

UDC: 332

DOI: 10.5281/zenodo.14035323

DIGITAL MODELS AND METHODS OF SYSTEMATIZATION OF INFORMATION IN SOCIO-ECONOMIC SYSTEMS

Kateryna Naumik-Gladka

Doctor of Sciences (Economics), Professor, Honorary Affiliated Scientific Fellow
of Fil Dr. Jan.U. Sandal Institute, Norway

ORCID ID: 0000-0003-0492-7631

Oleksandr Ihnatiev

Doctor of Philosophy in Public Management and Administration
Senior Lecturer of the Department of Military Training,
National University of Civil Defence of Ukraine, Kharkiv, Ukraine

ORCID ID: 0000-0003-3280-3468

Abstract. The authors analyze the existing models of information systematization in socio-economic systems and obtain several gradations: systems of systematization of financial information; depending on the architecture of information storage used in the system; depending on the model of combining information; depending on the method of processing of user requests; customer information systematization systems; systems for systematization of information about the external environment of the company; depending on the method of information analysis. The authors study the attitude of the HR specialists to the performance management and the measures of companies concerning improving of the performance management.

The authors propose to make information systematization systems in accordance with the CRM approach to solve the following problems: saving potential customers; control of employees and standardization of work with customers;

accumulation of statistical base, which is also very important for the successful development of any business; obtaining final solutions that can serve as a basis for building your own work system.

Keywords: socio-economic systems; model of information systematization; database; performance management.

Introduction. The systematization of data consists in combining many computers and servers into a single environment designed to solve certain types of problems, for example, scientific problems or complex calculations. Subsequently, a plurality of data distributed to computing nodes and repositories is accumulated in such a structure. Typically, applications running in a distributed computing environment access only one of the data sources. But, if there is a need for simultaneous access to several sources, there are difficulties, since these sources can contain heterogeneous data and methods of accessing them, as well as be located at a distance from each other. In addition, it is convenient for users who analyze the accumulated data to access a single source of information, forming requests and obtaining results in the same format. Thus, the main problem of the approach to storing information in distributed computing systems is the heterogeneity and remoteness of data sources. The solution of this problem is to create a centralized access point that provides a single interface for accessing all data sources of the computing cloud in real time. In this context, it is necessary to choose the most acceptable approach and the appropriate platform to ensure such consolidation.

Literature review. In the article (Simion, Eduard, 2022), the authors systematize knowledge from several cases of using of cybersecurity. The review begins from referring to the discovery of cases of military use. The review continues with several well-known and publicly disclosed cybersecurity methods. In particular, the authors consider classic options for protecting against distributed denial of service attacks and corresponding anonymization schemes.

In the article (Kitamura, Y., & Mizoguchi, R., 2004), the authors for the first time consider the ontologies that guide the conceptualization of artifacts from a

functional point of view. The basis for the systematization of knowledge is based on an extended ontology of the device and a functional ontology built on top of the ontology of the extended device. The authors also discuss the use of systematic functional knowledge in several application systems along with its advantages.

The authors (Michele Fabi, Myriam Kassoul & Julien Prat, 2024) provide an overview of the academic literature on automated market makers for decentralized exchanges. The authors consider the optimal design of automated market makers. The authors also discuss models that use diverse ideas in bilateral markets to characterize the equilibrium size of liquidity pools and the incentives of liquidity providers. Finally, the authors review recent research on the interaction between Miner Extractible Value and decentralized exchanges.

The study (Silva, E.C. da, Barbosa, F.C. & Carrino, A.L. 2023) examines the problem of systematization of the intellectual market. This work is aimed at demonstrating the implementation and systematization of an intelligent market focused on convenience. From the analysis of the collected data, it was possible to identify the main components of the autonomous market, as well as the challenges and opportunities for its implementation.

It should be noted that in the context of systematization of knowledge within socio-economic systems, such studies require further development. Accordingly, **the purpose of the article is** to observe digital models and methods of systematization of information in socio-economic systems in the modern conditions.

Methodology. In the context of analyzing models of information systematization systems, depending on the architecture of information storage used in the system, it is possible to distinguish such methods of their construction.

1. Transferring structures – this is moving existing analytical structures to one platform. This approach is attractive for companies that need to minimize costs. Such an operation allows you to reduce the number of servers used, as well as unload specialists who are engaged in their maintenance. It should be noted that transferring structures does not involve changing applications or integrating data. Data model,

metadata, transformation logic and reporting technology remain unchanged, only client interfaces are adjusted.

2. The project "from scratch" – the creation of a system of systematization of information in the company without the use of existing systems or other developments.

3. Define and transition to a corporate standard – install existing data storages or storefronts as a corporate standard, and translate other structures immediately or gradually to that standard. This approach can be recommended in cases where there is a takeover of one company by another. Accordingly, the master company data store becomes the corporate standard, and the absorbed company data store is built into it.

4. Synchronization. This approach is more acceptable if the company needs to standardize reference data for master data objects, such as customers, products, suppliers, etc., among a large number of operating programs. Quite often, data for these objects is collected and stored in different operating programs in different formats. As a result, numerous copies of identical records appear, which violates the consistency and standardization of the data.

5. Consistent data storefronts are another way to consolidate data storefronts without having to physically combine them. This is the introduction of measurements into the structure of each showcase so that they are coordinated with each other. This approach is based on the Ralph Kimball methodology, according to which companies create one temporary intermediate area from which data is used to fill the agreed data storefronts. The introduction of this new staging area consolidates redundant data, i.e. reduces costs.

6. Display case from data storefronts. If the company is decentralized, and only some corporate users need consolidated information, it is advisable to create a showcase of data from all existing analytical structures. According to this approach, the necessary corporate information is obtained from existing structures using standard ETL tools. The advantage of this approach is that you do not need to change existing environments. In addition, there is no need to increase the staff of specialists and purchase software tools, which can save significant money.

7. Distributed queries. Using this method, existing non-integrated analytical structures also remain in place. However, instead of the ETL and batch processing tools used to create a consolidated view, in this case, such a view is constructed using SQL (Structured Query language) tools.

In this case, it is necessary to further analyze the options for moving existing analytical structures to one platform. In this context, it is advisable to review information systematization systems depending on the information pooling model. Accordingly, the ability to distinguish several options for building such systems must take into account:

- 1) consolidation of storage systems located on several servers within a single server;
- 2) direct connection of several heterogeneous servers to a single storage system;
- 3) consolidation of storage based on the SAN architecture (Storage Area Network).

In particular, as for combining storage systems located on several servers, within a single server, this option of building information systematization systems reduces the severity of administration problems and saves the space that the equipment occupies by centralizing the storage system on one powerful server.

Main part. Distributed systems, and databases in particular, tend to offer a specific consistency model. These models ensure that if certain conditions are met, then the system can be expected to have certain properties (i.e. consistency or availability or network break resistance). In this context, a popular set of ACID properties is considered, which ensures that database transactions are processed reliably, compared to the opposite BASE model, which is derived from the CAP theorem, but seeks to provide a set of properties different from ACID [2; 12].

1. ACID concept. Traditionally, relational databases are able to provide and adhere to the ACID property, which is a set of properties that ensure that database transactions are processed reliably and that the database remains consistent under

concurrent access and system failure conditions. ACID is an abbreviation that stands for the following: atomicity, integrity, isolation, and durability.

Atomicity. Atomicity means that database transactions must adhere to the all-or-nothing principle.

Given that a transaction is a series of instructions to execute, for a transaction to be "atomic," all instructions must be executed, or if one (or more) instruction cannot be executed, the entire transaction fails and the database must remain unchanged.

Integrity. Integrity ensures that only valid data is written to the database, ensuring that if the transaction is successful, the database goes from one agreed state to another agreed state. If, for any reason, an error occurs during the transaction, any changes already made will be automatically reversed (rollback of changes), so that the database remains in a consistent state.

Isolation. Isolation requires transactions in progress and not yet recording their changes to remain isolated from any other transaction. Therefore, the execution of a transaction should not affect the execution of other parallel transactions. Isolation is important because while the transactions are in progress, the system state cannot be consistent, as the transactions only ensure that the system is in a consistent state after the transaction is complete. If the transaction was not started in isolation, it will be able to access data from the system that is not consistent at this point.

Durability. Durability ensures that any transaction in the database is permanent and will not be lost. The database should be able to restore confirmed updates made by the transaction in any form of system failure (hardware or software). Many databases implement durability by writing transactions in the transaction log so that these transactions can be executed again to replicate the state of the system that was just before the failure. The transaction is considered complete only after it is logged [6; 8].

The ACID model means the advantage of distributed database consistency, it is pessimistic and provides consistency at the end of each transaction. This strict consistency may not always be necessary depending on the requirements of the

subject area. In addition, this model is very difficult to implement in a distributed system without creating bottlenecks. For example, a two-phase commit protocol that ensures the atomicity of distributed transactions requires blocking the entire node while the node waits for notification. This model also creates a lot of overhead, as it involves sending a large number of messages. In addition, if we take into account the CAP theorem, ensuring strict consistency can only be ensured to the detriment of availability or resistance to network breaks. If we use a two-phase commit protocol, we are able to ensure node consistency, resulting in a loss in availability. Since different modules may have different consistency requirements, NoSQL databases tend to provide alternative consistency models [1; 2].

2. BASE concept. BASE, as its name suggests, is the logical opposite of ACID (a play on the English words base and acid, which denote chemical terms). While ACID focuses more on consistency, and in this sense is pessimistic, BASE focuses on affordability, providing consistency ultimately, and in this sense is optimistic. BASE is an abbreviation for a set of properties: Basic Available, Soft-state, Event consistency. This is inextricably related to the CAP theorem and was also proposed by Eric Brewer in 2000 at a symposium on the principles of distributed computing.

If you analyze the existing models of information systematization systems, it is possible to obtain several gradations:

- 1) systems of systematization of financial information;
- 2) depending on the architecture of information storage used in the system;
- 3) depending on the model of combining information;
- 4) depending on the method of processing user requests;
- 5) customer information systematization systems;
- 6) systems for systematization of information about the external environment of the company;
- 7) depending on the method of information analysis [7; 12].

In particular, if you consider the systematization of financial information in more detail, it is advisable to note that today they are presented on the market in the

form of package programs, company resource planning systems and standard ETL tools (Extraction, Transformation, Loading).

In turn, as for batch programs, they occupy an overwhelming market share and can be presented as analytical systems of the CPM class (Corporate Performance Management) and specialized financial systematization systems.

In particular, analytical systems of the CPM class are quite common. These are – systems such as ORACLE Hyperion, SAP BPC, IBM Cognos TM1. They have comparable functionality, and they have similar implementation processes. The choice between them is determined by the need to integrate with the information systematization systems already in the company, commercial expediency and personal preferences. Specialized financial systematization systems are less common. Among the most well-known in the domestic market, specific financial systematization systems should be distinguished such as Oracle Hyperion Financial Management (HFM) and IBM Cognos Controller. So, Oracle Hyperion has relatively more flexibility in settings (for example, in terms of systematization algorithms and basic consolidation corrections), which makes it more adaptive to a specific customer. However, this can lead to a longer and costlier implementation. The second system, IBM Cognos Controller, is less flexible and adaptive with respect to a given systematization algorithm and parametric settings, but also has such an advantage as a tuning speed based on the limits specified by the system. The lower prevalence of specialized financial systematization systems is largely due to the need to attract specialists of a certain profile to create them. Accordingly, such systems are narrowly specialized and have a longer payback period. However, the benefits of specialized financial systematization systems for the user are faster customization and, accordingly, reduced implementation costs [8; 12].

As for standard ETL tools, they are a type of key data warehouse management processes. Standard ETL tools provide the following features:

- 1) obtaining information from sources of external origin;
- 2) transformation of information and its adaptation to the needs of the business model;

3) uploads information to the data store.

As for standard ETL tools, they are a type of key data warehouse management processes. Standard ETL tools provide the following features:

1) obtaining information from sources of external origin;

2) transformation of information and its adaptation to the needs of the business model;

3) uploads information to the data store [2; 7].

Traditionally, ERP systems are built in accordance with the modular principle and allow you to cover all the key processes of the company's activities.

Among the common tasks of any information systematization system, the following ones should be distinguished:

1) selection of information sources;

2) development of a strategy for systematization of information;

3) assessment of information quality;

4) enrichment of information;

5) cleaning of information;

6) transfer information to the data store.

At the same time, it should be borne in mind that the work of any information systematization system in accordance with the stages of formation and use of consolidated statements is based on the following tasks:

1) collection and structuring of initial information;

2) the process of systematization of information;

3) use of consolidated statements [6; 7].

However, depending on the purpose of information systematization systems, their tasks can differ significantly.

So, if to study the tasks of the most relevant systems for systematizing information about customers at the present stage, the following are distinguished among their specific tasks:

1) collection of information from a variety of sources;

2) restructuring of customer information and bringing the structure to a single form;

3) search and systematization of similar records;

4) preparation of data storefronts and calculation of indicators.

In particular, information systematization systems built in accordance with the CRM approach allow you to solve such problems.

1. Saving potential customers. In small and medium-sized businesses, competition is very high. Companies are making significant efforts to attract customers. Compared to other costs of attracting customers, a significant budget is allocated. And it is very important that all these financial resources and efforts lead to the expected result.

2. Control of employees and standardization of work with customers. Without a common standardized CRM system, employee contacts are quite chaotic. The CRM system allows you to place information about all incoming and outgoing contacts in one repository, from where you can get it at any time.

3. Accumulation of statistical base, which is also very important for the successful development of any business. Thanks to the use of the CRM system, all working information is collected in one common database in a standardized form. It becomes possible to analyze work statistics, make heterogeneous reports (many of which are already in the form of CRM systems), that is, analyze work and plan further work more consciously.

4. Obtaining final solutions that can serve as a basis for building your own work system. Each CRM system has many ready-made tools that allow you to take work to a qualitatively new level. For example, the integration of the CRM system with telephony allows you to record all calls, save new contacts and analyze the quality of sales with customer requests. In small and medium-sized businesses most often work with customers is organized directly by the business owner. For the most part, he does not have experts and developments on the organization of work with clients. The implementation of the CRM system allows you to get an appropriate tool for making appropriate decisions regarding the work of the sales department [2; 6].

Conclusions. Thus, by solving the above problems, CRM systems enable:

- 1) obtaining a common company standardized base of contacts (customers, counterparties);
- 2) effective quality control of the sales department at any time;
- 3) obtaining statistics and analytics of efficiency of working with customer requests;
- 4) planning of improvement of work quality and development of business development strategy.

As for the tasks of larger systems of systematization of information about the external environment of the company, they are focused on solving the following problems:

- 1) analysis of consumer behavior;
- 2) evaluation of competitive actions;
- 3 research of market trends;
- 4) support of decision-making on selection of rational strategy and tactics of company activity;
- 5) development and implementation of the company's marketing complex;
- 6) monitoring the efficiency of the company.

The direct connection of several heterogeneous servers to a single storage system was made possible by the presence of high-end heterogeneous storage systems specially designed for information centers. They significantly reduce the cost of storage by creating a single high-performance storage system for multiple servers, eliminating the need to purchase separate devices for each server type. This approach also simplifies administration and saves the area occupied by the equipment.

When you use SAN-based storage consolidation, a shared pool of shared storage resources is created, including different types of storage. This model improves efficiency and reduces management complexity, increasing scalability, availability, and reliability of storage. In such information systematization, SAN acts as a mean to allow a large number of servers to have access to a single storage resource.

Discussion. In general, taking into account the above, it is relevant and necessary to develop and implement a system of systematization of information about customers within the system of systematization of information about the external environment of the company, adapted to the conditions of domestic business, in order to increase the competitiveness of the company. Such a system will provide the following technological advantages: ability to support different platforms; high performance due to distributed computing; almost linear scalability.

References

1. ACID Transactions. URL: <https://www.databricks.com/glossary/acid-transactions> (accessed: 28.07.2024).
2. Connolly, T., & Begg, C. (2004). *DataBase systems: A practical approach to design, implementation and management* (4th Edition). Boston, MA: Addison-Wesley Longman Publishing Co., Inc.
3. C. J. Date, A. Kannan and S. Swamynathan, *An Introduction to Database Systems*, Pearson Education, Eighth Edition, 2009.
4. Kitamura, Y., & Mizoguchi, R. (2004). Ontology-based systematization of functional knowledge. *Journal of Engineering Design*, 15(4), 327–351. <https://doi.org/10.1080/09544820410001697163>
5. Michele Fabi & Myriam Kassoul & Julien Prat, 2024. "Systematization of knowledge: constant function market makers," Chapters, in: Henrik Cronqvist & Desiree-Jessica Pely (ed.), *The Elgar Companion to Decentralized Finance, Digital Assets, and Blockchain Technologies*, chapter 10, pages 191-216, Edward Elgar Publishing.
6. Mullins, C. (2012). *Database administration: The complete guide to DBA practices and procedures*. Boston, MA: Addison-Wesley.
7. Peter Rob and Carlos Coronel, *Database Systems Design, Implementation and Management*, Thomson Learning-Course Technology, Seventh Edition, 2007.

8. Shio Kumar Singh, *Database Systems Concepts, Designs and Application*, Pearson Education, Second Edition, 2011.
9. Simion, Eduard, Systematization of Knowledge: Cybersecurity Techniques from a Military Intelligence Perspective (January 9, 2022). International Journal of Creative Research Thoughts, Volume 10, Issue 1, January 2022 ISSN: 2320-2882, Available at SSRN: <https://ssrn.com/abstract=4180622>
10. Silva, E.C. da, Barbosa, F.C. and Carrino, A.L. 2023. A study on the systematization and implementation of autonomous or self-service supermarkets. Seven Editora. (Oct. 2023).
11. The Latest Performance Management Statistics (2024 Update). URL: <https://www.selectsoftwarereviews.com/blog/performance-management-statistics> (accessed: 28.07.2024).
12. Zygiaris, S. (2018), "References", *Database Management Systems*, Emerald Publishing Limited, Leeds, p. 289. <https://doi.org/10.1108/978-1-78756-695-820181016>
13. Performance Management Statistics for 2024. URL: <https://www.myshortlister.com/insights/performance-management-statistics> (accessed: 28.07.2024).