IMPLEMENTATION OF CIDOC CRM ONTOLOGY IN UKRAINE FOR THE DEVELOPMENT OF THE DOCUMENTATION SYSTEMS AND DATA EXCHANGE

Vladyslav Pioro[1], Anastasiia Cherednychenko[1], Halyna Holubeva[2]

[1] Ukrainian Centre for Museum Development, 20-Mala Zhytomyrska str., 01001, Kyiv, Ukraine
[2] Ukrainian Centre for Cultural Studies, 9 Lavrska str., b. 5, 01015, Kyiv, Ukraine

mail@promuseum.org

Abstract. Following article describes the attempt to implement formal ontology CIDOC CRM in the context of reforming the cultural heritage and objects of cultural value inventory system in Ukraine. In particular, the standard became the basis for inventory processes and data structures formalization, development of the national metadata schemes etc.

Keywords: heritage, museum collection, documentation systems, data exchange, CIDOC CRM, formal ontology.

1. Introduction

We do realize that our experience in the suggested topic is not innovative in a general context. At the same time, it is extremely important for us to discuss the vision of the national digital transformation program and the description of its embodiment first steps. We hope that comments and advice of our colleagues will allow us to find the more effective approach in solving the problems Ukraine is facing today.

The system of cultural heritage objects inventory in Ukraine requires radical reforming. The approaches, that were inherited from Soviet Union and remained unalterable since 1980-th, do not correspond to the actual needs and challenges of the time. In particular, the main data massif on museum collections is stored: a) in analogue format (paper), b) locally, directly in particular museum. Complete and reliable information on quantitative and qualitative content of Museum fund of Ukraine (MFU) does not exist at the national level. Unified registry of MFU does not exist. There are approximately 130 000 objects in Ukraine, which have status of immovable heritage. Traditionally their data (inventory data) are stored on paper. Only about 10% of them are represented in States digital registry of immovable heritage. However, this registry, which was initiated almost 20 years ago, is actually simple spreadsheet static list of objects with very short object description by few attributes only ("Реєстр нерухомих пам'яток України" n.d.). It does not include spatial coordinates, detailed essential characteristics and information on the protection, does not reflect real condition of heritage object, information on monitoring and restoration processes etc.

This situation imposes essential restrictions on data access, significantly narrowing their usage potential. Furthermore, it might be said even about risks from the cultural heritage safety point of view. For example, the loss of the museum (caused by social turbulences, fire, natural disasters etc.) means in this context the irreversible loss of information about its collection. Partial occupation of the territory of Ukraine in 2014, unfortunately, became eloquent demonstration of this threat. Complete lists or at least exact data on the quantity of the museum collections items remained at the occupied territory are not exist. Ministry of culture estimates the loss in the governmental part of the MFU only approximately (10% – around 1 200 000 objects). The loss in non-governmental part of the MFU seems impossible to estimate at all.

This tragic situation led to realizing of the problem scale at the state level, testified the acute topicality of the inventory processes reforming and data operation in a cultural heritage area. This understanding determined further movement direction and activity plan for the period of 2014–2020.

2. Digital transformation in cultural heritage of Ukraine

The first official reaction on the articulated problem was solution that seemed obvious – if digital registries of cultural heritage and cultural values are not exist, then they have to be created. That is the task defined hot on the trail as one of the most important priorities of the Ukrainian government (Кабінет Міністрів України 2017).

Yet rather soon the understanding came, that this task in fact is not as simple as it looks at the first sight, it requires complex approach. Because nation-wide digital registries of cultural heritage is just top of the iceberg, element of complex system. The understanding where digital data will be taken from, who will generate them, in context of which processes, how their unification and aggregation will be provided etc. is extremely important for shaping such registries.

2.1 General principles

So the vision of what has to be done was substantially expanded. The overall purpose was defined as creation of the national infrastructure of cultural heritage objects data (NID), which in particular have to include:

- rules and standards of data generation, storage, publication and application, data exchange (formal ontology, metadata schemas, thesauruses, data exchange formats and protocols etc.);
- relevant organizational structure (subjects of data generation, storage, aggregation and publication, regulatory bodies, centers for education, researches and methodology etc.);
- hardware and software of different level (networks, platforms, sub-software, software application instruments for different needs, web-services, templates of data transformation etc.).

Main principles of NID creation and functioning were defined as:

- relevancy, authenticity, completeness and wholeness of cultural heritage object data;
- openness and accessibility of public character data;
- interoperability (in particular, technical and semantic);
- clear and transparent rules (Міністерство культури України 2017).

On the basis of these principles, digitalization of business processes at sectoral levels (museums, archives, libraries, protection bodies of immovable heritage etc.) is carried out and the development of digital interaction between individual actors for generation and use of data has begun.

The overcoming the sectoral isolation in the work with separate kinds of cultural heritage objects, the effective work with data of movable, immovable and intangible objects based on unified conceptual principles, the international integration based on Linked Data and Semantic Web were defined as important targets too.

In this connection importance of CIDOC CRM implementation as basic formal ontology in Ukraine was emphasized already at the level of shaping the NID general vision. So classes and properties of CIDOC CRM were used as the ground for metamodeling of other NID components, in particular for the inventory processes and structural data description.

2.2 Processes

Ukrainian museums, reservations and heritage protection authorities follow the traditional inventory procedures that have analogue format in their nature and oriented on entirely paper document circulation (“paper mentality”). They are unable to provide a new quality of data fixation and processing and to be a basis for digital interaction and formation of digital registries. Therefore, revision, formalization and digitalization of existed inventory procedures got started. Within the first phase of development NID basics, the key procedures in the area of heritage objects, which became transformation subject, were pointed out.

Museum inventory

In particular, in such basic list (core) of museum sector following procedures were included:

- Acceptance of museum significance object for temporary storage (as potential acquisition)
- Primary inventory of museum object
- Inventory management of museum object
- Special inventory management of museum object
- Acceptance the museum object for temporary storage
- Transfer of museum object for temporary storage
- Availability and condition of museum object check
- Withdrawal from inventory list

During 2014–2016 these procedures were revised in the context of digitalization needs and NID development. Modeling and formal presentation of their renewed versions were done. In particular, working processes diagrams, which reflect clear sequence of steps, subjects (counterparties) and their responsibility areas, connected data sets and changing their state etc. were created according to BPMN standard (Object Management Group n.d.).

By the way, in 2009–2011 there was an attempt to reform museum inventory procedures in Ukraine according to SPECTRUM standard. The initiative belonged to Ukrainian Centre for museum development. The appropriate agreement about the document localization between the Centre and Collection Trust was achieved in 2009. During the next two years, the preliminary translation of SPECTRUM (version 3) in Ukrainian was developed. However, the passivity of the museum community, lack of political will and changing the government afterward restrained implementation of this initiative. Only now, in the context of new attempt to formalize inventory procedures there is a chance to use in a certain way gained then experience (although it would be too soon to speak about implementation of SPECTRUM in Ukraine now).

Immovable heritage inventory

Similar work for formalization and digitalization of inventory procedures was started in the area of immovable heritage in 2017. To the basic core of procedures, which have a key character from the heritage object “life circle” fixation point of view, following procedures were attributed:

- Declaration;
- Reporting about an object, which potentially fit the criteria of value;
- Data verification (expert’s inference);
- Preliminary inventory;
- Permanent inventory;
- Monitoring procedures (routine, special, extraordinary);
- Withdrawal from inventory list.

These procedures are distributed between two scenario lines. Basic digital scenario (see green bloc at Fig. 1) describes consecutive inventory steps from detection of an object, which potentially may fit certain value criteria, to giving it the status of heritage object and including it to the corresponding list. The case foresees information to check and progressive forming of necessary data sets by authorized persons (experts). In particular, clear determination of the protection object, the boundaries of the object territory, the boundaries of the protection sectors, fixing the restrictions of economic activity and owner’s protection obligations have to become crucially important for inclusion the heritage object into the list of permanent inventory.

However, the heritage object, which is already registered, but still have information in old, paper documentation, demand special procedures to be included into renovated digital register. Foremost it is due to the need to verify the available information and to extend it according to modern data model requirements. For this purpose, the special procedure Declaration is foreseen (see red bloc at Fig. 1).

Renewed process map also foresees clearer regulation of monitoring procedures. The period of regular monitoring should be no more than 5 years. Out of turn monitoring may be carried out if necessary, in particular in the context of certain types of economic activities or at the request of local public activists. The result of monitoring should be the source for updating data in the system. In exceptional cases, according to the monitoring results, a decision may be made to exclude an object from the register.

Along with the regular and extraordinary verifications, a special monitoring of transitional period was foreseen. The subject of the verification in this monitoring is not so much the condition of the heritage object, as its compliance with modern criteria. Because since the soviet time a large number of objects gained heritage object status formally due to ideological reasons with the aim to soviet myths affirmation and materialization. Therefore, even after twenty-seven years of Ukrainian independence some phantoms of the past age, whose status has to be redefined, remains in the list of heritage objects.

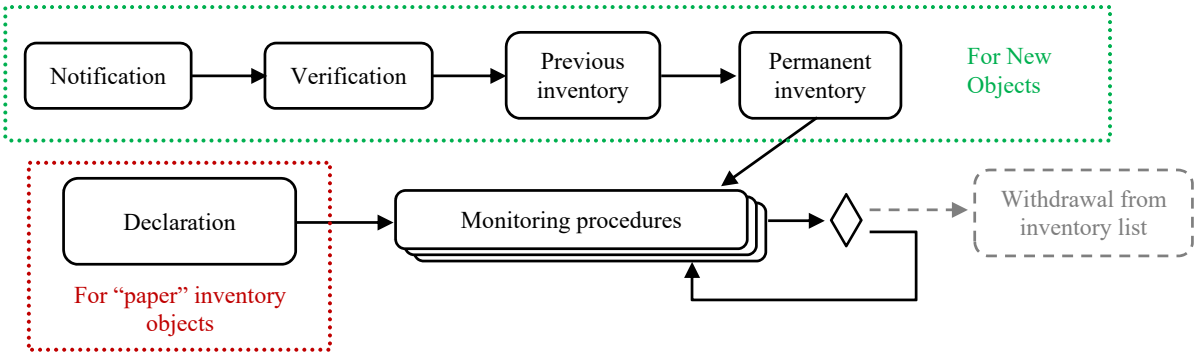

Fig. 1. General process model of immovable heritage objects inventory

So processes maps and schemas of digital interaction at the sectoral levels as well as in the context of shaping the general cultural heritage resource were shaped in the area of movable and immovable heritage based on the results of formalization procedures.

2.3 Data

Data formalization and unification became other important tasks in NID development. Paper documentation approved by appropriate institutions remained the only means for information ordering in the cultural heritage of Ukraine protection for a long time[1].

This situation obviously does not match actual needs of heritage objects digital inventory implementation. However, direct, literal implementation of already existed metadata standards in Ukraine also is not possible. After all such applicable standards as SPECTRUM XML schema and MIDAS Heritage were created in different context.

Considering the peculiarities of Ukrainian realities, the decision about complex of national metadata schemas creation was taken. It had to count the existent traditions of registering, to provide modernization processes and to be compatible with formal CIDOC CRM model at the same time.

Following steps were carried out for this:

1) Check points, in which certain incoming or outgoing data sets have to appear were defined within formalized inventory procedures scenarios and digital interaction processes. Requirements for mandatory

[1] The only exception was the library sector, where the family of MARC standards became widespread.

attributes (attribute groups), their entries quantity (elements examples), content features etc. were defined for every such case.

2) Every attribute as well as relation between the attributes, objects, historic events and inventory processes were described via CIDOC CRM classes and properties. It allowed to expand traditional data model and to apply eventocentric manner of object information organization. For example, the “Author (producer)” was transformed into set of elements that enable fixation of connected persons and their roles in certain historic events context.

3) At the base of obtained formalized model the derivative xml-schemes for application in context of specific processes and data exchange were developed. As the result of this work in particular, the national metadata schemes for description of the museum objects (CODE) and immovable heritage objects (HED) arose.

Every metadata scheme was constructed as connected XSD-files complex. The upper-level file contains description of user’s types and common elements. Further, these structures are sued in derivative XSD-files, created to serve specific inventory processes. They have different configuration of specific attributes, but unified general structure, in which these attributes are inscribed. Overall, their structure was modeled in such way, which provides fixing data in the context of inventory events, at the intersection of conventional lines “Object” and “Procedure” (see Fig. 2.).

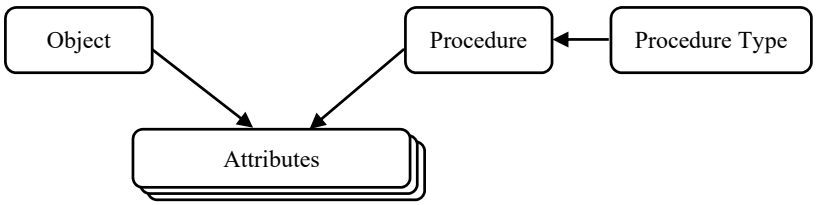

Fig. 2. Eventocentric model of object data fixation

General data packet structure in hierarchic representation looks as following (see Fig. 3).

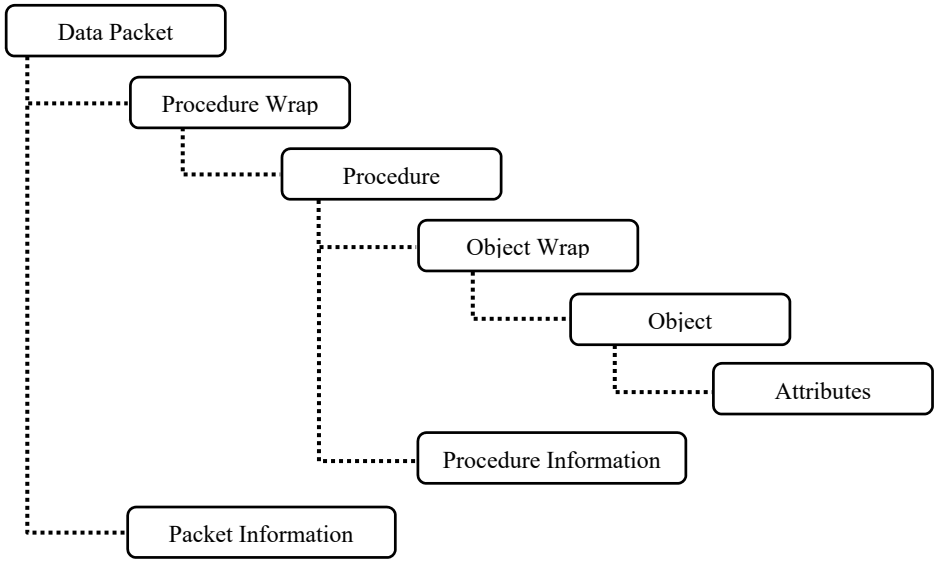

Fig. 3. General structure of XML-packet for data exchange in the context of inventory procedures realization

Such approach enables the cross-compatibility of data packets without relation to procedure type. It simplifies documenting the history of knowledge on objects, allows distinguishing separate data layers and

tracking their changes, which is important from heritage objects administration and collections management points of view.

Since the development of applied XML-schemes CODE and HED was based on data models that is represented via CIDOC CRM classes and properties, it provided in fact mapping CODE/HED – CIDOC CRM. CIDOC CRM orientation also causes the national schemes compatibility with the relevant international standards. In particular, in 2018 the development of CODE–LIDO and HED–CARARRE transformation templates started.

2.4 Thesauruses

Systematization of terminology for cultural heritage objects description, creation of national unified digital thesauruses and authoritative files started with the aim of data unification. However, this part of works, unfortunately, moves quite slowly. It is extremely important to transform the professional museum and heritage protecting community into active part of this process. Ukrainian center for museum development with the support of Ukrainian culture foundation in 2018 initiated with this aim development of web-platform for collaboration in creation of composite National thesauruses and reputably files and its publishing. Comparison thesauruses in Ukrainian with similar dictionaries in English, foremost the Getti foundation is the important task too.

2.5 System architecture

By now the general architecture concept of generation, aggregation and presentation system for cultural heritage objects information was developed and construction of its separate blocks was started.

Museum bloc

Museum institutions is obviously the main source of museum objects data. Highly trusted data have to be generated as the result of inventory procedures and museum collections digitization.

New game rules, metadata standardization and renewed base of museum inventory regulation in fact have laid basics for new market of specialized museum software in Ukraine. Considering a great number of museums in Ukraine and their diversity (of profiles, subordinations, ownerships, structures, scales, etc.), open market for specialized software instruments looks as the best solution for museum field. The unified single state museum inventory software system will hardly be effective and equally convenient for everyone. So general architecture of museum bloc (see Fig. 4) enables the usage of any software for collection description and inventory, which fit the general demands, and supports data exchange in unified format (CODE).

National register supposed to aggregate a part of most important data about museum collections. It has to obtain data from museum systems as the result of passing the clearly defined inventory procedure set. This register is considered as element of digital government system. Special API and protocols have to provide data exchange with other registers and departments (for example, with custom bodies to control the moving of cultural values through the States border of Ukraine). To provide this communication of digital state resources Trembita interoperability system based on Estonian X-road was created.

Data presentation public portal is constructed for providing the wide audience open access to information on museum collections.

Thesaurus portal in the frames of museum bloc is considered as separate element that have to be taken care by professional museum community.

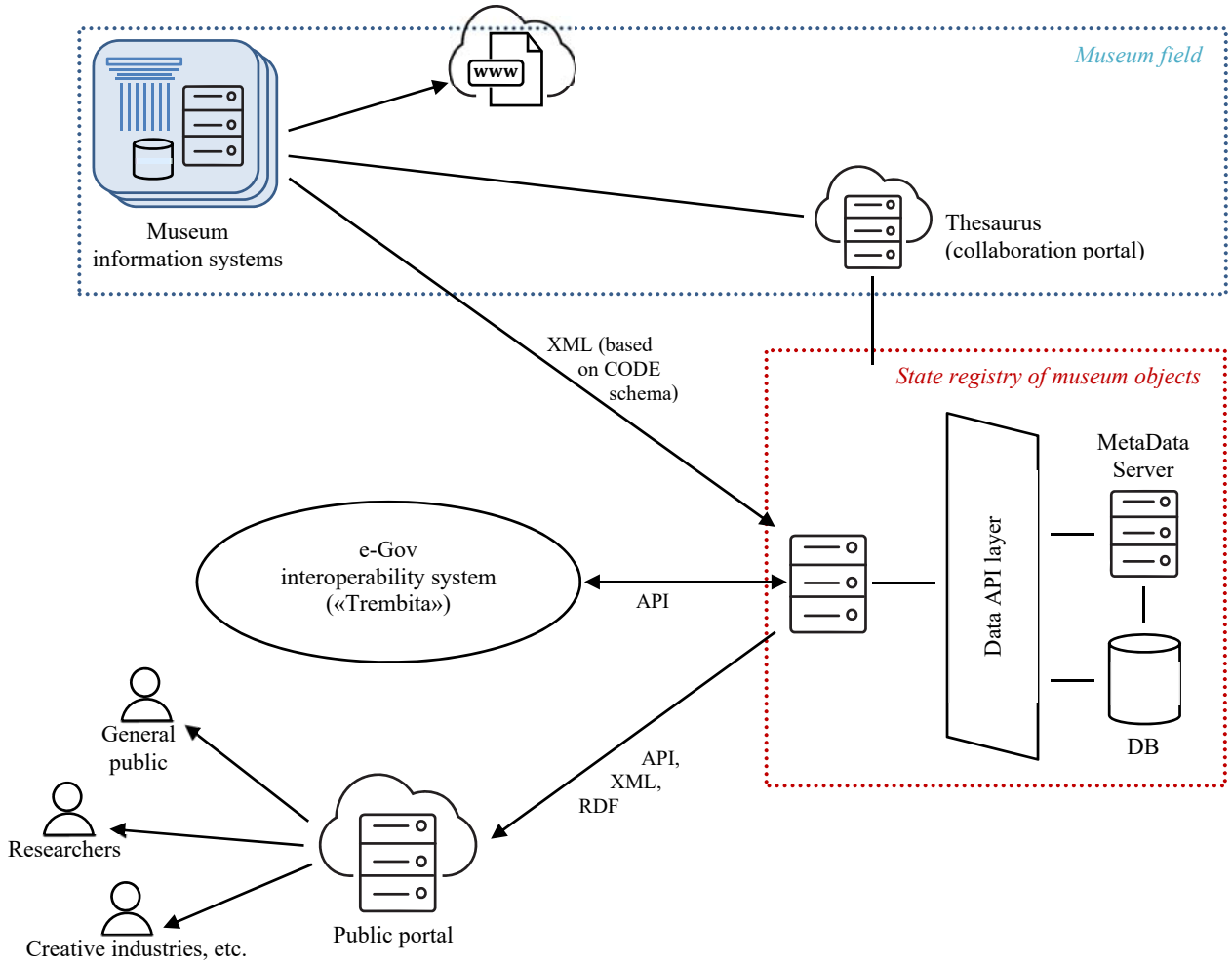

Fig. 4. Conceptual model of museum bloc

Immovable heritage bloc

Unlike the museum sphere, immovable heritage object protection and management in Ukraine tends to be more centralized. The main work of heritage objects inventory is relied on special protection bodies regional and local level.

These peculiarities of the sector management organization cannot but be reflected on vision of system's certain bloc architecture. In particular, the creation of centralized portal with limited authorized access to provide the inventory processes and forming the digital data sets is foreseen (see Inventory subsystem at Fig 5). Thesauruses forming and development for cultural heritage object description at present also assumed as function of cultural heritage protection central body.

At the same time, we seek to involve an active part of the local community in the task of identifying and protecting cultural heritage. Therefore, within the framework of the function of the public portal, it is foreseen to send a notification about a new object that potentially can meet the criteria of value (deserves to be entered into the register of a real estate), as well as the current state of the monument to the competent heritage protection body. According to the results of such messages, an appropriate accounting process can be initiated.

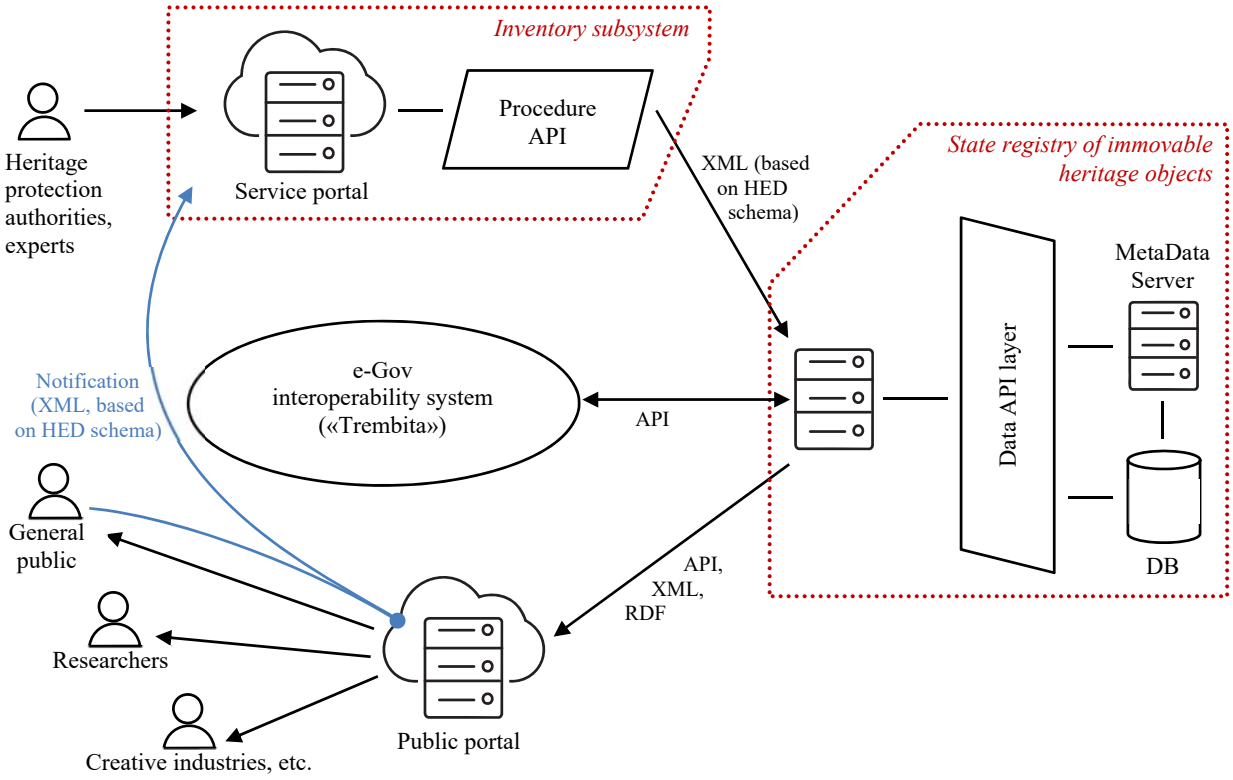

Fig. 5. Conceptual model of immovable heritage bloc

General data aggregation bloc

Data, collected from museums, immovable heritage security bodies, libraries, archives, etc., generally aggregated in one bloc, has to be the upper level of general system. This aggregation is expected to be done on the base of Linked Open Data. Data elements from branch registries marked by URI and described with RDF-triplets according to CIDOC CRM ontology will be represented on special cultural heritage consolidated resource (see Fig. 6).

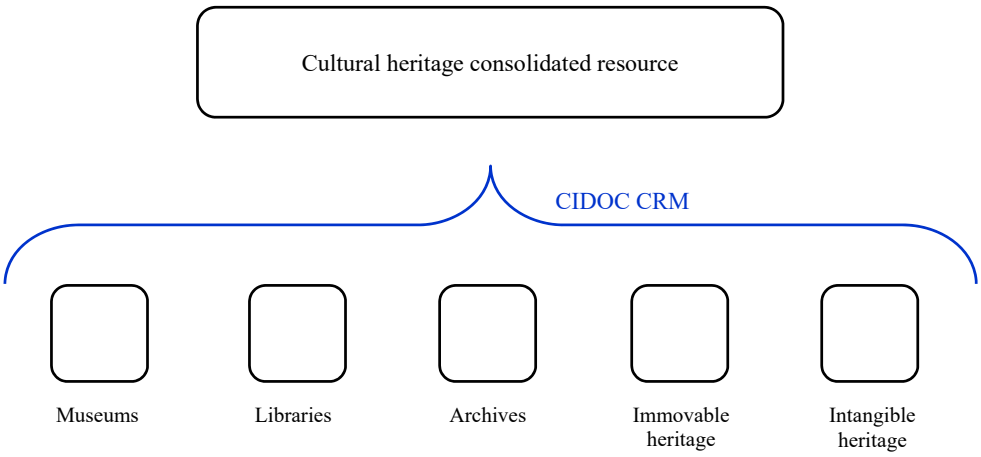

Fig. 6. Conceptual model of general data aggregation bloc

3. Intangible heritage

Digitalization problem in the area of intangible heritage deserves separate attention.

In 2008, Ukraine joined the Convention about protection of intangible cultural heritage. In became a powerful impulse for forming the inventory of intangible cultural heritage in Ukraine states system. In particular, it foresees the creation of corresponding digital registry. However, how exactly it has to be done remains the subject for discussion. Foremost, there is no common vision concerning inventory subject and necessary system detailing level.

Now the approach, according to which the registry reduces only to the level of data aggregation, prevails. From this point of view, the subject of inventory is in fact just artificial abstraction: generalized description of certain community practices, product of researches and interpretation. This is quite enough for representative lists formation, in particular, with the aim of their translation to the UNESCO level. However, such an approach means certain loss of detailing.

Alternative position anticipates architecture of the registry that has to include as aggregation as fixation of source material levels. This additional level will allow expanding the inventory subject understanding to traditions and practices specific instances reconstruction. It has to provide the preservation of digital medium (e.g. audio and video recordings, texts etc.) and accompanying metadata. Considering the specifics of the intangible cultural heritage process-oriented data organization model seems appropriate for such level (see, for example: Kettula and Hyvönen 2012). This approach will allow better recording of intangible heritage variability and versatility, tracking the links between objects, persons (communities), regions etc. in dynamics.

It has to be mentioned, that positive experience, that can be base for such expanded concept of intangible cultural heritage inventory system creation in Ukraine already exists. It is projects like “Bervy” (“Проект Берви. Українська автентика та фольклор” n.d.) and “Polyphony” (“Polyphony Project” n.d), which were dedicated to digitalization, systematization, meta-description and presentation of broad massif of folklore. Fixation of songs and rituals version with specification of extra data (connected persons, spatial data etc.) and relations between them became good practice.

4. Results

During 2014 – first half of 2018 the work focused mostly at the conceptual and application level. It concerned future data infrastructure and its separate element modeling. The formation of regulatory and methodological base was initiated. In particular, the Procedure of digital museum objects inventory was approved (Міністерство культури України 2016). Changes for regulations in the area of immovable cultural heritage protection were prepared. It became the ground for beginning the software and hardware instruments of different level creation. Series of pilot projects dedicated to the digital inventory of museum collection implementation were initiated in 2017-2018. The development of the platform for digital interaction in cultural heritage protection has begun. Software interfaces for data processing and transformation, presentation modules etc. are created.

This is complex, fundamental stage and its realization requires considerable resources and wide range of specialists coordinated activity.

However, the transition to the operational level, where “rules of the game” and instruments finally have to work, where data are generated and applied in daily work of protection bodies, museums, reservations etc. is also significant problem for the country.

The actors’ level of readiness for changes, their ability to realize new approaches and rules at the practice is the main challenge in this context. Ukrainian specialists, museum and heritage protection workers are missing relevant knowledge and competency very much. Universities do not offer necessary disciplines yet.

In this connection, Ukrainian center for museum development creates and implements training programs for professionals, organizes thematic seminars and workshops. The Center together with national ICOM committee and other partners are making efforts for development of professional network in Ukraine, formation of special working groups for solving hot issues of heritage digitalization.

We hope that our collaboration and integration with international professional community will become an important factor for the viability and initiated processes stable development, allow to provide a new quality of work with data in Ukraine in the future.

Bibliography

Кабінет Міністрів України. 2017. *Середньостроковий план пріоритетних дій Уряду до 2020 року та плани пріоритетних дій Уряду*. April 3. https://www.kmu.gov.ua/ua/diyalnist/programa-diyalnosti-uryadu/serednostrokovij-plan-prioritetnih-dij-uryadu-do-2020-roku-ta-plan-prioritetnih-dij-uryadu-na-2017-rik.

Міністерство культури України. 2016. *Порядок обліку музейних предметів в електронній формі.* September 9.

Міністерство культури України. 2017. *Концепція впровадження електронного обліку і презентації об'єктів культурної спадщини та культурних цінностей на 2017–2020 роки.* November 11.

"Проект Берви. Українська автентика та фольклор". n.d. Accessed August 1, 2018. http://bervy.org.ua.

"Реєстр нерухомих пам'яток України". n.d. Accessed August 5, 2018. http://mincult.kmu.gov.ua/control/uk/publish/officialcategory?cat_id=244910406.

Kettula, Suvi and Eero Hyvönen. 2012. "Process-centric Cataloguing of Intangible Cultural Heritage" Accessed August 1, 2018. https://seco.cs.aalto.fi/publications/2012/kettula-hyvonen-intangible-cidoc-2012.pdf

Object Management Group. n.d. "Business Process Model and Notation Specification". Accessed August 12, 2018. https://www.omg.org/spec/BPMN

"Polyphony Project". n.d. Accessed August 1, 2018. http://www.polyphonyproject.com/en.