Formation of Integrated Repositories of Social and Communication Data by Consolidating the Resources of Museums, Libraries and Archives in Smart Cities Projects

Oleksii Duda[a], Volodymyr Pasichnyk[b], Halyna Lypak[a], Nataliia Veretennikova[b], Nataliia Kunanets[b], Oleksandr Matsiuk[a] and Valentyna Mudrokha[a]

[a] *Ivan Puluj National Technical University, Ruska Str. 56, Ternopil, 46001, Ukraine*
[b] *Lviv Polytechnic National University, S. Bandery str., 12, Lviv, 79000, Ukraine*
[c] *Lviv National Library of Ukraine named after V. Stefanyk, Stefanyka Street 2, Lviv, Ukraine*

Abstract

One of the key components of process support in the resource and social and communication networks of a smart city is the organization of data warehouses and IT multidimensional data analysis – OLAP (operational analytical data processing). The use of this technology in the framework of information and technological support of processes in urban social and communication networks provides an opportunity to analyze the current state of social and communication resources of the city, as well as identify trends of their change and states by comparing data from different time periods and different collections.

The authors suggest an original information technology that promotes the formation of data warehouses, which accumulate consolidated information resources of libraries, archives, and museums. The key component of process support in the social and communication networks of a smart city is the information technology of multidimensional data analysis.

The procedure is based on the study of the characteristics of municipal social and communication resources, including the choice of attributes used for their parameterization. This approach ensures the effective consolidation of information resources established by the social memory of a smart city and allows you to simulate the structure of the data warehouse.

Keywords 1

Information technology; data warehouse; data structure; analytical processing

1. Introduction

Steady informatization of all life spheres encourages changes in the social, cultural, and political life of society and changes in social communications, which are transferred from the plane of interpersonal communication to the electronic remote information environment. In the context of the formation and rapid development of a knowledge-based society, such an environment is designed to provide opportunities for rapid and high-quality access to information that plays a role of significant social heritage. The formation of social and communication environment of a smart city encourages social institutions to seek new forms of their joint information activities to provide quality information services to users. Integration in the single information space of information resources of libraries, archives, and museums to preserve and transmit the cultural and historical heritage of mankind is recognized as one of the priorities of building an information society at the global level.

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine
EMAIL: oleksij.duda@gmail.com (O. Duda); vpasichnyk@gmail.com (V. Pasichnyk); halyna.lypak@gmail.com (H. Lypak); nataver19@gmail.com (N. Veretennikova), nek.lviv@gmail.com (N. Kunanets), oleksandr.matsiuk@gmail.com (O. Matsiuk), mudrokhavalentyna@gmail.com (V. Mudrokha)
ORCID: 0000-0003-2007-1271 (O. Duda); 0000-0002-5231-6395 (V. Pasichnyk); 0000-0001-9187-5758 (H. Lypak), 0000-0001-9564-4084 (N. Veretennikova) 0000-0003-3007-2462 (N. Kunanets), 0000-0003-0204-3971 (O. Matsiuk), 0000-0001-9508-227X (V. Mudrokha)

Some scholars such as N. Gwyn, J. Trant and others tend to use the term of convergence when it comes to the cooperation of libraries, archives, museums. This approach makes sense if we emphasize the functional, semantic and mental differences of these institutions. It is believed that the convergence of LAM as a process of convergence of unrelated objects, accompanied by the acquisition of common features, took place in the pre-digital period of cooperation between libraries, archives, museums, when it took the form of joint exhibitions, thematic associations, and other forms of interaction. If it is talked about combining digital collections of these institutions, when each exhibit, artifact, etc. is presented in the form of an electronic document and it is possible to access the entire set of these documents from a single access point, it is better to use the term of consolidation.

In general, consolidation means strengthening, uniting, integrating, grouping something. With regard to information consolidation, a distinction is made between physical consolidation, i.e. information actually gathered in one place, and logical consolidation that is distributed information, which is in a single repository with a common directory and equal access to it from the user's point of view. The concept of consolidated information is often used in the context of physical consolidation, when it refers to analytical and synthetic activities to select complete and relevant information to support decision-making in a particular problem area [1]. This direction is also synonymous with competitive intelligence. In the context of the objectives of our study, we are talking about the second type of consolidation, when information is selected from distributed sources and provided to a user through a user-friendly human-machine interface.

In the process of creating consolidated information, data is first transformed into information, then into knowledge and, finally into an information product [1].

From the point of view of our study a consolidated information product can be some electronic resources of certain social memory institutions, selected to facilitate the full satisfaction of a user request, and which can be represented by electronic data (information in any form), electronic programs (sets of operators or subprograms that ensure the implementation of certain tasks, including data processing) or a combination of these types in one resource. As a rule, they are remote access resources.

The social and communication environment of a smart city is a complex system where urban communications, information technologies of data transmission, databases and data warehouses are organically interconnected, furthermore, its important element is the consolidated information resource of libraries, archives, and museums. The social and communication environment of a smart city operates through the continuous processing and support of access to information that is consolidated in databases and data warehouses. After processing the consolidated data, their analysis takes place, which facilitates their transfer to a higher level of data management. A city can be distinguished as smart when social capital, traditional and modern communication infrastructure stimulate sustainable economic development and high life quality.

The information product created as a result of information consolidation, as well as the entire information technology consolidation system as material objects and information carriers are covered by the general definition of the resource, and therefore have specific qualitative and quantitative characteristics, which are defined later in this paper.

2. Analysis of the modern research state

With the development of information technology, which contributes to the evolution of information processing methods, including digital ones, there are new prospects for projects to combine information resources of libraries, archives, museums. Advances in information and communication technologies make it possible to present an artifact of uncontrolled origin in the form of an electronic document and integrate it into a data warehouse that consolidates the information resources of libraries, archives, and museums in a smart city.

The relevance of unification processes in the field of fundamental social institutions, which are libraries, archives, museums, is confirmed by the active discussion of this issue at various conferences, symposia, and meetings of professional associations.

For example, in the UK, similar activities were coordinated by the Museums, Libraries and Archives Council (MLA), which considered the issue of combining the resources of libraries,

museums and archives in the context of continuing education and formulated proposals on this matter for the government. It was funded by the Arts Council England Press Office until 2012, and then the MLA functions were transferred to the Arts Council and the National Archives.

The search for the most rational ways to preserve the national heritage from the funds of libraries, museums and archives of different countries has stressed the need to achieve a balance between the formation of information resources of regional and international importance [2].

Researches by Higgins S., Trant J., Duff W., Marty P., Tammaro A., Kirchhoff T., Mitchell E., T. Gill T, Van Hooland S., Verborgh R. are devoted to various aspects of consolidation of digital collections of social memory institutions and the specifics of their interaction in the online environment.

J. Trant [3] outlines a number of problems that accompany the integration of digital collections of libraries, archives, museums, in particular:

- the effectiveness of the use of digitized surrogates, as the question of certification of copies to confirm the conformity of the original arises; the indisputable advantages of digitization include the preservation of fragile old prints, an ability to reuse digital copies and provide wide access to them;
- management of digital collections, which should guarantee obtaining of the necessary information support by users, facilitate the search, detection and possibility of further use of information resources;
- usage support thanks to the ability to combine collections of libraries, archives and museums, and access to their resources from any device, anytime, anywhere is open;
- the need to study and apply the positive experience of integration of digital collections of libraries, archives, museums.

The development of effective repositories of social and communication data that combine the resources of libraries, archives and museums is impossible without the use of agreed metadata descriptions of these resources. Thorough research on how to describe metadata on digitized cultural heritage resources belongs to the belgian scientist S. Van Hooland. His research covers the evolution, quality and typology of metadata in the cultural heritage sector, their relationship to the semantic and ontological representation of data, as well as the creation of custom metadata [4]. Van Hooland widely developed the concept of Linked Data. In co-authorship with colleague Verborgh R. in the manual "Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata" (2014), they presented the basic concepts of metadata and related data standards, provided advice on their practical application regarding existing metadata; offered tools and explanations on how to achieve maximum results in the description of resources [5].

Some scholars (N. Gwyn, J. Trant and others) tend to use the term of convergence when it comes to the cooperation of libraries, archives, museums. This approach makes sense if we emphasize the functional, semantic and mental differences of these institutions. It is believed that the convergence of LAM as a process of convergence of unrelated objects, accompanied by the acquisition of common features, took place in the pre-digital period of cooperation between libraries, archives, museums, when it took the form of joint exhibitions, thematic associations, and other forms of interaction. If it is talked about combining digital collections of these institutions, when each exhibit, artifact, etc. is presented in the form of an electronic document and it is possible to access the entire set of these documents from a single access point, it is better to use the term of consolidation.

According to some authors, smart cities are a viable solution for combining information resources, human capital, social capital, and information and communication technologies to promote their development. The study aims to analyze the impact of smart management factors on the life quality, which has become possible using the methodology of structural equation modeling. The main focus is on measuring transparency, communication on life quality issues, which can help smart city leaders in developing public policies and actions of the municipal executive power to improve the life quality of citizens [6].

The next paper focuses on the study of the method of choice and use of communication means, which limits the understanding of modern use of information and communication technologies. A combinatorial perspective has been developed to determine the complexity of using information and

communication technologies to achieve communication goals and tasks, which can have a significant impact on the productivity and efficiency of individuals and corporations [7].

To coordinate the use of infrastructure and urban assets to build so-called smart cities, it is proposed to use various tools from sensors to services and artificial intelligence technologies. The smart city is suggested to be modeled as a multi-layered architecture in which the communication and network layer plays a central role for the implementation of intelligent control actions and secure information exchange [8].

3. Characteristics of the social and communication resource data storage of a smart city

Implementing projects of consolidation of information resources of libraries, archives and museums, a number of problems need to be solved such as formation of a system of effective information service for a user; providing a fundamentally new approach to the formation of information resources taking into account unpredictable and existing situations; the need to create institutions to ensure the processes of providing stable information services, development of tools and technologies for the consolidation of information resources, the formation of databases and data warehouses. Thus, the formation of social and communication networks of a smart city is occured.

A data warehouse that combines the information resources of libraries, archives, and museums is a subject-oriented, integrated, immutable, chronological set of data that can be a comprehensive source of information for IT_{IDP} and IT_{DSS} [9].

Data warehouses for storing information entities in the social and communication networks of a smart city have the following features:

- obtaining, normalizing, storing and presenting information in a detailed and aggregated form from various sources of social and communication resources of a smart city, including museums, libraries, archives, media, government agencies and municipalities;
- multidimensional presentation of data collections accompanied by metadata with descriptions of the data structure and structure of data warehouses for processes occurring in social and communication networks;
- availability of software agents for the implementation of procedures for importing data from existing databases and data warehouses of museums, libraries, archives, media, government agencies and municipalities and procedures for uploading them in universal formats; availability of procedures for complex analytical processing on the basis of IT_{IDP} and IT_{DSS} in order to obtain new data and knowledge;
- availability of software and algorithmic complexes, mobile and web applications for the implementation of the function of social communication and support of information services to residents and guests of a smart city;
- subject orientation of the presentation of data collections on the course of processes in the social and communication networks of a smart city.

The scheme of data aggregation in data warehouses for social and communication networks of a smart city is presented in Fig. 1.

The generalized representation of data warehouses on information and technological support of processes in social communication networks of a smart city has the following features:

Subject orientation. Data in warehouses are organized according to the main aspects of the activities of urban social and communication networks (organization or institution, in particular, museum, category, type, essence, or archive, type, document, essence, etc.). The correct organization of data warehouses of urban social and communication networks in accordance with the subject-oriented approach allows significant simplification of the procedures of analytical processing and increasing the speed of search and analytical queries.

Data are stored in specialized multidimensional databases based on *n*-dimensional cubes.

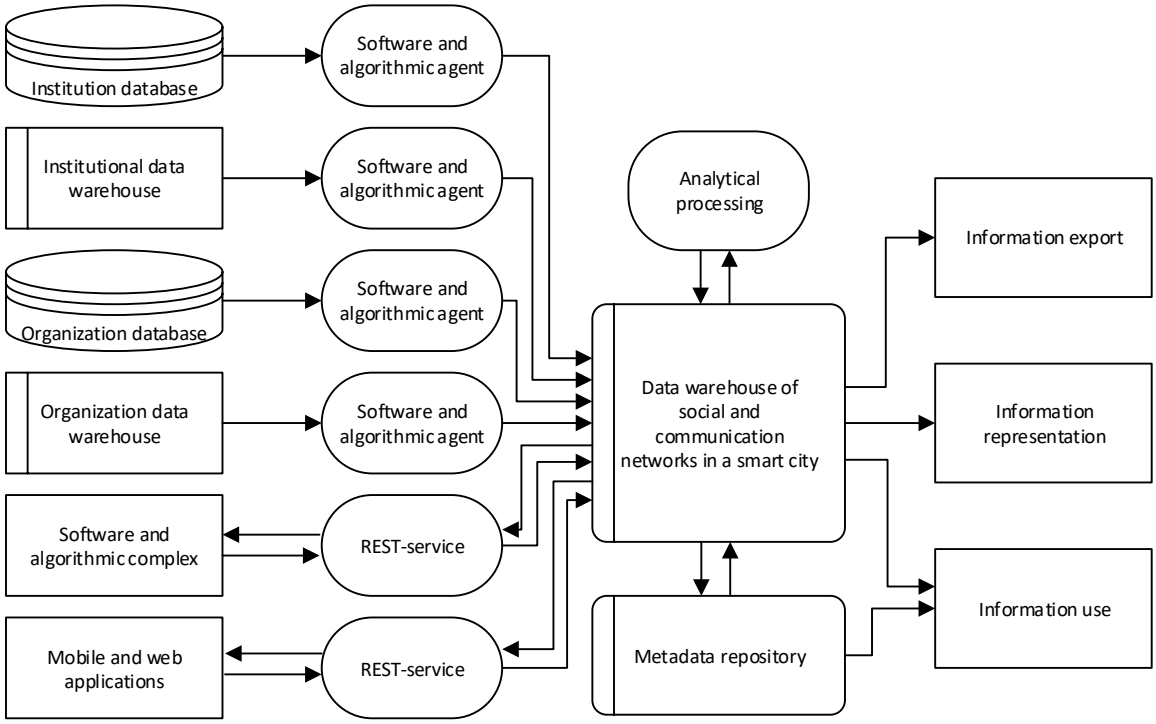

Figure 1: Scheme of data aggregation in data warehouses for social and communication networks of a smart city

Integration. The initial data on the data collections of social and communication resources of a smart city are extracted from the databases of institutions, organizations, software and algorithms complexes, mobile and web applications. The obtained data are validated, cleaned and filtered, normalized and aggregated (there is a primary processing and calculation of generalized and total indicators) and uploaded to the appropriate repositories.

Time data binding. The data stored in the warehouses on the processes that take place in the urban social communication networks must have unambiguous timestamps with reference to certain time periods. Data selected from online databases are stored in data warehouses in the form of historical collections according to a specific period of time. Usually, data from urban social and communication networks are characterized by high dynamics and transience of processes, which allows to analyze trends in the processes of their operation.

Irremovability. Getting in a certain historical collection of repositories, data on urban social and communication networks cannot be changed and adjusted.

A formalized description of data warehouses of social and communication networks of a smart city is presented in the form [10]:

$$DW = \langle NR, DB, RF, rm, RM, func \rangle, \quad (1)$$

where *NR* is a set of social and communication networks,

DB is a set of relationships and their schemes and restrictions, which contain information from the input databases (databases of institutions and organizations, including museums, archives, and media),

RF is a scheme of the set of relations of facts *rf*,

RM is a scheme of the set of metadata relations *rm*,

func is a set of decision-making procedures.

Gaining new knowledge is to extract data from the collections of data warehouses by implementing functions *func* on the facts *rf* considering the requirements that are directly formed in accordance with the needs of users. The connections between the sets *RF* and *DB* form data hypercubes, and their dimensions are the sets of the relations *RF* of the data warehouses of urban social and communication networks. Considering the peculiarities of the tools of intellectual analytical processing in the process of designing data warehouses for social and communication

networks of a smart city opens opportunities for more accurate disclosure and presentation of data collected in the relevant collections [10].

4. Information technology of multidimensional analysis of data in urban social and communication networks

One of the key components of process support in the social and communication networks of a smart city is the organization of storage and IT multidimensional data analysis – OLAP (operational analytical data processing). The use of this technology in the framework of information and technological support of processes in urban social and communication networks provides an opportunity to analyze the current state of social and communication resources of the city, as well as identify trends and changes by comparing data from different time periods and different collections.

The processes of generalization of detailed data collections collected in social and communication networks are usually used in OLAP smart city systems, which allows to obtain new knowledge from consolidated in various aspects of the analysis, interconnected information sets. The source data sets for analytical processing with the help of OLAP information on urban social communication networks are a large list of information entities, including archival documents, museum exhibits, publications in the media, etc.

IT OLAP is based on a multidimensional data model, its main basic essences are hypercube of data *rel*, dimension *D*, attribute *A*, cell *X*, value *rel(D,A)*.

A data hypercube, being an ordered set of cells, contains one or more dimensions. Each of the cells of the hypercube is defined by only one unique set of attribute values that are dimensions. The cell can contain an actual value, or be considered empty one containing Null.

The set of attributes that form one of the faces of a hypercube is called a dimension. For urban social and communication networks, important dimensions are, in particular, time and geography. The lists of years, quarters, months, days, hours, minutes, seconds are attributes for time measurement. The list of administrative and territorial objects such as regions, districts, settlements, neighborhoods, streets, houses, apartments are attributes for geographical measurement.

To access the data in the hypercube, the user selects one or more cells by specifying (fixing) the appropriate dimension values. The set of fixed dimension values is called the set of fixed attributes.

Multidimensional analysis of data on processes in urban social and communication networks, in turn, involves research and analysis of the characteristics of municipal social and communication resources, including the choice of attributes used to parametrize the information model (Fig. 2).

The set of categories and attributes SA used to describe the processes taking place in the social and communication networks of a smart city is presented as follows:

$$SA = \left\langle SA_{org} \cup SA_{clases} \cup SA_{res} \cup SA_{GEO} \cup SA_{usr} \cup SA_{query} \right\rangle, \quad (2)$$

where $SA_{org} = \left\{SA_{org}^{S}\right\}, S = \{Archives, Museum, Biblio, Governement, ZMI\}$ are attributes for each selected category of social and communication resources of a smart city, including archives (*Archives*), museums (*Museums)*, libraries (*Biblio*), government agencies (*Government)* and the media (*ZMI*);

SA_{clases} is a category of attributes of characteristics and properties of information essence classification of social and communication resources;

SA_{res} is a category of attributes of characteristics and properties of information social and communication resources;

SA_{GEO} is a category of geolocation attributes;

SA_{usr} is a category of attributes of the description of individual consumers' services of social and communication resources of a smart city;

SA_{query} is a category of attributes to describe information and search queries for social and communication resources of a smart city.

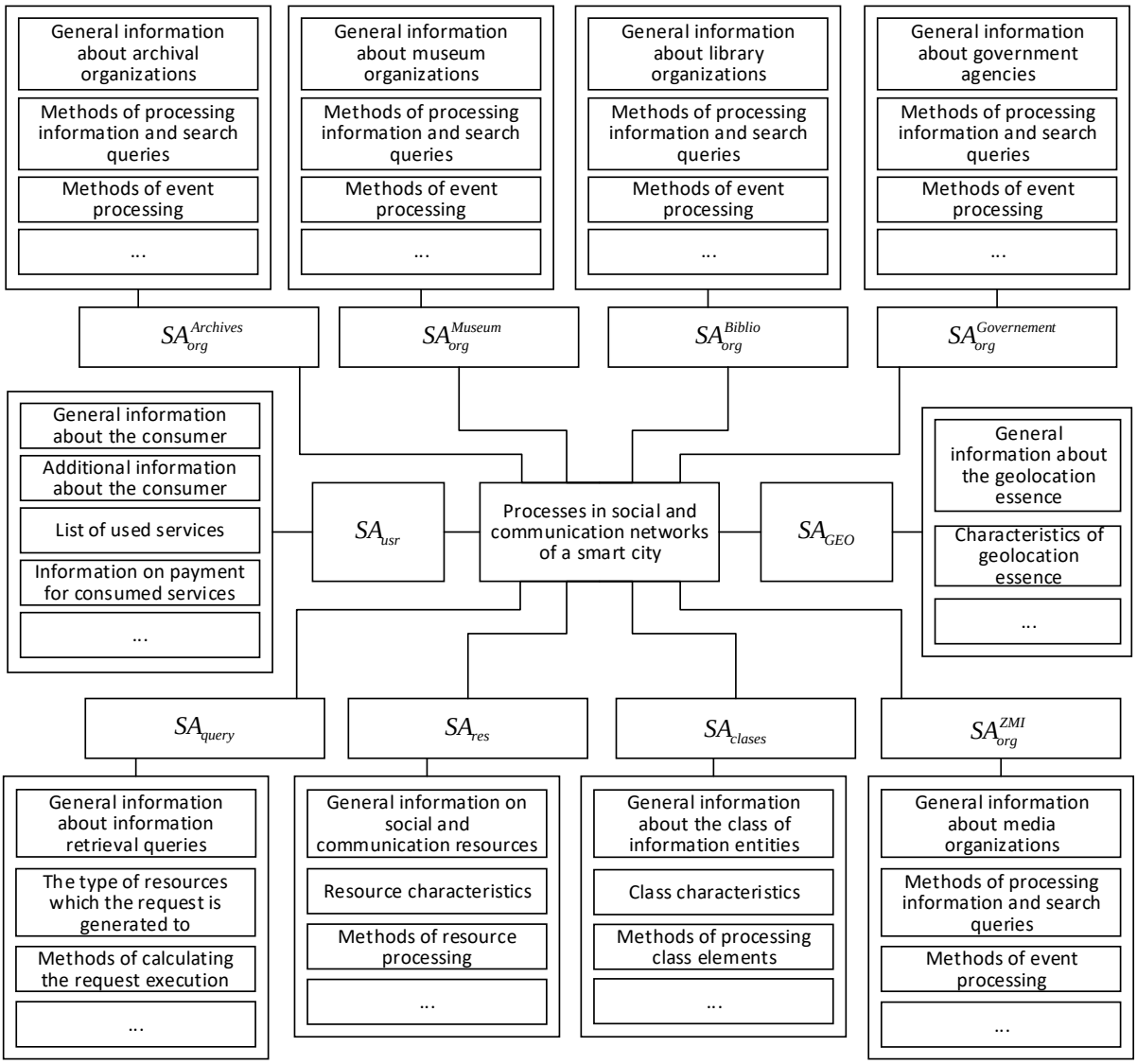

Figure 2: Attributes of the data hypercube for information technology support of processes in urban social and communication networks

A significant number of attributes used to describe the entities in the social and communication networks of a smart city are synonymous and have the same meaning. These attributes can be divided into the following subcategories:

Fonds is characteristics of social and communication funds;

Src is information on social and communication resources;

Docs is characteristics of the query results;

Geopoint is a geolocation point;

User is consumer information;

Provider is information on service providers;

Use is information on the consumption of services.

The information network of the process of providing information services in urban social and communication networks will be presented by a set of parameters (Fig. 3).

Each of the attributes for a separate category $Socio = \left(Socio_{Fonds}, Socio_{Src}, Socio_{Docs}, Socio_{Geopoint}, Socio_{User}, Socio_{Provider}, Socio_{Use}\right)$, provides a description of the relevant characteristics of the processes occurring in social and communication networks, and in total fully characterize the processes of providing services in social and communication networks of a smart city.

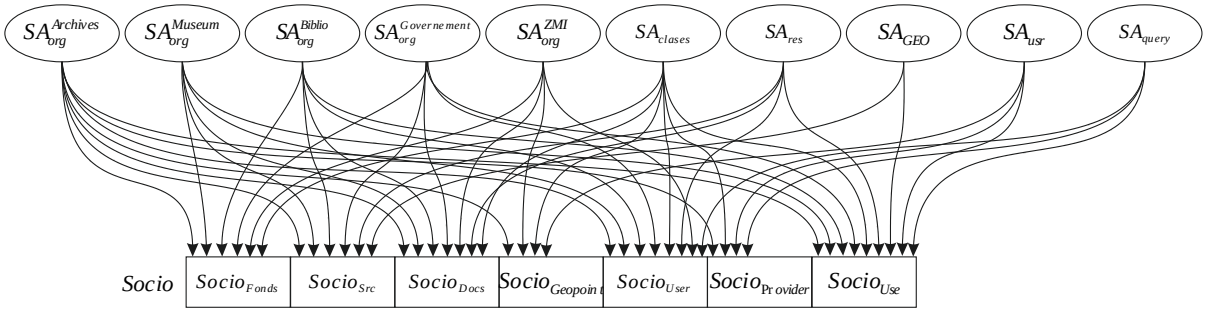

Figure 3: Aggregation of attributes of the process for providing information and service services in urban social and communication networks

The attributes selected to describe the processes were used to build a multidimensional information model of the set of social and communication networks of a smart city, presented in the form of a data hypercube:

$$S^{Socio}_{Prov}\left(Dim^{Socio}, Attr^{Socio}\right), \quad (3)$$

where Dim^{Socio} is a set of hypercube dimensions;

$S^{Socio}_{\Pr ov}$ is a a smart city service provided by a group of performers $Prov^{Socio} = \left\{Prov^{Socio}_{k}\right\}, k = \overline{1,t}$ using a number of urban social and communication resources $F_j, j = \overline{1,r}$, including archives, museums, libraries, government agencies and the media;

$Attr^{Socio}_{Dim^{Socio}_{i}} = \left\{Attr^{Socio}_{1_i}, Attr^{Socio}_{2_i}, \ldots, Attr^{Socio}_{n_i}\right\}, i = 1, \ldots, m$, is a set of attributes that are specific to the dimension Dim^{Socio}_{i};

$Attr^{Socio} = Attr^{Socio}_{1} \cup Attr^{Socio}_{1} \cup \ldots \cup Attr^{Socio}_{m}$ is a set of attributes that are typical for a hypercube;

$Dim'^{Socio} \subseteq Dim^{Socio}$ is a set of fixed measurements of the hypercube;

$Attr'^{Socio} \subseteq Attr^{Socio}$ is a set of fixed attributes of the hypercube.

It is introduced the notation $S'^{Socio}_{\Pr ov}\left(Dim'^{Socio}, Attr'^{Socio}\right)$ to describe the subset of the hypercube of social and communication data $Prov^{Socio}$ which is placed in accordance with the set of fixed values and to their access it is necessary to specify the set of fixed dimensions $Dim'^{Socio} \subseteq Dim^{Socio}$ and values of fixed attributes $Attr'^{Socio} \subseteq Attr^{Socio}$. The set of cells corresponding to the fixed attributes and measurements of social and communication data is denoted as $S'^{Socio}\left(Dim'^{Socio}, Attr'^{Socio}\right) \middle| S_{\Pr ov}{}'^{Socio} \subseteq S^{Socio}_{\Pr ov}$. The dimension key is called an attribute that uniquely identifies the tuple (string) of the hypercube dimension. The above analyzed sets of parameters for the description and information technology support of processes in the social and communication networks of a smart city allow to form the following dimensions of the multidimensional data model:

Dim^{Socio}_{1} is data on organizations providing services based on the resources of social and communication networks of a smart city from the set $Prov^{Socio} = \left\{Prov^{Socio}_{k}\right\}, k = \overline{1,t}$, in particular, $Attr^{Socio}_{Dim^{Socio}_{1}} = \left\{Attr^{Socio}_{i_1}, i = \overline{1,10}\right\}$ is a set of dimension attributes Dim^{Socio}_{1}, where $A^{Socio}_{1_1}$ is a type of service provider organization, including archives, museums, libraries, government agencies and the media; $A^{Socio}_{2_1}$ is a name of the service provider organization, $A^{Socio}_{3_1}$ is a settlement, $A^{Socio}_{4_1}$ is a form of ownership, $A^{Socio}_{5_1}$ is a microdistrict, $A^{Socio}_{6_1}$ is a street, $A^{Socio}_{7_1}$ is a street, $A^{Socio}_{8_1}$ is an identifier of the geographic information system reference point, $A^{Socio}_{9_1}$ is an identifier of a responsible person, $A^{Socio}_{10_1}$ is a phone number of a responsible person.

Dim^{Socio}_{2} is data on time periods, in particular, $Attr^{Socio}_{Dim^{Socio}_{2}} = \left\{Attr^{Socio}_{i_2}, i = 5\right\}$ is a set of dimension attributes Dim^{Socio}_{2}, where $A^{Socio}_{1_2}$ is a name of the time period, $A^{Socio}_{2_2}$ is the start date of the time period, $A^{Socio}_{3_2}$ is the end date of the time period, $A^{Socio}_{4_2}$ is a brief summary of the time period, $A^{Socio}_{5_2}$ is the parent time period.

Dim_3^{Socio} is data on classification of fundraisers, in particular, $Attr_{Dim_3^{Socio}}^{Socio} = \left\{Attr_{i_3}^{Socio}, i = 6\right\}$ is a set of dimension attributes Dim_3^{Socio}, where $A_{1_3}^{Socio}$ is a type of classification of fundraisers, $A_{2_3}^{Socio}$ is a name of class of fundraisers, $A_{3_3}^{Socio}$ is a functional purpose (direction of activity) of class of fundraisers, $A_{4_3}^{Socio}$ is a brief description of class of fundraisers, $A_{5_3}^{Socio}$ is a parent class, $A_{6_3}^{Socio}$ is an essence identifier of the geographic information system.

Dim_4^{Socio} is data on fundraisers, in particular, $Attr_{Dim_4^{Res}}^{Socio} = \left\{Attr_{i_4}^{Socio}, i = \overline{1,9}\right\}$ is a set of dimension attributes Dim_4^{Socio}, where $A_{1_4}^{Socio}$ is a type of fundraiser, it can be a legal or physical person, $A_{2_4}^{Socio}$ is a full last name of the fundraiser (for organizations) or surname, name and patronymic (for individuals), $A_{3_4}^{Socio}$ is a functional purpose of the legal person or the value of the individual person in the society life, $A_{4_4}^{Socio}$ is a brief description of the fundraiser, $A_{5_4}^{Socio}$ is a date of foundation (for organizations) or birth (for individuals) of the fundraiser, $A_{6_4}^{Socio}$ is a date of closing (for organizations) or death (for individuals) of the fundraiser, $A_{7_4}^{Socio}$ is a parent class of funds from the set of measurements Dim_3^{Socio}, $A_{8_4}^{Socio}$ is an organization holder of the fund from the set $Prov^{Socio}$, $A_{9_4}^{Socio}$ is an essence identifier of the geographic information system.

Dim_5^{Socio} is data on social and communication funds, in particular, $Attr_{Dim_5^{Socio}}^{Socio} = \left\{Attr_{i_5}^{Socio}, i = \overline{1,5}\right\}$ is a set of dimension attributes Dim_5^{Socio}, where $A_{1_5}^{Socio}$ is a name of the fund, $A_{2_5}^{Socio}$ is an identifier of social and communication services from the set Dim_1^{Socio}, $A_{3_5}^{Socio}$ is an identifier of the fundraiser from the set Dim_4^{Socio}, $A_{4_5}^{Socio}$ is unique attributes of the fund, $A_{5_5}^{Socio}$ is an essence identifier of the geographic information system.

Dim_6^{Socio} is data on the systematization of entities in the funds, in particular, $Attr_{Dim_6^{Socio}}^{Socio} = \left\{Attr_{i_6}^{Socio}, i = \overline{1,6}\right\}$ is a set of dimension attributes Dim_6^{Socio}, where $A_{1_6}^{Socio}$ is a name of the essence category, $A_{2_6}^{Socio}$ is a functional purpose of the essence category, $A_{3_6}^{Socio}$ is a brief description of the essence category, $A_{4_6}^{Socio}$ is a parent essence category, $A_{5_6}^{Socio}$ is unique attributes of the essence category, $A_{6_6}^{Socio}$ is an essence identifier of the geographic information system.

Dim_7^{Socio} is data on types of information elements for data storage, in particular, $Attr_{Dim_7^{Socio}}^{Socio} = \left\{Attr_{i_7}^{Socio}, i = \overline{1,4}\right\}$ is a set of dimension attributes Dim_7^{Socio}, where $A_{1_7}^{Socio}$ is a name of data type, it can be a database field, a text file, a binary file, URL, etc., $A_{2_7}^{Socio}$ is a brief description of a type of information elements, $A_{3_7}^{Socio}$ is a list of attributes of information elements type, $A_{4_7}^{Socio}$ is a list of characteristics of the type of information elements.

Dim_8^{Socio} is data on essence types in funds, in particular, $Attr_{Dim_8^{Socio}}^{Socio} = \left\{Attr_{i_8}^{Socio}, i = \overline{1,5}\right\}$ is a set of dimension attributes Dim_8^{Socio}, where $A_{1_8}^{Socio}$ is a name of essence type, it can be an archival document, a museum artifact, a library book, a newspaper article, etc., $A_{2_8}^{Socio}$ is a brief description of essence type, $A_{3_8}^{Socio}$ is a list of essence type attributes, $A_{4_8}^{Socio}$ is a list of characteristics of essence type, $A_{5_8}^{Socio}$ is a list of information elements used to store data from the dimension set Dim_7^{Socio}.

Dim_9^{Socio} is data on entities in the funds, in particular, $Attr_{Dim_9^{Socio}}^{Socio} = \left\{Attr_{i_9}^{Socio}, i = \overline{1,7}\right\}$ is a set of dimension attributes Dim_9^{Socio}, where $A_{1_9}^{Socio}$ is a name of the essence, $A_{2_9}^{Socio}$ is a time period to which the essence from the set of dimension belongs to Dim_9^{Socio}, $A_{3_9}^{Socio}$ is a social and communication fund from the set Dim_5^{Socio}, $A_{4_9}^{Socio}$ is a class of essences from the set Dim_6^{Socio}, $A_{5_9}^{Socio}$ is a type of set of essence funds used to store data Dim_8^{Socio}, $A_{6_9}^{Socio}$ is an essence identifier of the geographic information system, $A_{7_9}^{Socio}$ is the content of the information essence, or resource identifier.

The values of the dimensions (attributes) discussed above along each axis of the hypercube are hierarchically combined into one or more levels depending on the dimension. Due to this, hierarchical dimensions are formed for the purpose of further analysis of data by their aggregation, or detailing (Fig. 4).

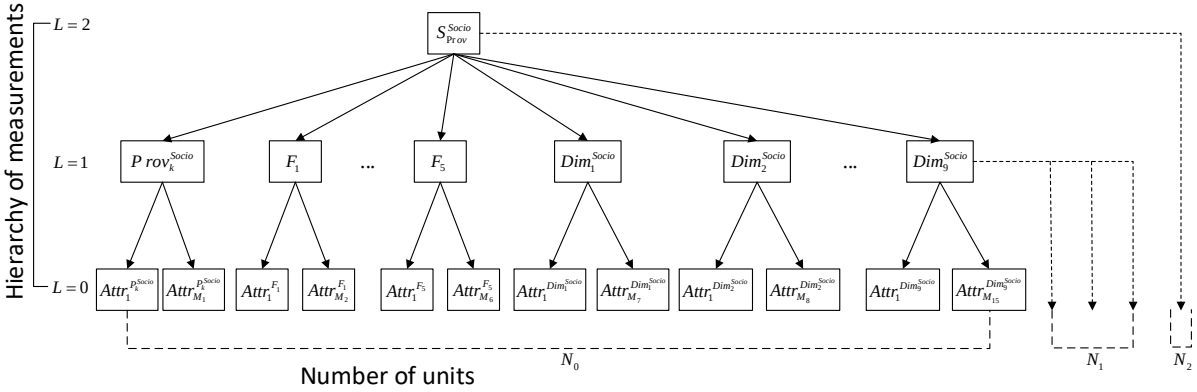

Figure 4: Aggregation of the hypercube of data of social and communication networks in the smart city. One-dimensional representation

Data aggregation of OLAP-cubes of social and communication networks of a smart city is a process of obtaining new aggregates (values) corresponding to the attributes of a certain level L based on lower-level values *L–1*. Such dimensions are axes of hierarchical dimensions, in particular $Prov_i^{Socio}$, F_j and Dim_k^{Socio} that contain the initial attributes of the lower hierarchical level (*L=0*) and can be supplemented by attributes of higher hierarchical levels from *L=1*. The slice operation is used to obtain the required subsets of social and communication cells $S_{Prov}^{Socio'} \subseteq S_{Prov}^{Socio}$ in order to cut off the unnecessary values by sequentially fixing the attributes.

5. Conclusions

The use of information and communication technologies contributes to the formation of data warehouses, which store images of artifacts or electronic documents. Thus, a consolidated information resource is formed based on the funds of libraries, archives, and museums of a smart city.

One of the key components of process support in the social and communication networks of a smart city is the creation of data warehouses and information technology for multidimensional data analysis – OLAP (operational analytical data processing).

Multidimensional analysis of data on the processes occurring in urban social and communication networks is based on the study of the characteristics of municipal social and communication resources, including the choice of attributes used for their parameterization.

The values of dimensions (attributes) are hierarchically combined into one or more levels depending on the dimension. Due to this, hierarchical dimensions are formed to further analyze the data by aggregating or detailing. This approach ensures the effective consolidation of the information resources of the social memory institutions in a smart city.

Thus, the paper presents the information technology of multidimensional data analysis in urban social and communication networks based on the methodology of building data hypercubes. This allowed the classification and parameterization of many categories and attributes used to describe the processes in the social and communication networks in a smart city and led to the formation of multidimensional data models used in the formation of prototypes of data warehouses. These repositories, in contrast to existing ones, are better adapted to the dynamic changes in the structure and states of social and communication networks of a large city.

Libraries, museums and archives, as intellectual elements of the social and communication environment of a smart city, get new effective tools for processing, storage, and analysis of existing information resources, which accelerates and improves the quality of analytical and synthetic

processing procedures, and users get remote access to documents from their funds through a single access point.

This approach simplifies and increases the efficiency of not only the institutions of social memory, but the social and communication environment in general, increases the efficiency of technologies for processing information flows circulating in it.

6. References

[1] N. Kunanets, V. Dobrovolska, N. Filippova, K. Parviz, H. Lypak, O. Duda, N. Veretennikova, M. Karp, C. Shunevych, L. Dubrovina, Designing the Repository of Documentary Cultural Heritage in: Advances in Intelligent Systems and Computing, Vol. 1293, 2020, pp. 1034–1044.

[2] S. Sanja, T. Giuffrè, A. Deluka-Tibljaš, Smart mobility solutions – necessary precondition for a well-functioning smart city in: Transportation Research Procedia, Vol. 45, 2020, pp. 604–611.

[3] J. Trant, Emerging convergence? Thoughts on museums, archives, libraries, and professional training in: Museum Management and Curatorship, Vol. 24, №. 4, 2009, pp. 369-387.

[4] S. Van Hooland, From spectator to annotator: possibilities offered by user-generated metadata for digital cultural heritage collections in: Immaculate Catalogues: Taxonomy, Metadata and Resource Discovery in the 21st Century, Proceedings of CILIP Conference, 2006.

[5] S. Van Hooland, R. Verborgh, Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata, Facet publishing, 2014.

[6] Julio Cesar Ferro De Guimarães, Eliana Andréa Severo, Luiz Antonio Felix Júnior, Wênyka Preston Leite Batista Da Costa, Fernanda Tasso Salmoria, Governance and quality of life in smart cities: Towards sustainable development goals in: Journal of Cleaner Production, Vol. 253, 2020.

[7] Y. Zhu, Combinatorial use of communication technologies in organizations in: Corporate Communications: An International Journal, Vol. 24, №. 4, 2019, pp. 623-635.

[8] A. Haidine, S. El Hassani, A. Aqqal, A. El Hannani, The role of communication technologies in building future smart cities, 2016.

[9] A. Sinaeepourfard, et al., F2c2C-DM: a fog-to-cloudlet-to-cloud data management architecture in smart city in: 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), 2019.

[10] V. Pasichnyk, N. Shakhovska, Data warehouses: Textbook, 2007 (in Ukrainian).