One Approach to Formation of Common Information Space Information Resource in Organizational Management Systems

Volodymyr Yuzefovych [1], Yevheniia Tsybulska [1], Oleh Andriichuk [1,2,3]

[1] *Institute for Information Recording of National Academy of Sciences of Ukraine, Mykoly Shpaka str. 2, Kyiv, 03113, Ukraine*
[2] *Taras Shevchenko National University of Kyiv, Volodymyrs'ka str., 64/13, Kyiv, 01601, Ukraine*
[3] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Peremogy ave., 37, Kyiv, 03056, Ukraine*

Abstract
This paper considers general principles of forming a Common Information Space (CIS) information resource in Organizational Management Systems (OMS) by building a hierarchy of interconnected information objects driven by the needs of OMS users. We propose a generalized scheme of a multistage information processing procedure for formation of the CIS information resource and a procedure for determining information needs of OMS users. Possible advantages and difficulties in implementing this approach are identified.

Keywords
Common Information Space, Organizational Management System, information resource, user's information needs

1. Introduction

This paper considers an issue of constructing a Common Information Space (CIS) as part of an Organizational Management System (OMS). This is a complex system for collecting, analyzing and processing information about an object under management, which has an internal environment and interaction with an external environment. These types of systems have a social (human) component, where the elements are active subsystems with their own local goals, own idea of the environment and management model [1, 6]. That is, a CIS serves as an information foundation for satisfying information needs of the OMS users.

Generally, a CIS is understood as a set of databases and data storages, technologies of their maintenance and use, telecommunication systems and networks that operate on the basis of common principles and rules and provide information interoperability to ensure the satisfaction of user's information needs [4].

Automated information resources unified into a common information space can be classified in a variety of ways. An important factor to consider is that such resources can vary in their structure, so they can be divided into several (elementary or complex) data units within a specific information resource. The common information space is engineered in accordance with the CALS concept (Continuous Acquisition and Life Cycle Support), which means that information support needs to be automated for all stages of the product lifecycle [4, 5].

Another important strategy for developing a modern common information space is to integrate it with intelligent subsystems supporting the decision-making process. In this context, the content and the structure of a common information space must satisfy an entirely new set of requirements in terms of functionality of all software components and their working procedures [4].

XXI International Scientific and Practical Conference "Information Technologies and Security" (ITS-2021), December 9, 2021, Kyiv, Ukraine
EMAIL: uzefv71@gmail.com (V. Yuzefovych); evts68@gmail.com (Y. Tsybulska); andriichuk@ipri.kiev.ua (O. Andriichuk)
ORCID: 0000-0002-6336-9548 (V. Yuzefovych); 0000-0003-3342-4507 (Y. Tsybulska); 0000-0003-2569-2026 (O. Andriichuk)

Thereby Common Information Space includes: information resources contained data (knowledge); organizational structures that ensure CIS operation and development; means of information interaction to ensure access to information resources (IR). The last two components form the CIS information infrastructure.

This infrastructure contains hardware and software that together must ensure reliable storage and quick access to information (in accordance with the user's authority) and data exchange. However, in order to meet the user's information needs, a CIS must provide, from the user's perspective, a high quality information resource (the first and, in our opinion, the key SIC component). This IR consists of a set of related information objects (IO) which can be "atomic" (indivisible) or composite. In this context, information object is an integral unit, which has a certain state, properties and functional purpose. [5].

2. Related works

Recently, there had been numerous works devoted to the issue of constructing a Common Information Space, including [2-7]. Most publications focus on creation and improvement of CIS information infrastructure. For example, Common Information Space proposed in [2] is a concept for a distributed database with an efficient, secure and flexible replication mechanism that features a context-based filtering mechanism (depending on the role, mission, and device) and takes into account bandwidth restrictions and disruptions of tactical networks. Also the Common Information Space acts as a generic mediation service, converting between different interoperability standards.

The authors of the paper [3] present feasibility of a Common Information Space as a way to interlink the existing data sources for shared access in management of complex production processes. Their analysis of of available architectural solutions demonstrates consistent of the service-oriented approach, especially when combined with a web-oriented data access system.

[4] discusses problems of user information interaction in a CIS of a specialized subject area, and offers tools for visualization of users request results. The authors define Common Information Space as an infrastructure for high-performance and high-throughput processing of sensor data streams and an environment for end-users to run applications and manage their results. In general however, problems of intelligent data processing in the CIS construction remain out of consideration in the above works.

In contrast, the paper [5] discusses how intelligent data processing techniques can be used for automated discovery of information system assets. The authors define concepts such as information objects, system assets, system events and their characteristics, and propose a technique that incorporates determining the types of object characteristics, the types of objects (system assets), and their hierarchy based on the statistical analysis of system events.

Also noteworthy is the paper [7], which proposes a way to determine the information needs of users, based on information analysis adopted from the Activity Theory. The author focuses on the development of effective and efficient mechanism for sharing information within and between organizations and proposes the special information sharing framework that takes proper account of the user's needs and their activities.

3. Background and problem formulation

It is obvious that any of the known OMS is a hierarchical (multilevel) system. Consideration of users information needs at different hierarchical levels of OMS shows that they differ significantly. This is due to the difference between functional tasks and responsibilities assigned to users. At the highest hierarchical levels, the tasks of the strategic level (within the system under consideration) are mainly solved. Accordingly, users of this level of IP deal mostly with generalized information objects.

As for the users of the lower hierarchical levels of an OMS, that generally solve tactical level problems, their information needs are mainly focused on data about objects that directly characterize current processes and events. These are objects of objective reality, which mostly (but not always) belong to the material world. Fundamental difference between such IO and the IO for users of the strategic level is the availability for direct observation (measurement, evaluation) of their main properties. Therefore, one of the components of the OMS organizationally, or at least functionally, is a subsystem for monitoring of external environment and internal system parameters.

The main tasks of the monitoring subsystem are following.

- Formation of a set of indicators that characterize the OMS state, the external environment state and are necessary to create user's information objects ("atomic" or composite);
- Collection (measurement) of indicator values and assessment of their quality (the notion of data quality is discussed below);
- Collected data pre-processing to improve its quality;
- Passing prepared data to an information-analytical subsystem (IAS) of the OMS for further processing and user's information objects formation.

On the other hand, even at a single hierarchical level of a management system, in addition to a staff that performs OMS direct functional tasks, there is a staff performing support and/or service tasks – personnel, financial, logistics and others. Such users (at a single hierarchical level) of the OMS information resource operate mostly on common information objects, but view them from significantly different angles. Thus, their information needs are met by data on common objects, but these data characterize completely different properties of information objects. So, it can be concluded that data on the same IOs are not able to effectively satisfy the information needs of all users of the hierarchical OMS. As a rule, data on IO, which are operated by users of higher hierarchical levels of the management system, are not subject to direct observation (measurement, experimental evaluation). That is why they should be formed in a certain way by means of analytical processing of simpler IO's data.

The purpose of this study is to develop a comprehensive (holistic) approach to the formation of the information resource of the OMS Common Information Space, focused on quality provide the information needs of hierarchical system users.

4. General principles of common information resource creation

It is obvious that complex and generalized IOs are initially formed on the basis of information received from a monitoring subsystem, via multi-stage processing in some information-analytical subsystem of OMS. Generalized higher-level objects are mostly formed from IOs of previous levels. Thus, non-stop step-by-step information processing with the creation of higher hierarchical level objects based on data on simpler IO ensures unity of the information space in terms of the procedure of forming an information resource in general. This unity is achieved by forming all information objects from the very beginning on the basis of common data from the monitoring subsystem, so, all decisions made by the OMS staff at different hierarchical levels have a common information basis.

Figure 1: The scheme of CIS information resource formationshows a generalized scheme of CIS information resource formation.

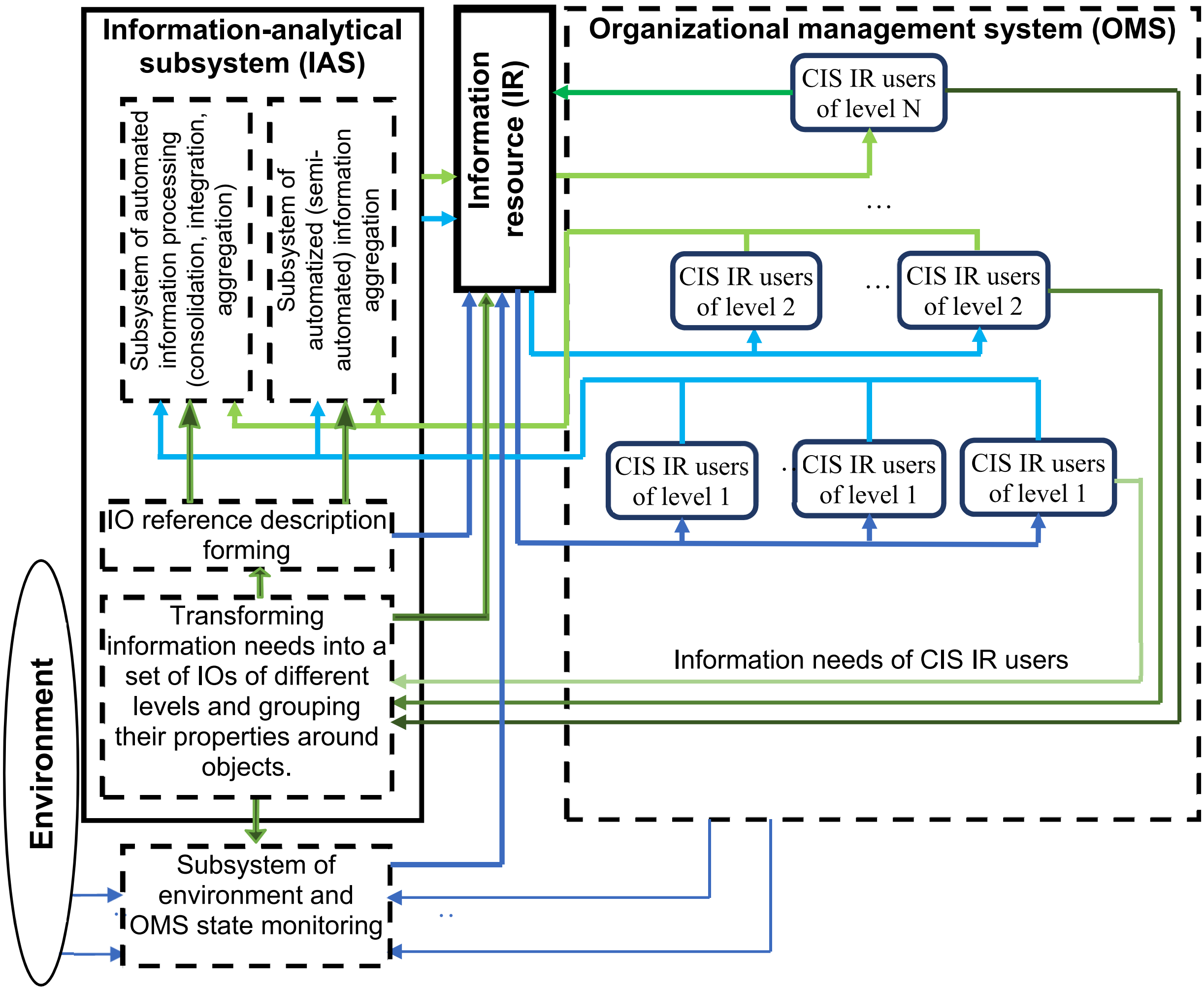

Figure 1: The scheme of CIS information resource formation

That is, this scheme provides to obtain IO of different levels of generalization by performing the following operations.

- Allocation of the components in the OMS information-analytical subsystem;
- Transformation of information needs into a set of IOs;
- Formation of object reference descriptions using their properties and requirements to them;
- Performing automatic or, if not possible, automatized (semi-automated) information processing.

Figure 2: A procedure for determining information needs of OMS usersshows a procedure for determining information needs of OMS users.

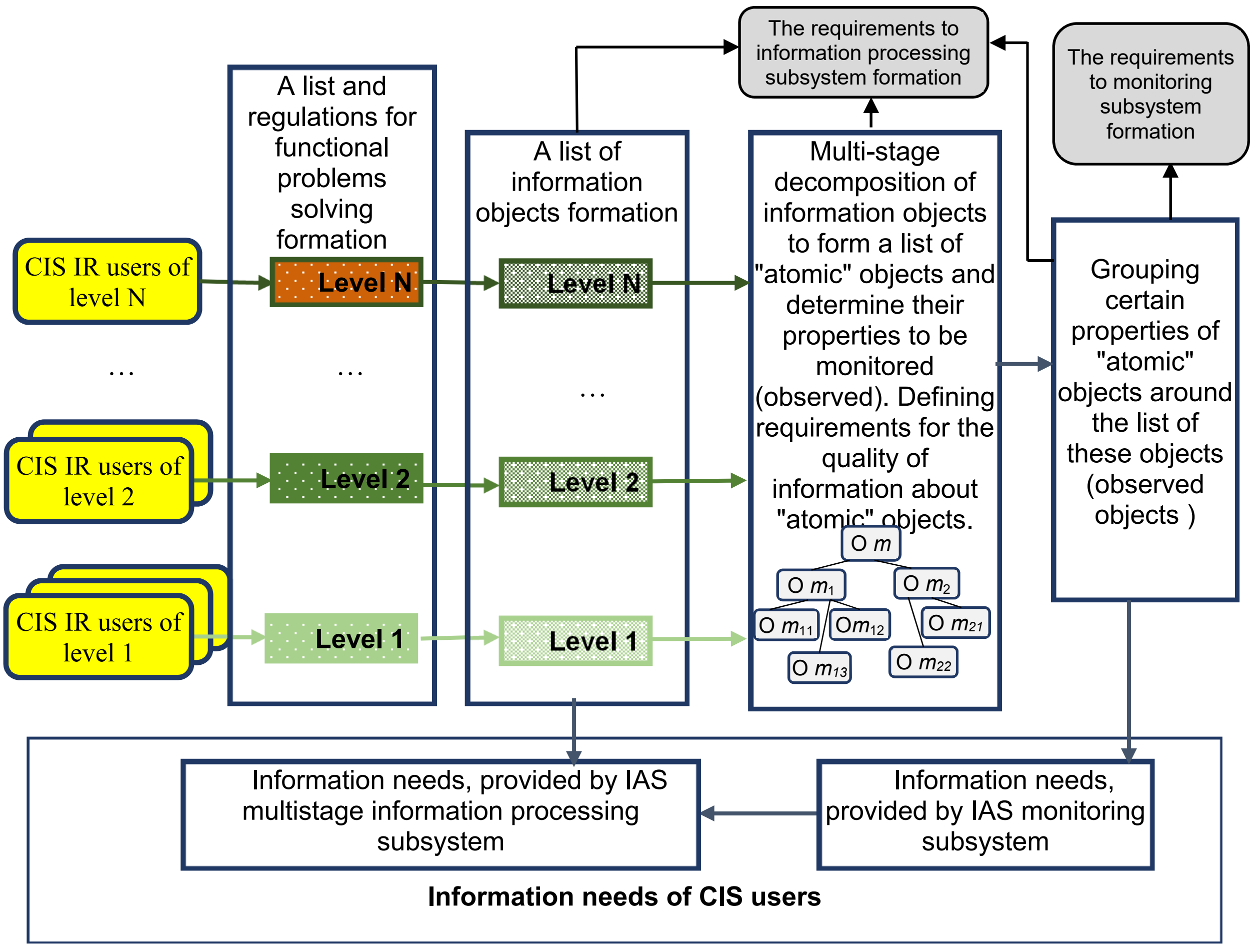

Figure 2: A procedure for determining information needs of OMS users

The first stage of determining information needs is to form a list of all functional tasks to be solved by OMS staff. It is also important to take into account the regulations for solving functional tasks at different hierarchical levels of the system.

It is obvious that tasks at the strategic management level are solved with a longer time horizon (in a longer cycle) than tasks of the tactical level. This is due to less dynamics of the strategic environment and more time required to implement strategic decisions. Therefore, higher-level IO data can be updated at a slower pace, which can reduce the burden on the monitoring subsystem.

Then, on the basis of functional tasks list, the set of all IO, which are used by users when solving these tasks, is determined. The resulting list of IO is subject to multi-stage decomposition with the formation of an exhaustive list of "atomic" objects, which in their various combinations allow to form any compiled IO that makes up the information needs. "Atomic" IO means information objects whose properties can be measured and evaluated by the monitoring subsystem tools. That is, the list of "atomic" IO is the information needs of OMS users, which must be provided by the OMS monitoring subsystem. Other IO or their individual properties should be formed by OMS IAS using multi-stage processing of formed before IO data.

Given that (as already mentioned) different users are interested in different properties of the same objects, the last step in determining information needs is to group several sets of properties of each "atomic" object to provide them to corresponding users. This grouping of properties will simplify the tasks executing by the OMS monitoring subsystem.

5. Data quality estimation and information transformation

An important task is to determine the requirements for the quality of information about "atomic" IO, which will essentially limit quality of the final information provided to the user.

Such measures of information quality include the following.

- Relevance - correspondence of data to the current moment (a period of information relevance is proportional to the speed of the processes that this information reflects (characterizes));
- Accuracy - proximity of the values of information units to the real value (determination error);
- Reliability - degree of objective reality reflection;
- Timeliness (as opposed to efficiency) - the formation of all the necessary information to users directly on demand;
- Completeness - compliance of the composition and volume of available information (taking into account its accuracy and reliability) to the information needs of users.

Estimation of the accuracy of data obtained by technical tools, as a rule, is not a problem and is equal to the sensor accuracy. This information can be considered reliable. Accuracy and reliability indicators of the information obtained from other sources are established experimentally. It is important that all input information contains data on initial accuracy and reliability. This makes it possible to improve (justifiably increase or decrease) its quality indicators in the process of joint processing of all incoming information.

Taking into account the regulations for performing the tasks of the OMS staff allows using the information timeliness measure instead of the relevance indicator. Let's consider this issue in more detail. Obviously, the relevance of information is provided by the promptness of its acquisition, and before that – by the speed of data extraction, collection and transmission. This ultimately ensures the ability of OMS IR to reflect the real dynamics of events. At first glance, the high speed of data acquisition is an essential requirement for solving the problem of information collection by the OMS monitoring subsystem, because it ensures availability of the most up-to-date information at any time. However, the requirement of high speed of information collection (together with other requirements) will require significant resources for data acquisition (extraction), refinement, processing, storage, transfer and analysis. As a result, these resources can make up the specific weight of all system resources allocated to solving the OMS functional tasks.

Taking into account the above, it seems more rational to abandon the requirement to ensure high speed of obtaining information in favor of the requirement to provide management units with all the necessary information for decision-making in a timely manner. That is, information about the system indicators or external environment, which maximally meets the requirements of relevance, completeness, accuracy and reliability, should be available not at any moment of time, but precisely when it is needed to make decisions. We will analyze the possibility of obtaining practical advantages by applying this "ideological" approach.

By definition, the OMS central (target) task is making management decisions. Each step in the management cycle is invoked sequentially and requires time for its execution. The pace of making decisions together with time required to send them to units for implementation and receive a response should not exceed the speed of their execution (implementation "into life"), except for cases when the current decision is aimed at the change or cancellation of the previous one. That is, the time required to implement decisions objectively limits the expedient pace (interval) of their development. As is known, a characteristic feature of a large number of complex systems is the need for significant amounts of time to implement management decisions, especially at the strategic level. In addition, the results may not be visible immediately.

Figure 3 shows a hierarchical diagram that presents the passage and transformation (aggregation) of data about the system and its external environment for providing it to different levels of management units. It is obvious that as the data gets generalized more, it becomes less dynamic. That is, it can be argued that more generalized data (IO) actually characterize slower processes that are important for the system. This also explains the fact that the need to make strategic (systemic) decisions regarding the OMS operation, which are based on generalized IO's data, occurs less frequently than the need to make tactical (current) decisions.

So, for each specific OMS, it is possible to determine a pace (periodicity) of updating data on the IOs of the CIS information resource, depending on their generalization level. In fact this pace will correspond to a pace and work cycles of the relevant OMS management units. In most cases, the cadence of obtaining indicator values can be determined in advance. This approach to solving the problem of data collection regarding the OMS IR information objects with a predetermined pace of updating indicator values fully corresponds to the system "regular" operation.

For example, in special-purpose systems (armed forces, emergency services or civil defense, etc.), the regulation of updating indicators is established by a special document - a report card [8, 9], which defines the list of necessary data (indicators) for various system components and the intervals for updating their values. In conditions of destructive (especially organized) external influence on the management system, it is advisable for each of its levels to define a list of data that require a faster OMS response and should be determined at an increased rate. Change in the data determination and update mode can be introduced by setting a certain mode of the system operation as a whole, for example – everyday (regular) operation mode, intensive (emergency) operation mode. In military systems, where the effective functioning is supercritical for any nation, various modes of combat readiness (for example, "constant", "increased", "full") are used for this purpose.

Also, in the general case, the control system can use so-called signal indicators, which recognize the appearance of some new negative factors and can characterize their intensity. Such indicators do not have a determination period. Therefore, for their timely processing, the OMS level to which their values should be transferred, the degree of importance (priority) and the time that can be spent on their transfer to the IR for further analysis and decision-making must be determined. The time that must be spent on obtaining the indicator value (measurement, pre-processing, transmission, etc.) must be deducted from the start time of its use in the next OMS control cycle to obtain the start time of its determination by the monitoring subsystem.

Therefore, when solving the task of collecting data on system indicators, it is necessary to take into account that such information must be available exactly when it is needed to make a management decision, in the necessary volumes, in the most complete form, and be up-to-date. Too low of a rate of updating data will lead to its obsolescence and, accordingly, to a decrease in the efficiency (timeliness) of the decisions made. Too high update rate will lead to unnecessary (non-productive) loading of the OMS monitoring subsystem, data transmission channels, means of their processing and storage, etc. The latter is fundamental from the point of view of forming the CIS information resource.

Thus, to ensure data relevance, it is advisable to determine in advance the time intervals for updating data about determined information objects. The list of such data and the intervals of their receipt may change depending on the conditions of system operation. The setting of a certain OMS operation mode can be the signal for these changes.

If an OMS is unable to process all indicators in real time, or if a person is involved in their analysis, it is advisable to determine the priority of data processing in advance by providing (assigning) an indication of IO importance and store such data to a CIS information resource.

Also the completeness indicator can be assessed by comparing the list and amount of data about IO, which determine the information needs of OMS users, and the amount of information available in the information resource, taking into account its quality.

As already mentioned above, Figure 3: The stages of information processing at the OMS IASshows a procedure of multi-stage information processing at the OMS information-analytical subsystem in order to form a CIS information resource. At the first stage, the information received from the monitoring subsystem is combined and integrated with the information already available in the system. Thus, we get a description of the basic characteristics of user’s information objects. At the next stage, information is grouped in accordance with the information needs of system users. Further, there are gradually formed generalized information objects, which correspond to the tasks performed by users of various levels. These objects are created on the basis of individual properties, “atomic” objects and generalized objects of a lower level.

Based on the analysis of Figure 3, providing of this procedure is possible through integrated application of the following.

- Methods and ways of pre-processing information (assessment of "input" data quality; formalization of heterogeneous data from different sources);

- Methods and ways of identifying objects in the external environment and consolidating their property measures around these objects;
- Methods of combining heterogeneous information (numerical and linguistic) with different accuracy and reliability to improve its quality;
- Methods of IO parameters smoothing by a numerical series in time obtained from one or more sources taking into account their errors and opportunities to ensure the quality of information;
- Methods for predicting IO parameters and IO state in general (in particular - the use of time series forecasting models);
- Methods of IO grouping (for example, clustering methods) for further recognition of complex (generalized) information objects;
- Methods of IO data aggregation with the formation of complex (generalized) information objects;
- Recognition methods for information objects of different levels.

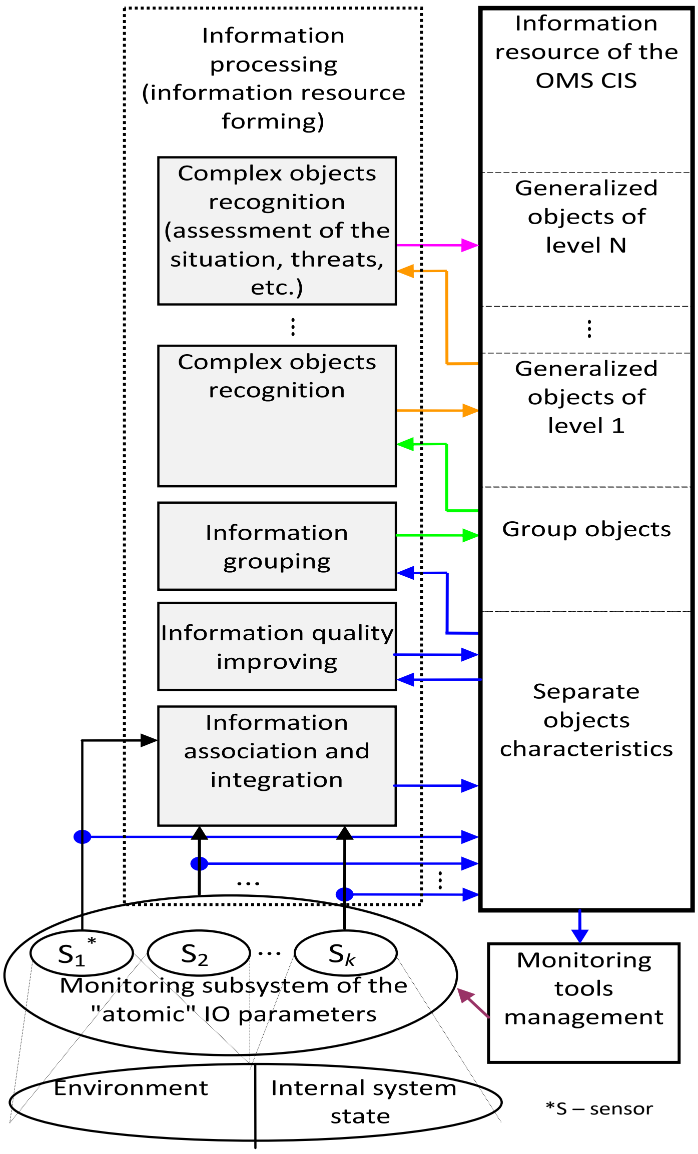

Figure 3: The stages of information processing at the OMS IAS

It should be noted that methods of information pre-processing are closely related to the methods of its further processing. For example, when using fuzzy logic or fuzzy integrals to process information, input should be presented and stored as fuzzy sets.

Returning to Figure 1**:** The scheme of CIS information resource formation, we note that when choosing specific methods of information processing, preference should be given to those that provide automatic problem solving. For example, the paper [10] considers existed and proposed new method of automatic clustering, which allows you to solve the problem of moving objects grouping without the participation of the operator. The method proposed in [10] is characterized by the fact that, under certain conditions, it does not require the construction of complete clustering tree.

Also, a new linguistic method for forecasting non-stationary dynamic time series is proposed in [11]. It has the minimum number of input parameters needed to "run" a prediction model. At the same time, these parameters have a "transparent" meaning for users, have a simple setting and can often be set in advance without a detailed "manual" analysis of a specific time series.

6. Proposed method characteristics

In general, the proposed approach to the formation of the OMS CIS information resource, as a hierarchy of interconnected information objects of different level generalization for different level SOU users, has the following advantages.

1. Target orientation to identifying and meeting the information needs of different level users , which corresponds to the main goal of the CIS.
2. High automation level of an OMS CIS information resource formation process.
3. Ability to accumulate information "around" the information objects that make up the users information needs, to provide quick access to data on the current state of such objects. Ability to use your own (specific to this group of users) searching terminology.
4. Absence or minimum of redundant information received at the OMS CIS user request.
5. Ability to quickly target detailed analysis of information objects state due to direct (searchless) access to the objects components that determine it.
6. Lower dynamics of changes in high-level information objects allows EEA users to continue solving tasks for some time, in the absence (partial loss) of information about "atomic" objects or objects of lower hierarchical level. That is, formed in this way CIS information resource is characterized by increased indicators of information viability.
7. Ability to recover information about lower-level information objects, in case the procedure of obtaining higher level data is reversible.

Disadvantages of the approach include the following.

1. Significant increase of the amount of data to be accumulated and stored in the OMS CIS.
2. High analytical complexity and ambiguity of the multi-stage procedure of CIS information resource formation.
3. The need to provide additional specialists with high qualification requirements in the OMS organizational and staffing structure for the formation and maintenance of information resources.

7. Conclusions

The paper proposes a comprehensive approach to the formation of the OMS common information space information resources as a hierarchy of interconnected information objects that are determined by the information needs of OMS users and together are able to provide them. This approach makes the process of filling the information resource object-centric, which is convenient for solving the OMS monitoring subsystem tasks. On the other hand, it is client-oriented, where the user of its information space acts as a OMS "client".

Also this paper presents a general scheme for CIS information resource formation. It provides for the allocation of components in the OMS information-analytical subsystem, transforms information needs into a set of IOs, forms a reference description of such objects with use their properties and requirements to them, performs automatic or, if not possible, automatized (semi-automated) information processing. This approach allows to ensure the unity of the information space from the point of view of the process of forming its information resource, and not only the information infrastructure.

A procedure for determining information needs of OMS users is developed. This procedure consists in the analysis of the functional tasks performed by the staff, the selection of information objects necessary for their solution and the formation of the appropriate sets of hierarchically related information objects and their properties that are of informational interest for each user.

The stages of information processing for the information resource of OMS CIS creation are determined. They include integration and aggregation of information, assessment and improvement of the information quality, creation of "atomic", group, and, on their basis, generalized information objects of different generalization levels. These information objects and their environment constitute a Common Information Space designed to fully and effectively meet the information needs of OMS users.

A list of information processing methods necessary to ensure maximal automation for the process of OMS CIS information resource creation in accordance with the proposed approach, is given.

Since any system characterized by functional specialization of elements can be considered as hierarchical, proposed approach can be extended to other socio-technical systems.

Acknowledgement

This work was supported by the ECHO project which has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement no 830943.

References

[1] O. Dodonov, V. Putyatin, S. Kutsenko, D. Lande. Construction of the generalized structure of information organization management system (in Ukrainian), Mathematical Machines and Systems 3 (2017): 3-22, URL: http://dspace.nbuv.gov.ua/handle/123456789/125607.

[2] Fabian Angelstorf, Stefan Apelt, Nico Bau, Norman Jansen, Sylvia Käthner. Shared Information Space, in (2017) International Conference on Military Communications and Information Systems (ICMCIS) 15-16 May 2017. doi:10.1109/ICMCIS.2017.7956489.

[3] Pavel V. Senchenko, Yury B. Gritsenko, Oleg I. Zhukovsky, R.V. Mesheryakov. Architectural Principles of Common Information Space Development for Control of Complex Production Processes. URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8687007.

[4] B. Baliś, T. Bartyński, M. Bubak, G. Dyk, T. Gubała, M. Kasztelnik. The Common Information Space: A Framework for Early Warning Systems, in Cracow Grid Workshop (2012), Kraków, Poland, 22 October 2012. URL: http://dice.cyfronet.pl/products/cis.

[5] A. Fedorchenko, E. Doynikova, I. Kotenko. Towards Intelligent Data Processing for Automated Determination of Information System Assets, in Handbook of Research on Intelligent Data Processing and Information Security Systems, 2019, pp. 147-160. doi:10.4018/978-1-7998-1290-6.ch007. URL: https://www.igi-global.com/chapter/towards-intelligent-data-processing-for-automated-determination-of-information-system-assets/243040.

[6] O.G. Dodonov, O.S. Gorbachyk, M.G. Kuznietsova. Dynamic Reconfiguration in Automated Organizational Management Systems, in Selected Papers of the XX International Scientific and Practical Conference "Information Technologies and Security" (ITS 2020), Kyiv, Ukraine, December 10, 2020, pp. 129-141. URL: http://ceur-ws.org/Vol-2859/paper11.pdf.

[7] A.A. Alhefeiti. A Systematic Method of Developing Information Sharing Systems Based on Activity Theory. Ph.D. thesis, Informatics Research Centre, Henley Business School, University of Reading, GB, 2018. DOI: 10.48683/1926.00082873.

[8] V. Afanasiev, A. Raychenko. Regularity of cyclicity in the principles of management. (in Russian), URL: https://cyberleninka.ru/article/n/zakonomernost-tsiklichnosti-v-printsipah-upravleniya.

[9] M. Marusina, A. Tihanovskiy, V. Tkalich, O. Ushakov, A. Chernyaev. Metrology, certification and standardization. Electronic textbook. (in Russian), URL: http://de.ifmo.ru/bk_netra /page.php?tutindex=1.

[10] V. Yuzefovych. Solving the problem of moving objects grouping using clustering methods in the monitoring systems (in Ukrainian), Data Recording, Storage & Processing 17 (4) (2015): 31-37.

[11] D. Lande, V. Yuzefovych. A Linguistic Approach to Time Series Forecasting (in Ukrainian), Data Recording, Storage & Processing 24 (1) (2022): 13-22.