. g., average value of a variable over a year) and a vector (e. g., a set of values from which the average is calculated). The «Actions» group consists of the description of performing a specific statistical service, where the result can also be a variable and/or a vector.

A similar architecture for managing the workflow process is discussed in [9]. According to the concept, a business process is defined using a knowledge base that describes a specific domain. An administrator, who is not an IT expert, sets up the base by defining entities, actions, their relationships, interaction formats, etc., and registers and describes the services. The user, in turn, provides the description and purpose of the workflow processes, sets input parameters according to the domain. The execution of the workflow process follows a declarative programming concept, where the program represents the description of the expected result, as opposed to an imperative approach that directly implements instructions. Therefore, before executing the «Plan» the most optimal route for executing the workflow process is determined, composed of services, each responsible for a specific objective. The determination of the service sequence is based on a similarity indicator, with the algorithm being constructed by comparing the characteristics of input and output data, conditions, and the effects of workflow states on the output data.

Thus, the solutions presented in [9] can be utilized in creating a standard shared service for implementing statistical services. However, practical implementation of the proposed idea requires addressing the following issues, which have been overlooked and need further investigation:

1. *Data storage scheme in the Statistics Repository and its location:* This decision will have a significant impact on all types of metadata and the ability to reuse statistical services by other users. Two main approaches are possible. The first approach is to have the Statistics Repository located at the user's side, which would make the metadata description of entity structures complex due to the need for synchronizing metadata with the actual state of the user's data. The second approach is to create the Statistics Repository on the service side by offering a standardized structure or multiple structures. In both cases, ensuring the confidentiality of statistical data becomes an important consideration.

2. *Storage of the statistical service context:* This includes the methodology, calculation application specifics, rules for evaluating result quality, and other related information. In the context of statistical research, as defined in [10], it is necessary to store declarative metadata, which carries the main methodological load. According to Bo Sungren's definition [10], these metadata outline measurement aspects of statistical data quality, such as content (as a prerequisite for assessing data reliability), accuracy (as a prerequisite for precision and reliability), and data availability (as a prerequisite for user accessibility). Let's define them as statistical metadata.
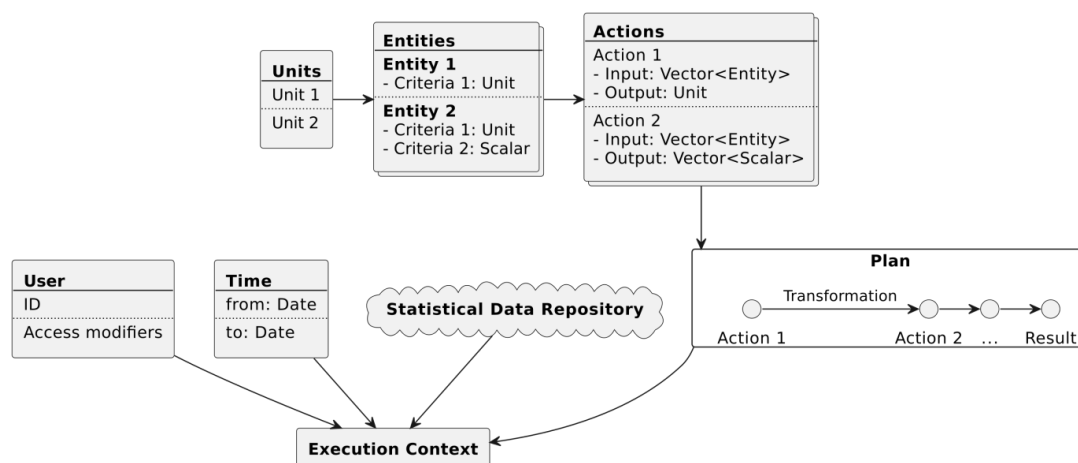

Fig. 3. Sequence of metadata usage in working with the service

The combination of these metadata values provides users with the ability to assess the usefulness of the data for their intended purpose. It is important to determine the need for preserving statistical metadata (e. g., based on user requests when using the Statistics Repository on the service side) and establish the structure of the corresponding entities.

3. *Preservation of the context of result acquisition:* This involves storing information related to sample parameters, quality assessments, and the method used for performing the statistical service. While this aspect is partially related to the previous point, it also emphasizes the importance of documenting the method and its application specifics.

4. *Cataloging of statistical services:* The utilization of a statistical service will contribute to accumulating metadata descriptions of both atomic and aggregated statistical operations, which can be reused when using the Statistics Repository with a standardized structure. Descriptions of accumulated statistical calculation algorithms, representing statistical services, can be stored in a separate catalog, provided to users for utilization, or used as a basis for creating new calculations.

The proposed conceptual approach to creating statistical tools requires further elaboration in terms of the identified research questions. From the issues discussed in the article, the following questions will require clarification for practical implementation:

– Grouping of metadata and defining the relationships between them. To describe the business process, it is advisable to categorize metadata based on their purpose of use. The discussed groups describe the entities on which the process is performed, the sequence of actions (plan) for conducting statistical calculations, and the metadata responsible for user access rights.

– Defining roles (administrator and user) in interacting with services and metadata. Creating and managing the business process is carried out by the user, who may not necessarily have IT skills. Actions for calculations are stored in a pre-existing knowledge base, and the processes of validation and data integrity are entirely entrusted to statistics experts.

Thus, the considered approach to categorizing metadata of the statistical process does not cover all possible usage needs. This imposes limitations, according to which:

– The statistical process should be decomposable into atomic procedures.

– The data source should acquire a clear structured form (e. g., a relational database).

– All atomic procedures should be programmable and pre-registered in the knowledge base.

The use of a service-oriented architecture for building statistical processes allows for standardization by unifying the functions of individual services. The availability of practical implementation opportunities for software products in a service-oriented architecture using cloud environments serves as a guide for further research into the use of a service-oriented approach in building statistical processes.

4. Conclusion

An approach to the development of external statistical services as a set of software products for solving analytical tasks by groups of independent researchers is proposed. The proposed approach assumes that such a service should have its typical data storage structure with fixed types and attributes, to which the researcher imports their own data,

selects a calculation execution scheme, initializes their execution, and receives results. Taking this storage scheme into account, the main groups of metadata required to describe the implementation of a statistical process are identified. A statistical business process typically consists of a set of procedures that implement certain statistical operations and data transformations, and the provided set of metadata allows for the realization of the process goal and activity.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Data availability

The manuscript has no associated data.

References

1. *Modernization of official statistics.* UNECE. Available at: https://unece.org/statistics/modernization-official-statistics Last accessed: 30.03.2023
2. ESS Vision 2020 Shared Services. European Commission. Collaboration in Research and Methodology for Official Statistics. Available at: https://ec.europa.eu/eurostat/cros/content/ess-vision-2020-shared-services_en Last accessed: 30.03.2023
3. Lumpova, T. I., Ostapchuk, O. E. (2016). Vykorystannia statystychnoi posluhy dlia perekhodu do servis-oriientovanoi arkhitektury statystychnoho vyrobnytstva. *Statystyka Ukrainy, 2,* 6–13.
4. Pakari, S., Kheirkhah, E., Jalali, M. (2014). Web Service Discovery Methods and Techniques: A Review. *International Journal of Computer Science, Engineering and Information Technology, 4 (1),* 1–14. doi: https://doi.org/10.5121/ijcseit.2014.4101
5. Torma, S., Villstedt, J., Lehtinen, V., Oliver, I., Luukkala, V. (2008). *Semantic web services – a survey.* Helsinki University of Technology. Available at: https://www.semanticscholar.org/paper/SEMANTIC-WEB-SERVICES-%E2%80%94-A-SURVEY-Villstedt-Oliver/b10040ffe9b7e8a2d9b82ee473248935c00ec9ac
6. Sycara, K., Paolucci, M., Ankolekar, A., Srinivasan, N. (2003). Automated discovery, interaction and composition of Semantic Web services. *Journal of Web Semantics, 1 (1),* 27–46. doi: https://doi.org/10.1016/j.websem.2003.07.002
7. Sperley, E. (1999). *The Enterprise Data Warehouse, Volume I: Planning, Building and Implementation: 1 (Hewlett-Packard Professional Books).* Prentice Hall, 368.
8. Herasymenko, S. S., Holovach, A. S., Yerina, A. M. et al.; Herasymenko, S. S. (Ed.) (2000). *Statystyka.* Kyiv: KNEU, 467.
9. Petrenko, A., Bulakh, B. (2018). *Intelligent Service Discovery and Orchestration.* 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC). Kyiv, 1–5. doi: https://doi.org/10.1109/saic.2018.8516723
10. *Guidelines for the Modeling of Statistical Data and Metadata* (1995). United Nations Statistical Commission and Economic Commission for Europe. Methodological Material. Geneva, 31. Available at: https://unece.org/DAM/stats/publications/metadatamodeling.pdf Last accessed: 29.03.2023

Tetyana Lumpova, PhD, Professional Pre-Higher Education Institution «Optical-Mechanical College» of Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, ORCID: https://orcid.org/0000-0002-7576-4289

✉*Ihor Kasianchuk, Assistant, Department of System Design, Applied System Analysis Institute, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine, e-mail: kasyk3@gmail.com, ORCID: https://orcid.org/0009-0000-2215-149X*

✉*Corresponding author*