Załącznik nr 1.3 do SIWZ/Opisu Przedmiotu Zamówienia **(ZMIENIONY 2017-07-07)**

Description of the Object of the Contract

e-services OMNIS

Task 1.5: Implementation of the metadata hub - purchase of metadata hub with licenses

Contents

[1. INFORMATION ABOUT THE PROJECT 2](#_Toc483476632)

[2. GENERAL DESCRIPTION OF THE OBJECT OF THE CONTRACT 3](#_Toc483476633)

[2.1. Introduction 3](#_Toc483476634)

[2.2. General Arrangement 7](#_Toc483476638)

[2.3. Implementation methodology 14](#_Toc483476639)

[3. SPECIFICATIONS OF THE OBJECT OF THE CONTRACT 16](#_Toc483476640)

[3.1. Introduction 16](#_Toc483476641)

[3.2. Configuration of and feeding the Multibrowser Omnis 17](#_Toc483476642)

[3.3. Feeding in full-text data 22](#_Toc483476643)

[3.4. Sandbox 23](#_Toc483476644)

[3.5. Converters 25](#_Toc483476645)

[3.6. Record life cycle 26](#_Toc483476646)

[3.7. Multibrowser 27](#_Toc483476647)

[3.8. Data model 43](#_Toc483476648)

[3.9. OMNIS Linked Data 44](#_Toc483476649)

[3.10. UID/URI 47](#_Toc483476650)

[3.11. Data access APIs 48](#_Toc483476651)

[4. AUTHORIZATION 53](#_Toc483476652)

[5. ADMINISTRATIVE INTERFACE 54](#_Toc483476653)

[6. TESTS 55](#_Toc483476654)

[7. SCHEDULE 56](#_Toc483476655)

[8. DOCUMENTATION 58](#_Toc483476656)

[9. GUARANTEE AND WARRANTY 59](#_Toc483476657)

[10. COPYRIGHTS 60](#_Toc483476658)

[11. ESTIMATED DEMAND FOR THE SERVICE OMNIS 60](#_Toc483476659)

[12. PERFORMANCE AND SCALABILITY 62](#_Toc483476660)

[13. GLOSSARY 63](#_Toc483476661)

[14. Diagram translations 65](#_Toc483476662)

[Attachements 66](#_Toc483476663)

# INFORMATION ABOUT THE PROJECT

This object of the contract is the first of the four tasks specified in the project titled “OMNIS e-service” carried out by the National Library (hereinafter referred to as BN) within the Digital Poland Operational Program for the years 2014-2020, Priority Axis no. 2 “E-administration and open government”, Measure 2.1 “High availability and quality of public e-services”, co-financed within the Digital Poland Operational Program by the European Regional Development Fund and the State Budget.

The basic activities within the project include implementation of 4 separate and integrated   
e-services at the highest, i.e. the fifth degree of maturity.

1. Integrated Multibrowser OMNIS (A2C type service),

2. e-ISBN publishing repository (A2B type service),

3. Cloud-based service POLONA w Chmurze for libraries (A2C type service),

4. Cloud-based service POLONA w Chmurze for scientists (A2C type service),

# GENERAL DESCRIPTION OF THE OBJECT OF THE CONTRACT

## Introduction

The object of the contract is the construction and implementation of the information system for the e-service of Integrated Multibrowser OMNIS under the name of “OMNIS Metadata Hub”. Hereinafter, the following terms are used interchangeably: Integrated Browser, Multibrowser or OMNIS Browser for the e-service of integrated multibrowser OMNIS and the Metadata Hub for the OMNIS Metadata Hub.

The main recipients of the e-service will be primarily individual users (students of various education levels including universities, employees of academic institutions, teachers, librarians), who through the service will be able to search nationwide library catalogs for the desired book, magazine, journal, article, map, graphics, photography, manuscript or any other type of document as well as get access to inventories of selected archives and museums. The browser shall also provide the full-text searching of books, magazines and articles in digital form to enable a high quality exploration of library information resources. Potentially, there may be 5.9 million recipients of the service. The Contractor shall ensure efficient operation of the system with the above number of users, as described in detail in Section 11. ESTIMATED DEMAND FOR THE SERVICE.

The construction and implementation of the hub are analytical tasks involving development of conversion tables that are the object of other public procurements, work on data modeling, and system testing on the user’s side that shall be carried out by the Contracting Authority and is not the Contractor’s task. It is also assumed that the hub will be connected to further data sources right from the very beginning of the implementation and after its completion. The general time perspective for the Object of the Contract spans from the date of entering into the contract with the Contractor selected under the open public procurement procedure to 2018-09.

As a result of the Contractor’s work, an e-service shall be created whereby a single access point will be provided for searching through metadata/bibliographic descriptions and content of publications available on the market as well as those at the stage of preparation for publishing and those collected in resources of all Polish libraries. As for publications planned for publishing, it will be possible to launch the e-service owing to integration of the e-ISBN service already available at the National Library in such a way that data regarding publications that publishers transmit upon obtaining an ISBN number are available in the Omnis system browser. The Contractor shall carry out the above-mentioned integration with e-ISBN service expanded with the functionality that allows publishers to provide the National Library system with an electronic version of the publication under development and, additionally, offer the possibility of full-text search using indexing of its content in the OMNIS system.

The estimated demand for services offered by the Integrated Browser OMNIS in the first place takes into account the readership level in Poland and number of people with access to the internet, which is the key condition for using the e-service, as well as results obtained in the study on the Omnis e-service potential that was carried out by the National Library in March, 2016 (study titled: Potential of the new e-service dedicated to individual users). The results of the demand analysis show that the projected number of users will be 3.07 million in 2025 with the total number of activations of the e-service amounting to 5.6 million in 2025.

As far as programming work is concerned, the Contractor shall closely cooperate with specialists from the National Library within the period specified in section Schedule. Details of tasks covered by the cooperation are explained in relevant sections. The above-mentioned programming work aims at implementing the OMNIS metadata hub and the Multibrowser based on this hub (hereinafter referred to as the Integrated Browser, Omnis Browser or Browser) together with the Proxima mobile interface, that involves the following activities with the time-frames specified in the section Schedule:

1. Designing and implementing by the Contractor of the infrastructure configuration with interoperability of the hub and the BN systems (Sierra/Millennium, Academica, Polona, e-ISBN) and external data sources taken into account;
2. The Contractor shall carry out analytical work consisting in building an analytical model of the system embracing such elements as specification of indices that have to be built, specification of Actors and Use Cases, and elements of the General Arrangement of the Contract.
3. Designing and implementing, by the Contractor, software for the Metadata Hub with interoperability of the Hub and BN systems (Sierra/Millennium, Academica, Polona, e-ISBN) and external data sources taken into account;
4. Allowing specialists from the National Library to evaluate and modify conversion tables in line with results of analysis of library data exchange protocols and metadata formats used in Poland, i.e. at least such as ProLib, Mak+, Virtua, Dlibra, Aleph, and metadata formats such as Marc21, Dublin Core, and ONIX. The analysis made by the Contracting Authority is not the object of the contract in question and its result will be provided to the Contractor in accordance with the Schedule and then, based on it, the Contractor shall implement an adequate configuration so that its evaluation and modification is possible at the test stages.
5. Based on information provided by specialists from the National Library, implementing adequate modifications in the configuration in the course of developing conversion tables for metadata formats necessary for data exchange (metadata formats such as Marc21, Dublin Core, and ONIX). As a result of cooperation between the Contractor and Contracting Authority, there will be created conversion tables offering the possibility of further evaluation and allowing free exchange of information and switching of data formats, which is necessary in the context of tests and each subsequent introduction of new data formats.
6. The Contractor shall develop the main data set OMNIS and rules for data display on the basis of previously built conversion tables;
7. The Contractor shall develop the OMNIS Hub mechanism within the scopes of its three individual modules, as presented in Section 2.2 “General arrangement” and Section 3. “Specifications of the Object of the Contract”, with their interactions taken into account.
8. The Contractor shall develop and implement an OMNIS browser for stationary devices;
9. The Contractor shall develop and implement an OMNIS browser for the Proxima mobile devices;
10. The Contractor shall ensure connection of BN data sources to the Hub and a similar test connection for selected external data sources.
11. While working on interfaces for the Multibrowser and its mobile version “Proxima”, the Contractor shall take into consideration the GUI/UX solutions prepared by the Contracting Authority under the task “Design of UX access interfaces” and allow verifying them by the Contractor so that the final implementation of the Multibrowser Omnis takes into account components necessary for its operation that have been developed under the above-mentioned task and makes it possible for the Contractor to implement them at a later date. The task “Design of UX access interfaces” is outside the described object of the contract and its result will be provided to the Contractor in line with the Schedule.
12. The Contractor shall provide browser interfaces based on the design prepared by the Contracting Authority and delivered to the Contractor in line with the Schedule.
13. Granting a warranty and providing warranty services on terms specified in Section “Warranty”.

Software of the Multibrowser Omnis and Hub shall operate in a virtualized environment based on multi-core servers of a standard design, a SAN with a disc array and tape library, and a 10 Gb Ethernet LAN. The ordering party informs that they use open source system tools that will run on the free Citrix XenServer platform. Preferred systems for this platform are CentOS 7.0 or higher, Debian Wheezy 7.0 or later, Ubuntu 12.04 or later. Preferred databases are: Microsoft SQL Server Express, PostgreSQL, MariaDB, Apache Derby. As an essential requirement, it is necessary to ensure data replication and availability of the basic system functions with a Backup Centre found at another location. The functioning of the Multibrowser is regarded as the primary function of the system. That is, in the event of a failure, there shall be a duplicate version of the main module (See Diagram 1) with the frozen state of the backup copy in the backup center. Copies will be created once per 24h and should include a fully functional, searchable startup module. For full-text and semantic modules there shall be no replication but only once-daily backups shall be run.



## General Arrangement

The Contractor shall build an information retrieval Hub comprising processes of metadata acquiring, processing, correlating and sharing, as necessary for activating the e-service of Integrated Browser OMNIS and handling connections with other systems (including other OMNIS services) that use data exchange protocols and are based on metadata found in the Hub. The work includes:

1. Integrating access to metadata of all library collections owned by BN,
2. Integrating access to metadata of publications owned by other libraries,
3. Integrating access to metadata of forthcoming publications and *in commerce* publications from the e-ISBN service,
4. Integrating full-text searching of digital items from resources of the National Library and Polish libraries,
5. Integrating searching of texts generated within the cloud-based Polona w Chmurze service for scientists,
6. Integrating searching of resources of selected cultural institutions that make their resources available over the OAI-PMH protocol.

The system shall support the following data sources with tools for data acquisition described in Sections 3.1-3.3:

1. Electronic library catalog including the National Library catalog, so far based on the Sierra System (Innovative Interfaces Inc.), and catalogs based on the systems Alma (ExLibris), Aleph (ExLibris), Voyager (ExLibris), Prolib (Grupa Sygnity), Mak+ (Instytut Książki), Virtua (Innovative Interfces Inc.), Horizon (SyrsiDynix), Patron (MOL), Libra (MOL), Sowa (Sokrates Software), Mateusz, Koha, Co-Liber.
2. Digital library polona.pl and other Polish digital libraries insofar as integration of their data is possible, as well as the cloud-based service Polona w Chmurze that is under development within another task as a separate procurement. In case of data found in the digital library polona.pl and the emerging service Polona w Chmurze for libraries, data are and will be available within the open data set that is accessible through open APIs, as available at https://polona.pl/api/ (<https://danepubliczne.gov.pl/dataset/polona-pl>).
3. e-ISBN service and publishing repository
4. The Polona w Chmurze service for scientists, accessible as the Polona Scholar tab of the polona.pl digital library.
5. Inventories and catalogs of archives and museums in Poland, provided that their resources are accessible over the OAI-PMH protocol.

The system shall offer the following functions:

1. Downloading and displaying metadata from libraries of all types found in Poland, as specified by the Contractor, with maintaining information about the source.
2. Downloading and displaying metadata from different library systems (Sierra (Innovative Interfaces Inc.), Alma (ExLibris), Aleph (ExLibris), Voyager (ExLibris), Prolib (Grupa Sygnity), Mak+ (Instytut Książki), Virtua (Innovative Interfces Inc.), Horizon (SyrsiDynix), Patron (MOL), Libra (MOL), Sowa (Sokrates Software), Mateusz, Koha, Co-Liber),
3. Downloading and displaying information about availability of a given copy,
4. Downloading and displaying information about the location of the library,
5. Simple and advanced searches with definable faceted filters, based on all above-mentioned data sources,
6. Full-text search within digital publications with contextual display of phrases,
7. Access to the bibliographic description with a possibility to export it in a selected format for librarians (exchangeable formats, MARC, ISO) and scientists (BibTex, EndNote, Zotero, etc.).
8. Referring to the full text in case of items in the public domain or licensed items (using the interface of the digital library Polona),
9. Referring to the booking site in the Academica system, i.e. a service currently offered by the National Library at <https://academica.edu.pl/>, which address may be changed as a result of the current development work on the system and if this is the case, it will be provided by the Contracting Authority,
10. Presenting availability of collections at the nearest libraries, depending on the user’s location (location services).
11. Mobile access to services of the Multibrowser OMNIS,
12. Scalable, definable and automatic clustering of bibliographic records at the level of copies, realization and embodiment based on the model built on assumptions of the FRBR/LRM model as developed by the Contracting Authority, on the assumption that its evaluation and modifications will be carried out in cooperation with the Contractor.
13. Supporting own, universal and scalable metadata model, hereinafter referred to as the OMNIS data model, based on mapping from various formats, particularly MARC21 (https://www.loc.gov/marc/), Onix (http://www.editeur.org/8/ONIX/) and Dublin Core (http://dublincore.org/specifications/),
14. Open APIs allowing third parties to reuse metadata,
15. Open APIs offering the possibility to carry out projects covering metadata semantization and integration of data from diverse sources, including external sources available in the Linked Open Data Cloud.
16. Graphical user interface compatible with at least five leading Internet browsers,
17. The accessibility service at the presentation end of the Multibrowser meets AA level requirements of WCAG 2.0,
18. Publishing a defined data set of the Hub as Linked Open Data as well as the possibility to use the Linked Open Data option to publish selected data sub-sets from individual Hub modules.

The Contracting Authority expects that system will have the following non-functional features:

1. Performance sufficient to process a set of min. 20 million records with defined response times at the project completion,
2. Reliability and transactionability - a failure or unavailability of any subsystem shall not lead to data corruption,
3. A failure of any single hardware item should not result in the system unavailability exceeding 2 minutes,
4. 24/7 access to the service and functions of the Multibrowser (with taking into account all assumed regular data updates, which should not affect the access to the service or noticeably reduce its performance),
5. Providing full documentation of the software (system user and administrator manual,   
   architecture documentation, developer documentation, implementation documentation, etc.), as described in detail in section “DOCUMENTATION”.

According to the general assumption, the task that is the object of the contract shall be divided into separately operating elements. Each element should be capable of operating on its own on the assumption that the full functionality of the Hub is attained only when all elements operate together. The main OMNIS metadata Hub module shall receive data from selected external institutions, in particular at the level of MARC21 (https://www.loc.gov/marc/), Onix (http://www.editeur.org/8/ONIX/) and Dublin Core (http://dublincore.org/specifications/). It is assumed that data will be uniformed by converting them to a common format compatible with the OMNIS data model that will be used for storing metadata. The OMNIS data model shall be based on the implementation of the DBN (Deskryptory Biblioteki Narodowej / National Library Descriptors - see section 3.8 “Data Model”) where formal ontology will be used in order to ensure semantic integration of individual metadata sets acquired by the Hub. The Contractor shall provide the possibility of saving source data in their original format and then processing them within the Hub and ensure the possibility of updating data at any time. The OMNIS common data model assumes the possibility of its continuous development.

The second part of the Hub mechanism, named the LOD semantic module, will be built using the main index and the determined semantic mechanisms, including conversion tables and mapping to the RDF format. The Contractor shall take into account the mechanism of data flow that exists between the main and semantic modules as presented in section “Specifications of the object of the contract” The Contractor shall ensure that performance of the LOD semantic module is sufficient for it to be used together with the on-line Multibrowser and through third-party services integrated with the Hub at the API level. The Contractor allows creating a static image of operations carried out by the semantic module and making it available as needed for such services where the module will be capable of returning results of its operation for further use in predetermined (daily) intervals. It shall be possible to use such results in LOD-based applications (including elements of the Multibrowser OMNIS).

The Contractor shall ensure that the LOD semantic module is capable of re-integrating data from the main module and third-party services in predefined (daily) intervals.

The last part of the Hub shall be the Index of full-text data FULLTXT (the FULLTXT module) that is to be linked to the main index at a level of unambiguous identification of the object and integration of descriptive metadata in line with the OMNIS model. At the same time, data acquisition and the very structure of the model may form a completely independent mechanism. The FULLTXT shall be an effective module where full-text data acquired from distributed sources is stored and linked at the identification level to the main module. The main module stores both identification data for a given collection and information about all copies of a given item present in all distributed sources. The full-text copy stored in the FULLTXT module shall be perceived as one of many copies available in the main module. The Contractor shall ensure that the operation of the FULTXT module does not interfere with continuous operation of the multibrowser and other modules and at the same time, as part of the integrated search, it is capable of searching full text database (searching for a phrase of the question asked in the multibrowser or the Hub API). The main source for this index shall be an element of the National Library Repository where digitized objects of e-publications are stored for the needs of the cloud-based e-service Polona w Chmurze (Polona dla Bibliotek) including metadata for object identification and full-text data but the Contractor shall take into account other heterogenic data sources on the assumption that they are or will be made accessible.

Such a structure aims at ensuring the highest possible performance with maintaining at the same time full functionality of the system as well as ensuring operation of basic functions if any index is not available. The Contractor shall provide mechanisms for the integrating and sharing of data from individual modules and data exchange between individual modules with the need of sharing and integrating specified subsets (data sets defined together with the Contracting Authority) taken into account. The Contractor shall provide mechanisms for both integrating external data at inputs of individual modules and sharing data from individual modules, also at the level of interfaces of the browser OMNIS.

Irrespective of the above described elements, the Contractor shall provide a mechanism for acquiring data from heterogenic data sources and transferring the same to the main resource of the Hub or the set of full-text sources, or any other mechanism that satisfies relevant technical conditions. The mechanism or set of mechanisms presented in detail in section 3.2 shall be arranged as a separate application that may be launched even outside the environment of the metadata Hub so that, when a proper configuration is used, it is possible to use it for defining a new data source, establishing a connection with it, introducing conversion tables and providing results in a required format. As part of the data Hub, such a device shall allow a preliminary conversion of source data and transferring them for further work on developing a structure in the structure model in the OMNIS data model and formal ontology used for its presentation.

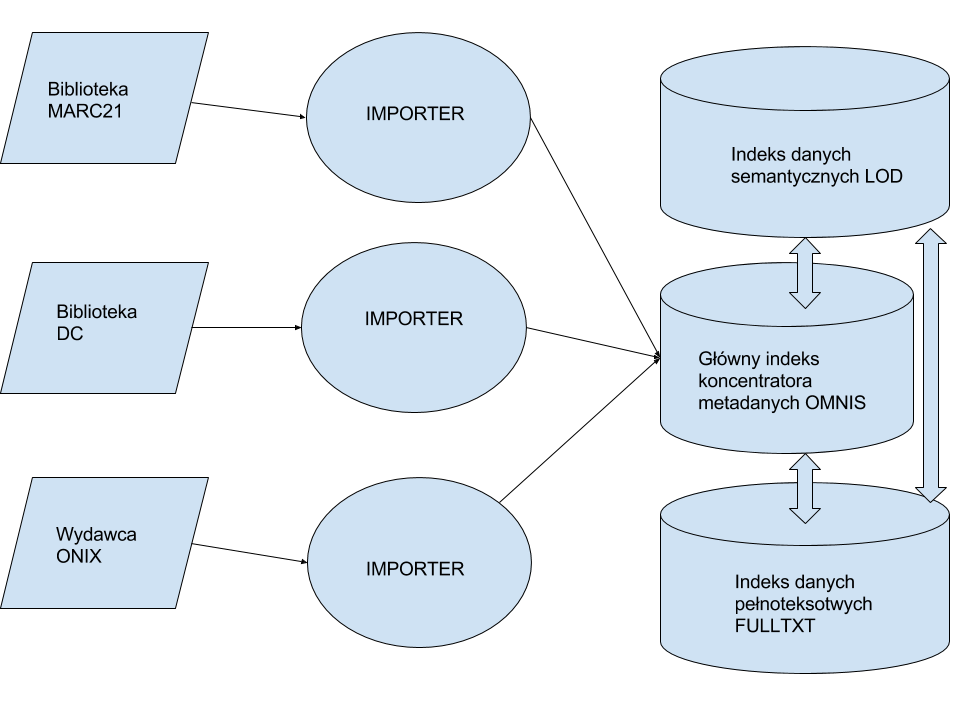


Diagram 1 Block diagram of the metadata Hub structure with three separate indices that exchange data between them.

## Implementation methodology

The object of the contract is the construction of the metadata Hub. The process of the developing and adapting of the existing software shall be carried out using the incremental-iterative model. This methodology assumes that consecutive prototypes of individual system elements will be built and end-users will be actively involved at successive stages of developing the solution, which meets requirements of the standard PN-EN ISO 9241-210:2011 Ergonomics of human-system interaction -- Part 210: Human-centered design for interactive systems. A detailed information about the iterative approach that the Contractor will be required to apply is to be found in section SCHEDULE.

In the programming and design work, the Contractor shall take into account requirements found in the above mentioned standard and PN-EN ISO 9241-151:2008 or equivalent as well as guidelines of WCAG 2.0 or equivalent, that is conveying all indications of their guidelines applicable to the order. It is mandatory to use the WCAG 2.0 and ISO guidelines, or equivalently, that will convey all indications of these guidelines. In the context of entire contract, the Contractor is obliged to use the above mentioned ISO standards, ISO standards listed in the remaining chapters of the contract description, WCAG 2.0 standards or equivalent to all listed ones, that will convey all indications of these guidelines. Wherever in the description of the contract, reference is made to the ISO standards or the WCAG 2.0 guidelines they should be read with words “or equivalent”. In this case, the main focus shall be placed on the system aspect that contains the Multibrowser. A significant impact on the design will come particularly from implementation of the following aspects defined by the standard:

1. basic search and advanced search
2. full-text search
3. displaying detailed information about current search mode
4. resistance to user’s mistakes, prompting of terms
5. sorting of search results
6. explaining criteria used for sorting results
7. scope of search, filters, pagination

The Project provides for the UX tests that additionally shall take into account standard guidelines such as WCAG2.0. The tests do not fall within the scope of the presented object of contract and the Contractor is not required to carry out detailed UX tests. However, it is necessary to take into account recommendations developed within the work on the graphical interface for the Multibrowser and its mobile version Proxima.

The Contractor is also obliged to adapt interfaces and documentation of systems for nomenclature of the eIDAS Regulation and M460 standardization mandate.

# SPECIFICATIONS OF THE OBJECT OF THE CONTRACT

## Introduction

The metadata Hub system shall perform three basic tasks of the data management system, namely:

1. Collecting information consisting in gathering, registering and recording of acquired data and then storing them on permanent data carriers in a form that facilitates using such information for further processing.
2. Processing of information where it is submitted to logical operations related to structurizing in the OMNIS data model (built on assumptions of FRBR and DNB-compliant formal ontology) and integrating additional data from other sources based on integrating the formal ontology from the OMNIS data model with other formal ontologies used in the LOD cloud (including DBPedia Ontology, BIBFRAME Ontology, EDM OWL Ontology, FOAF, dcterms);
3. Sharing acquired and processed information through the Multibrowser with a mobile interface or via an open API.

The basic premise for carrying out these tasks is automation. The user’s role should be limited to specifying a data source and inputting necessary information about it (format, mapping of metadata), and providing configuration settings related to the OMNIS data model. Further data processing and enhancing in the Hub should take place without the user’s involvement.

## Configuration of and feeding the Multibrowser Omnis

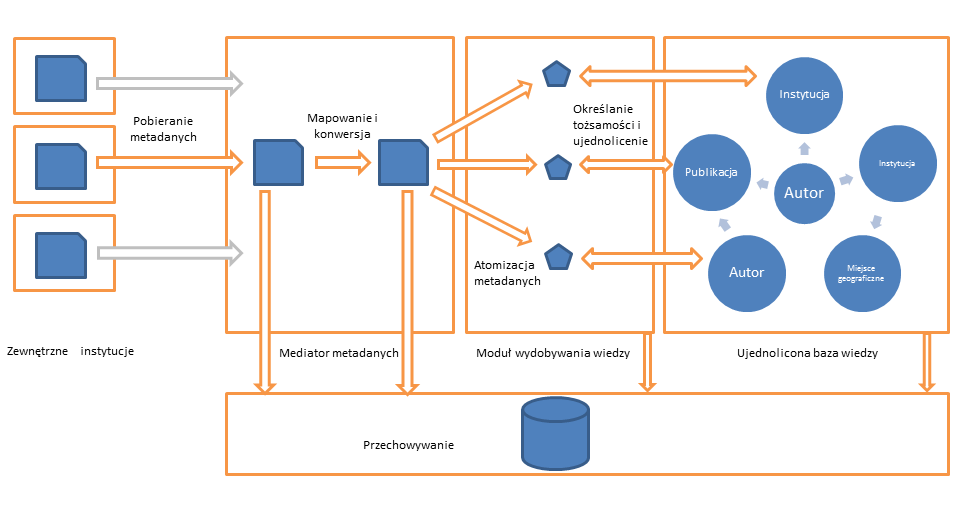


Diagram 2. Block diagram of feeding the metadata Hub.

The Contracting Authority expects that the conversion mechanism, when combined with the importer, will form a solution that is capable of downloading data and transforming them in line with the editable definition of formats (editable conversion table) and offer the possibility of launching it after it is disconnected from the remaining infrastructure of the metadata Hub. The Contractor is required to combine these two elements in such a way that they may be used as a separate application that may be run on a virtual server.

The Contractor shall develop appropriate interfaces for providing the Multibrowser Omnis with all data and metadata from databases available in the BN, including data from full-text indices. The current infrastructure of the National Library employs data sources of different types and therefore this will be an introduction to constructing a distributed data Hub. At this stage, it is possible to define the following sources and mechanisms of potential integration that the Contractor shall take into account in the course of work:

1. catalogs and bibliographic databases (protocol Z39.50, static file of exchange format, API)
2. Electronic library catalog including the National Library catalog, so far based on the Sierra System (Innovative Interfaces Inc.), and catalogs based on the systems Alma (ExLibris), Aleph (ExLibris), Voyager (ExLibris), Prolib (Grupa Sygnity), Mak+ (Instytut Książki), Virtua (Innovative Interfaces Inc.), Horizon (SyrsiDynix), Patron (MOL), Libra (MOL), Sowa (Sokrates Software), Mateusz, Koha, Co-Liber.
3. Resources of National Library’s Digital Repository (API, OAI-PMH)
4. Data made available by the digital library Polona and interlibrary loans system Academica (API, OAI-PMH)
5. The e-ISBN system and commercial databases of emagazines and ebooks.
6. Data from full-text indices

**Acting in collaboration with the Contracting Authority, the Contractor shall develop a mechanism of integration with external metadata sources including library catalogs, bibliographic databases, libraries and digital repositories in institutions cooperating all over Poland and Polish institutions abroad. The Contracting Authority shall provide initial list of such potential data sources and the Contractor shall integrate not more then 15 of them, pointed by Contracting Authority, with the mechanisms of the data Hub in line with section SCHEDULE. Additionally, the Contracting Authority shall closely cooperate with the Contractor to develop necessary mapping and conversion tables for individual data sources and supported metadata formats. For each data source connected in the course of work, the Contractor shall ensure that the system records, based on parameters entered by the data administrator, information about the geographical location of the institution that provided relevant data and it is possible to use such information about location to produce a map of data sources.**

The Omnis server shall carry out automatic and periodic polling of all metadata sources. If new records or records modified after the latest data download are found, a process of copying the data from the source has to be launched. The protocol designed for this communication option is the OAI-PMH standard developed by Open Archives Initiative.

If the OAI-PMH protocol is unavailable for a given data source, the Omnis system shall be provided with a properly configured wrapper for acquiring data from the source and emulating OIA-PMH functionalities.

For sources where it is possible to anonymously check availability of a physical copy of a given publication, such information shall be added to relevant records.

After downloading data records, the server shall convert and atomize metadata and then de-duplicate and identify atomic entities in line with the OMNIS data model based on the DBN concept (e.g. titles, authors, institutions - basic entities of the OMNIS model that are compliant with the structure of the National Library Descriptors: http://przepisy.bn.org.pl/deskryptory/zasady-tworzenia-deskryptorow-bn).

The metadata acquisition system for the main module of the metadata Hub and full-text modules provided by the Contractor shall support two following options of data acquisition and offer the Contracting Authority the possibility of administrative connecting of next data sources.

Method no. 1: “Pull”

In Method 1, it is assumed that the project in question requires active involvement of the Omnis system automatics within the area of searching solutions used in the Integrated Browser Omnis.

When simplified, the procedure employed for expanding information resources of the browser would consist of the following elements:

* + - 1. One-off defining of data sources at cooperating institutions and proper configuring of conversion mechanisms by the service operator
      2. Automatic polling of all data sources by the server. If new records or records modified after the latest data downloading are found, a process of copying the data from the source is launched.
      3. If it is not possible to use the OAI-PMH protocol for a given data source, it is necessary to use (at the server end) a properly configured wrapper that acquires data from the source and emulates functionalities of OAI-PMH or API.

Method no. 2: “Push”

In Method 2 it is assumed that the project in question requires permanent activity on the part of institutions that hold collections (public and scientific libraries connected to the Omnis system) as far as searching via the Integrated Browser Omnis services are concerned.

When simplified, the procedure employed for building information resources of the Browser would consist of the following elements:

* + - 1. The cooperating institutions make their own selections of record ranges to be transferred (new or modified).
      2. Records are exported from local repositories to files with possible conversion along the way
      3. Using a standard protocol such as ftp or http upload, institutions initiate upload of metadata records saved as files to a temporary work space created in the Omnis system (separate for each institution).
      4. The rest of the process follows the “Pull” procedure.

Irrespective of this, the Contracting Party shall ensure that the application is capable of saving the source response as a static cache that may be restored on demand if contact with the data source is lost. If connections employ information exchange protocols, the application periodically establishes a connection with the data source to update data.

Software developed by the Contractor shall be equipped with an importer tool for source metadata capable of being run in a virtual environment for a required number of data sources. For example, for a single type of library systems. The application makes it possible to configure access to the data source with data exchange protocols such as OAI, API Rest, Z39.50 or by exchanging the static file. The data are translated through a prepared metadata mapping sheet, conversion tables (pre-prepared in cooperation with the Contracting Authority (in line with the section SCHEDULE)) for converting data to the pattern compliant with the OMNIS data model that forms basis for the formal ontology, and once they have been processed, application transfers them to the Hub for further processing and enhancing.

The Contractor shall ensure the possibility of further development of the data acquisition module by including the options of defining next data sources, configuring new conversion tables for such new sources as described in section 3.4, and adding them to the Hub even after completion of the Project.

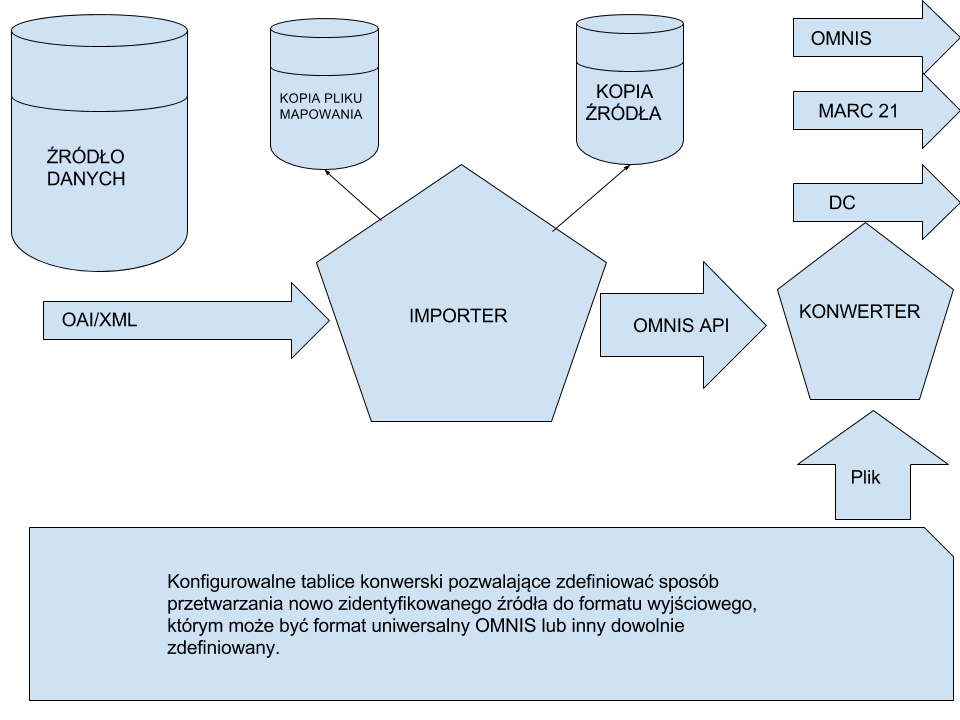


Diagram 3 Functional diagram of the metadata importer

The importer module provided by the Contractor shall be equipped with tools/protocols needed for importing metadata from external sources using the pull and push variants. The minimum requirement from the sources that are not libraries is the ability of downloading data using the OAI-PMH protocol and this protocol has to form basis for operation of the importer. Other offered data download methods (for sources other than cultural institutions that are not libraries, except for the OAI-PMH protocol):

1. Establishing a direct connection with the API of a given data source and this solution should also be ensured by the system.
2. Support for transferring the static exchangeable file in an ISO 2709:1996 format
3. The SRU protocol
4. AtomPub
5. Other as proposed by the Contractor

## Feeding in full-text data

For the full-text data, the Contractor shall equip the metadata Hub with a data download system, whose principle of operation is generally the same with the one described in sub-task 3.2. Considering the data type, the Contracting Authority assumes that such a system may be different from the one described above. This tool shall perform three basic tasks of the data management system, namely:

1. Collecting information consisting in gathering, registering and recording of data and then storing the acquired data on permanent data carriers in a form that facilitates using such information for further processing.
2. Information processing involving logical operations on information and integrating information with other indices of the Metadata Hub.
3. Sharing acquired and processed data through the Multibrowser with a mobile interface or via an open API.

The full-text importer module as developed by the Contractor shall be equipped with tools/protocols needed for importing data from external sources using the pull and push variants. Data shall be downloaded through:

1. A direct connection with the API of a given data source and this solution should also be ensured by the system.
2. Support for transferring the static exchangeable file
3. The SRU protocol
4. AtomPub
5. Other as proposed by the Contractor

It is assumed that the primary source of full-text data will be the National Library repository system, whose current interfaces are described in Attachment 3.

The software for full-text data importer as developed by the Contractor shall be equipped with a tool for indexing the very content of the acquired document and saving it as a static cache in which it will be later possible to search for information using an open API and the Multibrowser. An example of such operation is downloading a \*.pdf file with a text layer and saving this layer alone as a cache for this source. The application shall be ready for handling the following text file formats: pdf, txt, p-pub, mobi, alto xml.

Similarly to task 3.2, the Contractor shall ensure the possibility of further development of the full-text data acquisition module by including the options of defining next sources, configuring relevant access APIs, and adding them to the Hub even after completion of the Project.

Apart from the completely different nature of acquired data, the distinctive element of the full-text data feeding-in module is the need of ensuring information about the data source and transferring them to the metadata indexing module so that a relevant reference to the source is found in this module during the data search. This solution allows highlighting the copies for which a full-text version exists and provides access to the source of such data.

In order to access the source, the user should be referred to the website of the institution of origin of a given object. The metadata Hub is not required to offer direct access to the developed database with full-text content. The exception may be applied to administratively authorized applications of the National Library that obtain access to this index directly and not through tools or interfaces of the Metadata Hub.

## Sandbox

The application or more precisely the web portal sandbox is designed for verifying the correctness of assumptions adopted in the conversion tables created for new data sources. It is a test version with the same GUI as the production GUI where a sample of data from a newly connected source is processed for a determined time. This tool allows checking correctness of the conversion and making necessary corrections using other tools in order to achieve a desired final effect. The application sandbox should operate at a temporary address, separate from the main address of the multibrowser but available to those external users who know about it.

The pattern of operation should be similar to the one in the nearest diagram where close communication between the user and converter administrator is assumed in order to facilitate making corrections in the conversion table. The Sandbox is not responsible for such communication but only for displaying converted data and making them available to the user. The Contractor should take into account the possibility of introducing clear division for individual data sources so that it is possible to quickly select data form the source the user is interested in.

According to estimates, in order to achieve relevant test results it is necessary to import at least 10% of data from each source and such a level is a minimum requirement. The assumptions are made that this parameter is configurable and the application may be applied to individual sources. This means that the converter system administrator will be in a position to make a decision if, for a given source, a sample is created in the application sandbox or all data are converted to the main application.

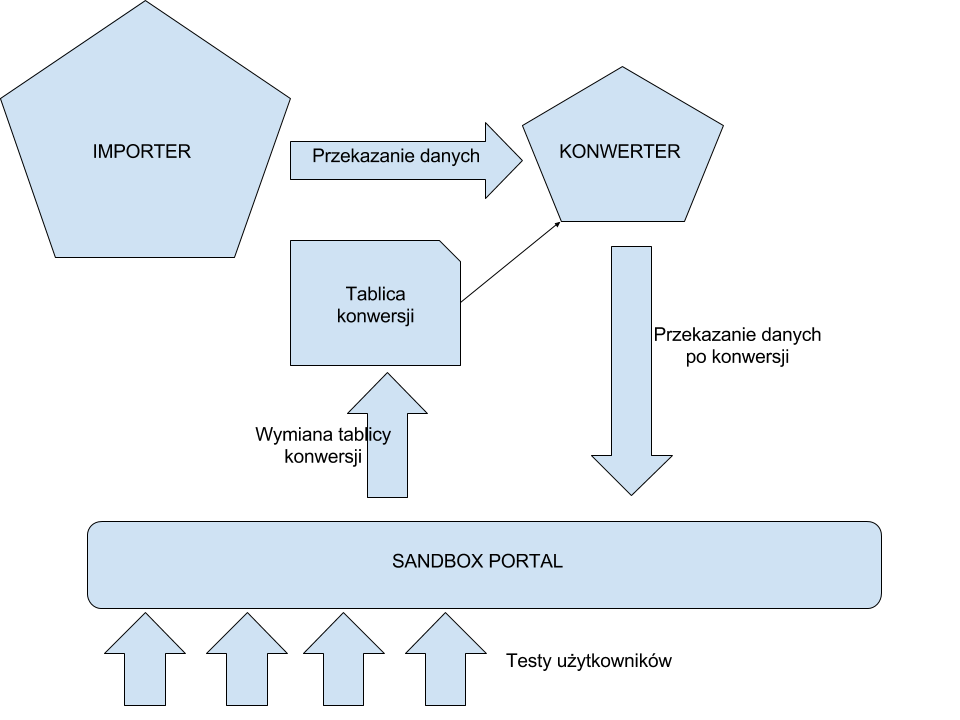


Diagram 4 Functional diagram of the application Sandbox

## Converters

The Contracting Authority expects that the conversion mechanism, when combined with the importer, will form a solution that is capable of downloading data and transforming them in line with the editable definition of formats (editable conversion table) and offer the possibility of launching it after it is disconnected from the remaining infrastructure of the metadata Hub. The Contractor is required to combine these two elements in such a way that they may be used as a separate application that may be run on a virtual server.

The conversion tables shall be developed in close cooperation with the Contracting Authority. The Contractor shall ensure that the Contracting Authority may edit these tables after completion of the programming work. It is assumed that the Contracting Authority has knowledge about data formats and necessary conversions and the Contractor’s task consists in developing a mechanism for carrying out conversion on the basis of guidelines specified by the Contracting Authority.

Conversion tables for individual formats are provided for the functionality described in point 3.2 and allow one to define a conversion of the connected data source to the OMNIS format and, at the same time, to convert such data to other formats defined by the Contracting Authority that are made available for downloading by end users. Within the Converter tool, the system shall offer the possibility of adding next conversion tables and defining any number of output formats with maintaining the same data transfer protocols even after completion of the project. The Contractor shall offer access to data at least in the unified format OMNIS and formats MARC21, DublinCore, ONIX, BibText, EndNote and Zotero.

The Contractor shall ensure the possibility of configuring conversion tables by defining a file with a Contractor-specified structure and saving it in a selected location so that the system could read it as a next usable mapping table. Additional points will be awarded for a solution where the converter module is designed so that it is possible to create such a file at the system level, using the graphical interface.

## Record life cycle

The application is required to support the option of regularly enhancing of the already selected record by both structurizing it and acquiring enhanced information from intermediate systems. The diagram below where a “record live cycle” is presented shows one of optional ways that may be used by the system OMNIS to acquire data and gradually enhance them by exchange of data with the module BN Repository module - publishing repository e-ISBN. A simple record in the ONIX format is converted to the OMNIS open format and then structured and placed in a relevant location. If some linked structure already exists, a simple record is connected to it and if such a structure does not exist, it is established. Further linked systems download data about the new record based on the set of data defined at the conversion stage as identification data, match them and then gradually enhance the record. In this example, the library catalog imports basic metadata in order to complete them with more detailed information and saves them in its database while notifying the OMNIS Hub of the new changes. Then the record is downloaded again and supplied with information that was missing before.

Within the work carried out, the Contractor shall take into account the scenario where the data source was updated (e.g. migration to a new system) and a necessity arouse to make changes within applications used for data downloading as described in sections 3.2-3.5. In this case, the record should be replaced with the one from the new version of the system. There should be ensured the possibility of detecting potential duplicates for data sources in order to prevent the same records, from the same sources but from different systems from simultaneous operating. Information about the record in a previous version may be saved (or need not to be deleted) but inquiries should be handled using the latest version of the given record.

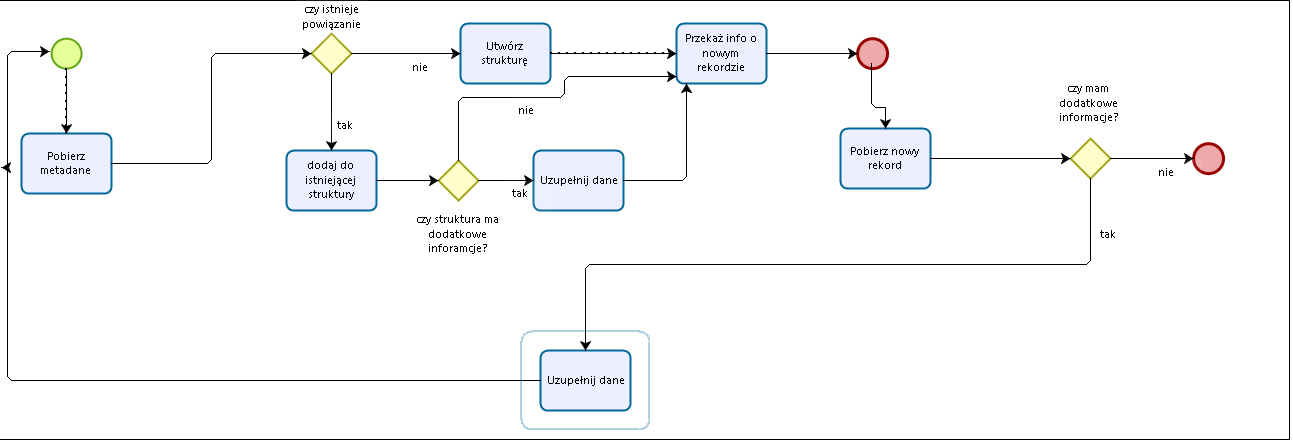


Diagram 5 Record life cycle

## Multibrowser

###### Principle of operation

The search form has to offer solutions typical for popular web browsers. As a response to the user’s query, there should be displayed a result list with a wide selection of faceted filters that allow one o filter results depending on the mode of access to a given publication, location of the publication and set of filters generated in accordance with descriptive metadata regarding publications that match data for publications as specified in the data model OMNIS, as well as a list of associated results.

It is required that the result lists in the first place display collections available in the open internet including the BN repository and then (for collections protected by copyright) in the Academica loan system, and finally those available in paper versions. In case of protected collections, it must be possible to switch to the relevant booking process in the Academica loan system at the reader-selected institution in Poland or to find a library with a physical copy.

The Proxima service shall use information about the present location of the reader (obtained automatically from the mobile or standard browser) or other location specified by the reader to offer the possibility of sorting results from the shortest to the longest geographical distance to either a library that holds a physical copy, or an Academica terminal.

The Multibrowser Omnis has to be adapted to at least five leading Internet browsers and additionally shall have a version optimized for mobile devices. Popular internet browsers are those that have the highest share in the market of internet browsers that is: Chrome, Safari, Firefox, Internet Explorer, Opera.

Mandatorily, results returned by the multibrowser OMNIS have to be indexed by the most popular universal browsers and at least by Google, Bing and Yandex.

Basic browser functions have to available without logging in. However, users who use their accounts should have access to the full functionalities of tools offered in this service. As part of using this service, the user shall have access to intelligent tools that offer the possibility of advanced and personalized search. Individual user accounts shall allow one to use personalized settings such as using saved search criteria, generating messages notifying the user of new publications available in the system and meeting the set criteria, building one’s own collections of publications, etc., that will be determined within the cooperation with the Contracting Authority in line with the SCHEDULE (Cooperation between the Contractor and Contracting Authority regarding the designing of the UX tests as described in the task covering other procurement under the Project and pertaining to “Design of the UX for access interfaces”, and implementation of guidelines developed within this task with the assumption that the preliminary deliverables of the task “Design of the UX for access interfaces” will be transferred and verified, and the Contractor will provide its suggestions). Using the dedicated mobile interface Multibrowser OMNIS, apart from the described above functionalities, the system shall also show the nearest library where a physical copy of the desired publication is present. In case of the e-service, the intended result is achieved when information is obtained in line with specified source-dependent criteria and a reference is given to a proper location (e.g. a digital library, loaning system Academica, the nearest library offering the title, etc.)

The Contractor shall design a mechanism whereby information about availability and location of a given publication is provided. The Contractor shall ensure the possibility of personalizing functionalities of the OMNIS browser where the degree of personalization will be determined in cooperation with the Contracting Authority in line with the SCHEDULE (Cooperation between the Contractor and Contracting Authority regarding the designing of the UX tests as described in the task covering other procurement under the Project and pertaining to “Design of the UX for access interfaces”, and implementation of guidelines developed within this task with the assumption that the preliminary deliverables of the task “Design of the UX for access interfaces” will be transferred and verified, and the Contractor will provide its suggestions). At the interface level, the Contractor shall ensure the possibility of integrating data from other sources in order to create applications that combine data from different sources (mashup applications). The Contractor shall implement the graphical concept of the interface (graphical elements compliant with the common graphical concept of the National Library) provided by the Contracting Authority.

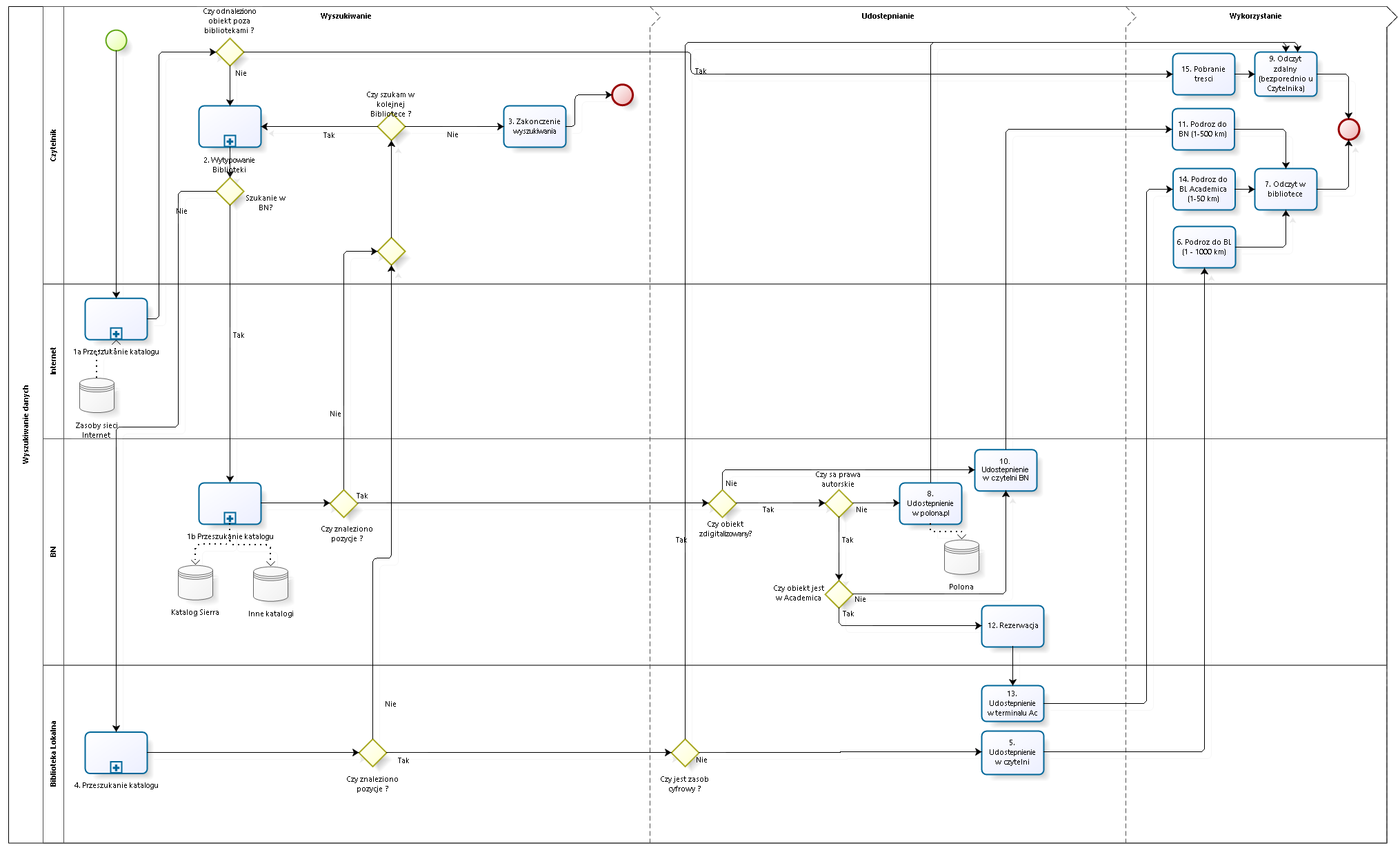


Diagram 6 Search

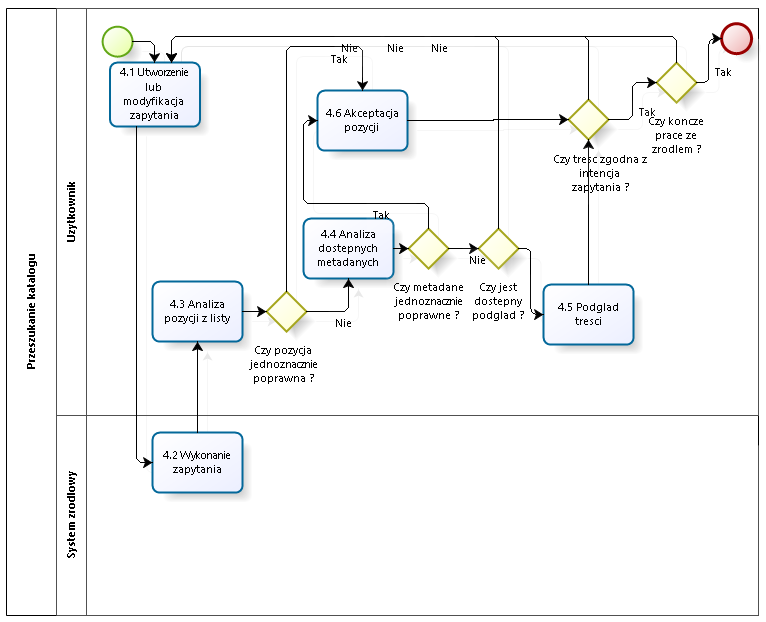


Diagram 7. Catalog search

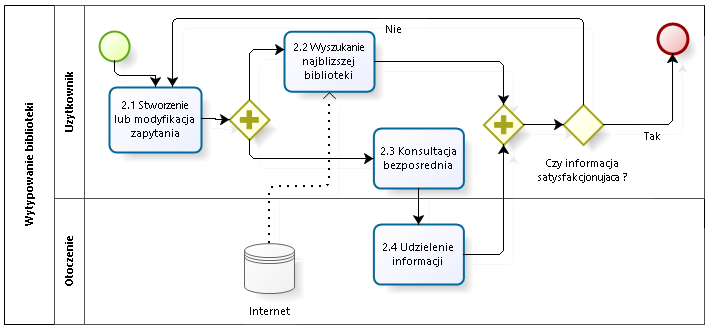


Diagram 8. Search for a library

###### Graphical interface

The Contractor shall cooperate with the Contracting Authority in work on the appearance of the OMNIS browser as the only element with important components of the user interface. The Contractor shall provide functional elements for which the Contracting Authority shall provide a graphical interface.

In the scenario of searching for a single copy, at least three levels of the interface are assumed, which should allow operating the search functions. The detailed scope of displayed metadata and indices of faceted filters shall be determined in cooperation with the Contracting Authority, depending on finally accepted data model. At this stage, it is known that it will be necessary to implement a “data sources” filter for making inquiries about objects from a given data source and allowing one to search only collections from a given source.

The Contractor shall provide the user with the possibility of logging in at least through an e-PUAP account, as described in section 4, and saving successive searches as inquiries on his/her account. Then, at the next logging in, the system shall provide information about new objects that satisfy requirements of the saved search query. Search query results are displayed as search results.

At the first level of search results, the user will receive information about a given entry, i.e. integrated information about a given title and author(s) and topics. This is meant to provide general information about objects that potentially may meet conditions of the query. At a higher level, after narrowing down to the title, author or topic and selecting a specific work, there will be provided information about its all known versions distinguished in accordance with information about edition or date of publication, translation or language of the work and other differentiating data. This is aimed at obtaining information about all embodiments/editions of a given work and allowing the user to select the most appropriate for him/her. After entering the level of a given edition, the user receives information about copies of the given edition of the work. At this level, the user should be provided with complete information about all copies of a given edition that are present in the network of libraries together with information about physical location of a given copy and, if possible, its availability. This way, it should be possible to find a library with a selected copy of the desired work.

The Contractor shall develop an interface for the mobile version of the Multibrowser Proxima, whose main task is searching for a copy of the desired work that is nearest to the user’s location. The interface should offer the possibility of searching at different levels, as it is the case with the main interface, but its main functionality should be localizing. A similar functionality shall be available as a separate subpage in the interface of the main browser. Using the dedicated mobile interface, the Multibrowser OMNIS “Proxima”, apart from the described above functionalities, shall also show the nearest library where a physical copy of the desired publication is present. Consequently, it will be possible to completely integrate information about collections of libraries with online catalogs and significantly improve the reader service quality within the national library network.

In the course of programming work, the Contractor shall take into account the possibility of connecting to the browser interface via the API so that the query entered in the external form could be redirected to the result list of the main multibrowser. An example of such a functionality is the embedded search window in the main page of the National Library where after selecting the option of searching collections, the user is redirected to the multibrowser interface displaying a list of results that match his/her query. The Contractor shall offer an option of making such a query using an embedded “data sources” filter, i.e. searching for information desired by the user only within collections of the selected institution.

At the request of the user, it shall be possible to make a search in the text layer consisting in searching full-text resources indexed within the metadata Hub. Results of the text search may present fragments of the text that meet satisfy the search conditions together with information about the copy from which given content is taken.

The following drawing (diagrams 9 to 15) present examples of appearance of the browser and its basic functions. Using it as the basis, a graphical version of the browser shall be developed where, as far as the functional aspect is considered, the Contractor shall implement a provided graphical design - it is assumed that the Contractor and Contracting Authority will cooperate in the designing of the GUIs described in the task covering other procurement under the Project and pertaining to “Graphical designs of the access GUIs”, and guidelines developed within this task will be implemented. According to the assumptions of the cooperation, preliminary results of the task “Graphical designs of the access GUIs” will be delivered to the Contractor and verified and the Contractor will make suggestions so that the design of UX for the Multibrowser OMNIX takes into account necessary graphical elements developed within the above-mentioned task and it is possible for the Contractor to implement them at a later date.

The minimum set of functionalities includes:

1. Logging in
2. Downloading data of a single copy
3. Downloading record structure data
4. Prompts for the phrase searched for
5. Displaying related records and relations
6. Going over to source systems
7. Displaying a single copy record
8. Displaying an edition record
9. Displaying a work record
10. Information about availability and source of a digital version of a given copy
11. Simple search
12. Full-text search
13. saving search query
14. Sorting search results
15. Filtering search results
16. Multi-level search result list
17. Information about the copy location
18. Pagination of the result list
19. references to hits in content of the documents
20. Presenting information about the place where the library object is present and availability of the copy
21. Downloading and displaying information about the location of libraries
22. Full-text search with contextual display of phrases
23. Simple and advanced searches with intelligent filters
24. De-duplicating search results
25. Clustering of search results in line with assumptions of the FRBR model
26. The option of displaying all versions of the document present in the grouped search results
27. Exporting bibliographic description in a selected format
28. Access to full-text digital versions of works from the public domain
29. The Proxima service - information about availability of a given item at the geographically closest libraries

The Contractor will take into account the need for Responsive Web Design (RWD) approach during the GUI development, that is, the design of an information architecture, interaction, and display of content, so that users can seamlessly use implemented services on screens with various diagonal screens with different real and effective resolution, specific for the web browser of the device. The mobile interface "Proxima" is supposed to be a multi-lookup browser responsive GUI version of the desktop GUI, which will be based on the sample screenshots below. The final version of the entire graphic design will be submitted to the Contractor in accordance with the Schedule and may differ from the version on the sample screenshots.

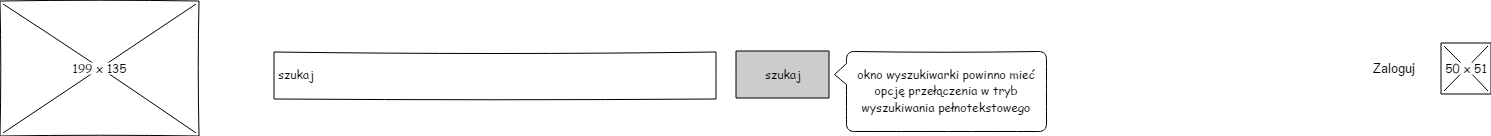


Diagram 9. GUI Level 1. Browser window.

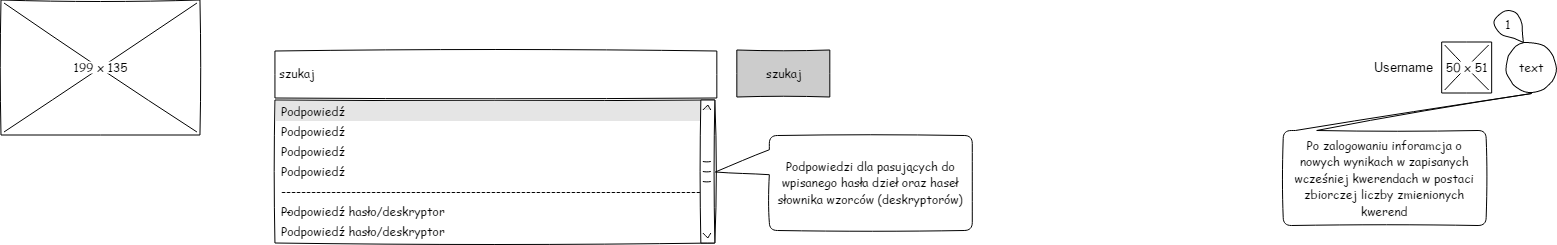


Diagram 10. GUI Level 1. A logged-in user and prompts

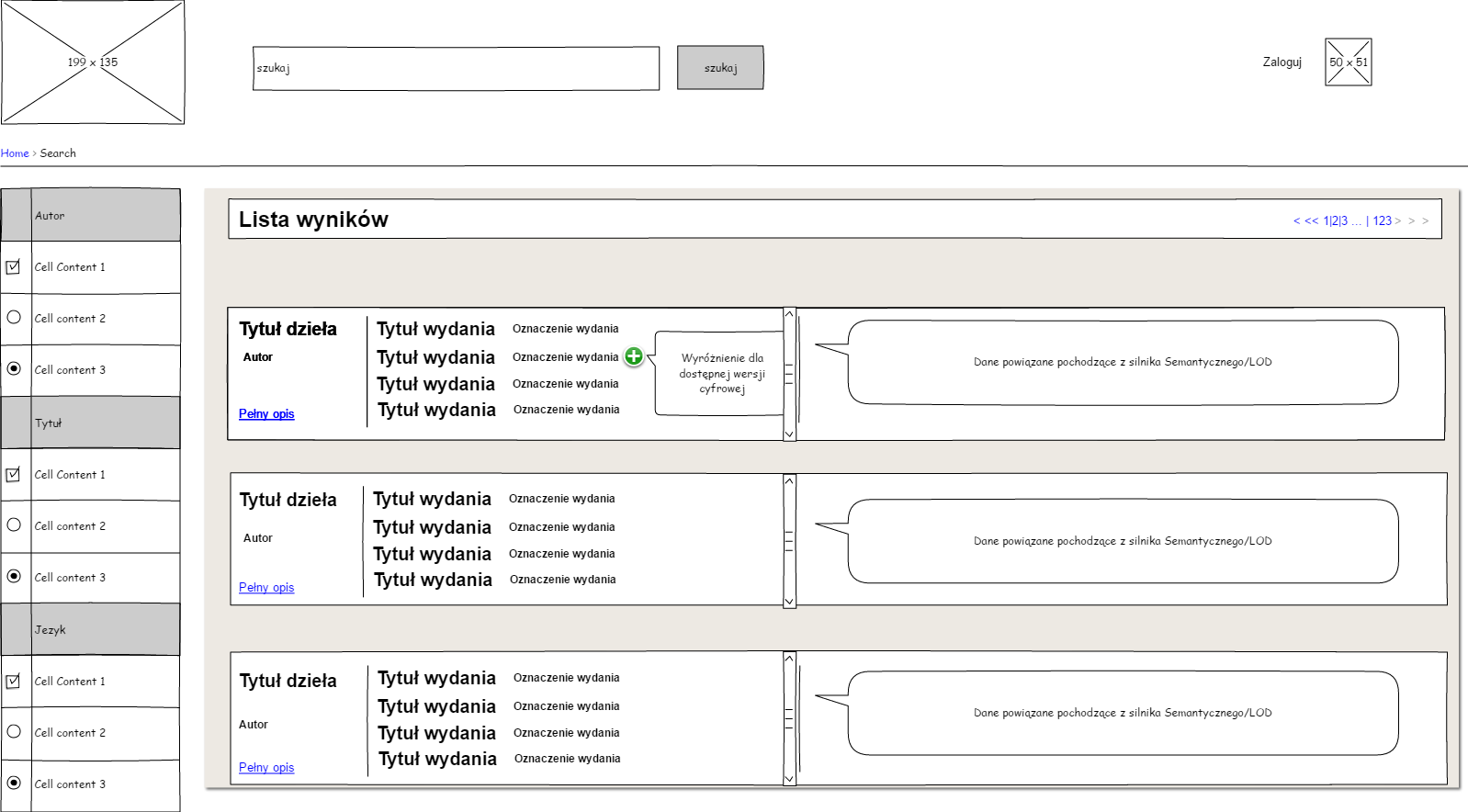


Diagram 11. GUI Level 2. Result list.

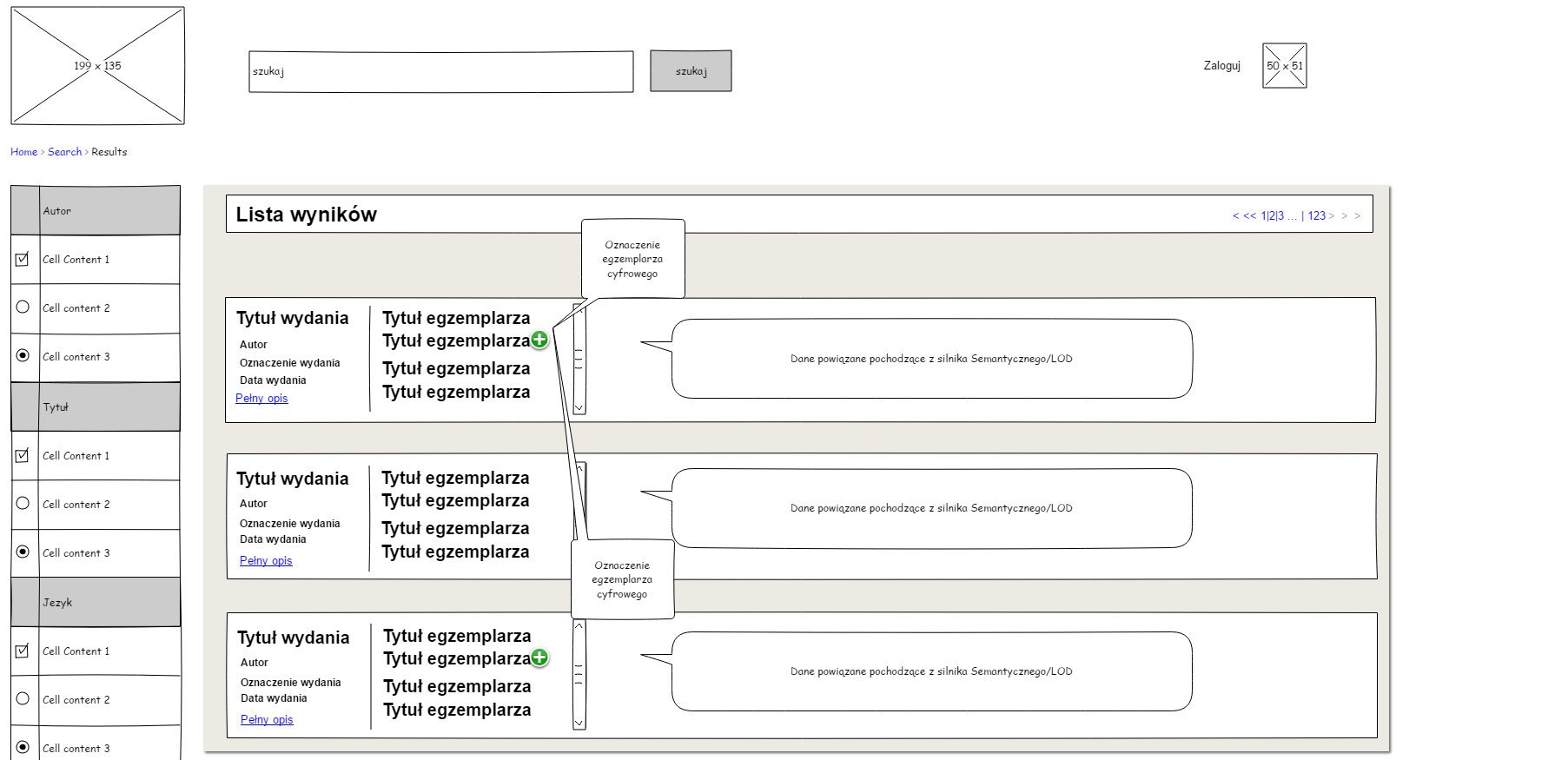


Diagram 12. GUI Level 3. Result list with the list of copies.

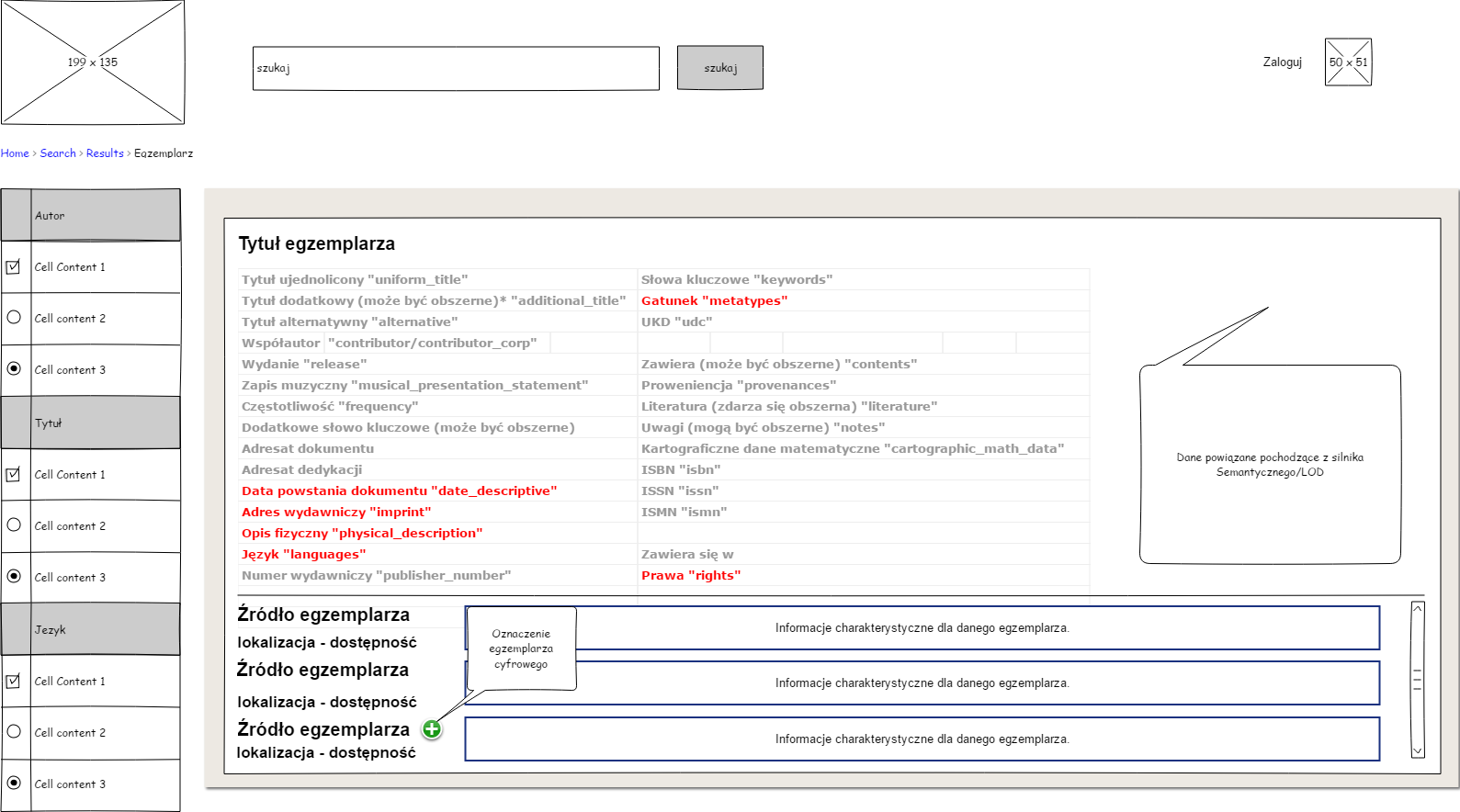


Diagram 13. GUI search: a single copy level with a location list



Diagram 14. GUI: Visualization of the locations on the map

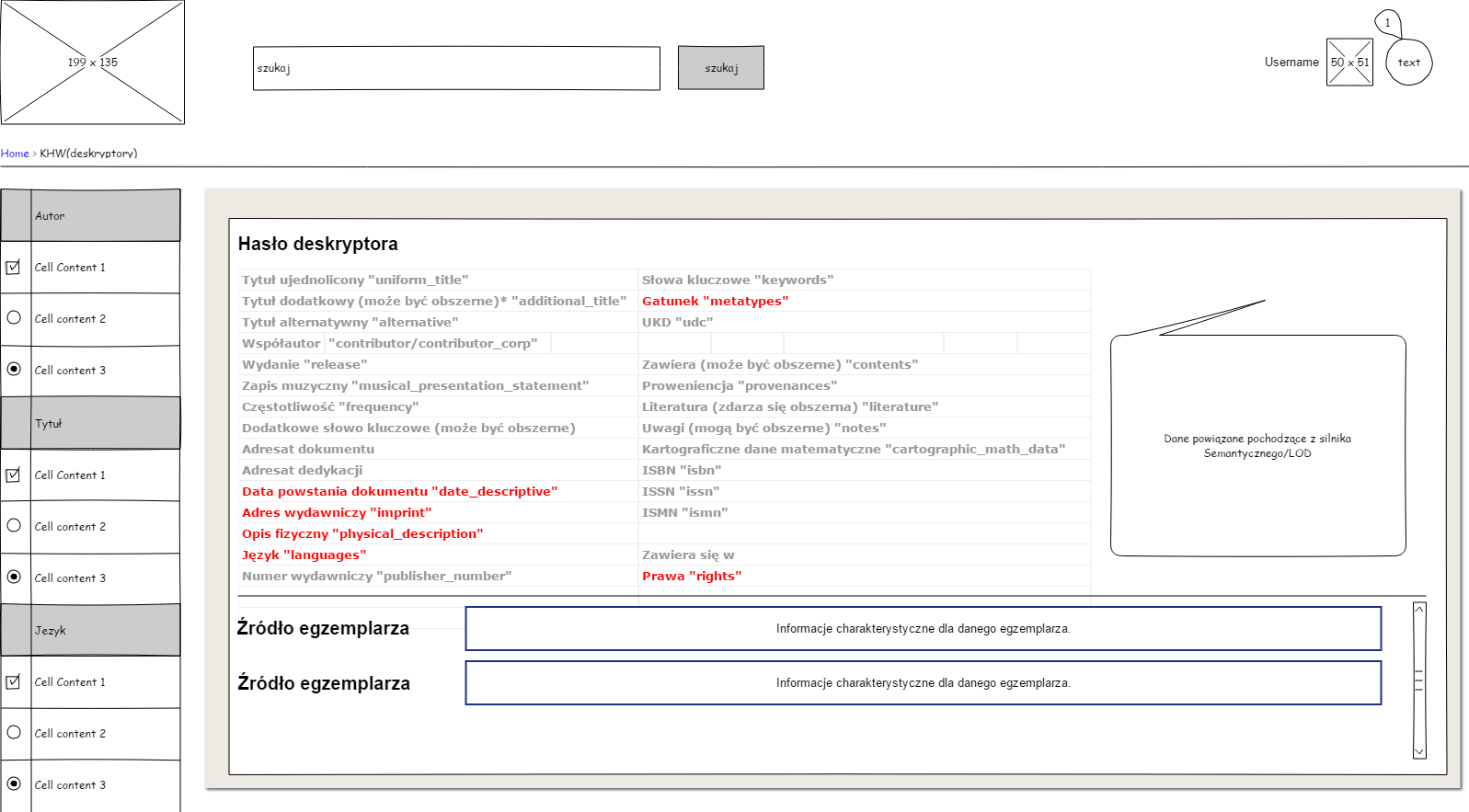


Diagram 15. A model entry.

## Data model

Within the main Hub module, the Contractor shall build structures based on data downloaded from distributed sources, in line with concept of the FRBR model, i.e. create clusters that are other forms of the same work (e.g. a book, a play recorded as a video, a movie), and references to clusters interrelated in accordance with a pre-defined algorithm based on rules consistent with the formal ontology of the OMNIS model (e.g. other works of the same author, works of a similar subject or inspired by a given title). After such a process, source objects should have their representation in the multibrowser index or be linked to objects that already exist in the multibrowser so that it is possible to take them into account in user’s searches. Owing to this approach, the user in return to their query shall see all institutions holding such an item in their collections and additionally be given a full ability of navigating related authors, institutions and other works associated, for example, in the FRBR process.

The Omnis data model shall be based on the DBN (National Library Descriptors) concept employed in the National Library and elements of the FRBR-LRM model (https://www.ifla.org/files/assets/cataloguing/frbr-lrm/frbr-lrm\_20160225.pdf). This does not mean the FRBR-LRM model concept is to be implemented in its entirety. In the course of work and cooperation between the Contractor and the Contracting Authority, there will be determined appropriate levels of detail for the model, depending on availability of metadata of specific types. The main assumptions of the FRBR model are the division into entities and relations between them. In this respect, the diagram E-R where individual entities (matching individual descriptors): <http://przepisy.bn.org.pl/deskryptory/zasady-tworzenia-deskryptorow-bn> and their attributes: <http://przepisy.bn.org.pl/deskryptory/atrybuty-deskryptorow> are linked with one another, is the starting point for the DBN concept, the OMNIS data model, and the relevant formal ontology. It is assumed in the OMNIS data model that a possibility exists to expand such relations and create clusters based on data from other sources. The detailed data model and formal ontology shall be provided by the Contracting Authority. The contractor shall take into account the possibility of evaluating and optimizing them in line with a previously described Converter mechanism, whereby it is possible to edit input and output data sets.

## OMNIS Linked Data

Development of application designed for publishing structured and uniformed data in line with the LOD (Linked Open Data) methodology and releasing such data for free circulation of information in the Linked Data cloud where, in the minimum variant, at least the Schema.org (http://schema.org) dictionary and its micro-data tags are used, with the assumption that other dictionaries may also be employed.

The solution OMNIS Linked Data shall be based on elements of the semantic module within which the Contractor shall implement a separate RDF Store for integrated data obtained from other OMNIS modules as well as from other data sources and Linked Open Data (LOD) together with its own SPARQL endpoint and LD interface. In the module, it should be taken into account that data integrated in the RDF Store is obtained both from the main module and the full-text module. For data from the main module, the Contractor shall ensure conversion to the RDF format based on the conversion tables that are to be delivered under a separate subtask (other task that, according to the schedule, should be completed before commencement of implementation of the Hub considered under this contract - the product shall be delivered by the Contracting Authority). For data from the full-text module, the Contractor shall ensure the possibility of integrating separated entities matching the data model, obtained from the main module, that were extracted with a tool implemented under a separate task (other task that is scheduled for completion by the end of March, 2018 - the product shall be delivered by the Contracting Authority). In the minimum variant, OMNIS data shall be described using the Schema.org dictionary with its wide selection of microdata tags.

The semantic module has to be equipped with a Contractor-provided Data Integrator for the RDF Store that is capable of integrating data from different Open Linked Data services. Under the cooperation with the National Library team, there shall be carried out necessary integration of formal ontologies of data sets and the OMNIS data model being integrated based, among other things, on the conversion tables that are to be delivered under a separate subtask (other task that, according to the schedule, should be completed before commencement of implementation of the Hub considered under this contract - the product shall be delivered by the Contracting Authority). Integrating formal ontologies to a common ontology model assumes that the Contractor will make it possible for the National Library team to evaluate the developed solutions and making corrections to previously arranged conversion tables. As an result, there should be created and efficient service offering data of the OMNIS model based on its formal ontology and combining data from other Open Linked Data services and using their multiple ontologies, (including the DBPedia Ontology, BIBFRAME Ontology, EDM OWL Ontology, FOAF, dcterms), based on mapping data within the formal ontology used as the basis for the OMNIS data model.

Additionally, for certain data subsets from the LOD cloud, the Contractor shall provide the possibility of storing data dumps as RDF files in order to ensure smooth integration with the semantic module of the Hub and the Hub-based applications if such data are unavailable. Such data will be used for developing solutions presented via the multibrowser OMNIS and the Contractor shall ensure the possibility of their integration.

The semantic module has to be fitted with a data publishing mechanism (Data Republisher) allowing one to republish linked LOD data in the cloud (http://lod-cloud.net).

As part of the exchange of data from the semantic module of the Hub, structured and uniformed data stored in the OMNIS system are published in line with the LOD (Linked Open Data) methodology and released for free circulation of information in the Linked Data Cloud. Additionally, for data from the RDF Store, the Contractor shall provide the option of storing data dumps as RDF files in order to ensure reliable access for web servers.

Using APIs, the service shall enhance search mechanisms and provide an SPAQRL endpoint as well as offer a GUI for visualization of dataset elements.

Using as well as offer a GUI for visualization of dataset elements.

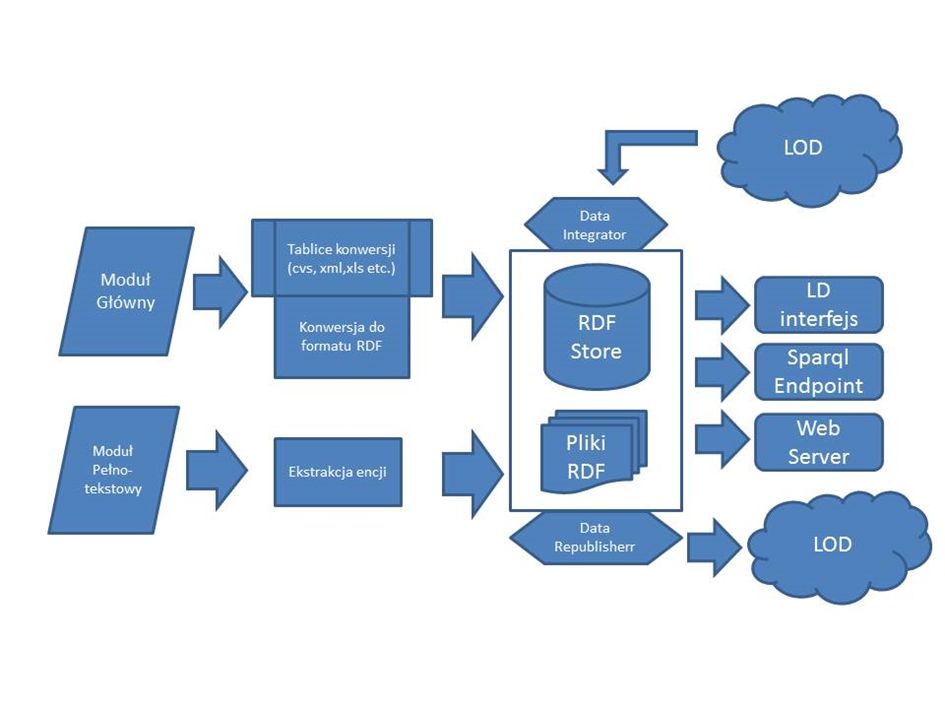


Diagram 16. Schematic diagram of connections between the LOD module and the metadata Hub.

The contractor shall provide the following functionalities:

1. RDF Store for integrated data obtained from other OMNIS modules as well as from other data sources and Linked Open Data (LOD)

2. APIs service that shall enhance search mechanisms and provide an SPARQL endpoint allowing for querying and extracting data from the above mentioned RDF Store

3. Converters that shall ensure conversion to the RDF format based on the conversion tables, using elements of formal ontologies typical of Semantic Web solutions, that are to be delivered under a separate subtask (other task that, according to the schedule, should be completed before commencement of implementation of the Hub considered under this contract - the product shall be delivered by the Contracting Authority)

4. Markup of OMNIS data using the Schema.org dictionary with its selection of microdata tags

5. Data Integrator for the RDF Store that is capable of integrating data from different Open Linked Data services with OMNIS data

6. Data publishing mechanism (Data Republisher) allowing to republish linked LOD data in the cloud (<http://lod-cloud.net>), ensuring that data stored in the OMNIS system are published in line with the LOD (Linked Open Data) methodology and released for free circulation of information in the Linked Data Cloud

7. For data from the RDF Store, the Contractor shall provide the option of storing data dumps as RDF files in order to ensure reliable access for web servers

8. For certain data subsets from the LOD cloud, the Contractor shall provide the possibility of storing data dumps as RDF files in order to ensure smooth integration with the semantic module of the Hub and the Hub-based applications if such data are unavailable

9. GUI – LD interface for visualization of dataset elements and querying the dataset

## UID/URI

A persistent identifier forms a basis for operation of a modern IT system. It is a code that uniquely identifies a given electronic resource and, at the same time, may be quoted in order to share information in a longer period. Identifiers are of a tremendous value for the metadata Hub system. It is unacceptable to store millions of object without ensuring their persistent and unique  
 identification. The Contractor shall ensure that the system offers such uniqueness and persistency even when resources are transferred to another web location. If in the course of development this structure becomes disrupted for any reason, it is necessary to maintain its backward compatibility with the previous version.

The URI is meant as a “concise string of characters that identify abstract or physical resource”; the UID is only an alphanumeric identifier of a given resource that is usually found at the end of the URI. The Contractor shall follow the guidelines for the standard URI format http://{domain}/{type}/{concept}/{reference}

• {domain} means a combination of a host and appropriate sector

• {type} means the type of the resource being identified.

• {concept} may by a collection, type of identified element of the real world, or a title of a concept model

• {reference} a specific element, term or concept

## Data access APIs

###### One access API

The Contractor shall provide an application API such that it may be used both by the user’s GUI for downloading data generated and completing the prompt list, and the external systems that act as clients of the API service and download data from the metadata Hub. The exception is information linked to the logged-in user that is accessible only to the API of such a user and where such an API is made available only to external systems (e.g. the GUI of the browser). The API is required to be capable of returning data in all supported data formats in a JSON or XML file using the open protocol http.The client that asks for a given resource should specify the desired target data format (Onix, Marc 21 or the open format OMNIS) and the preferred data exchange format (JSON, XML or other). The provider of APIs has to ensure their high efficiency and capability of simultaneous handling multiple queries both for external users and cooperating systems.

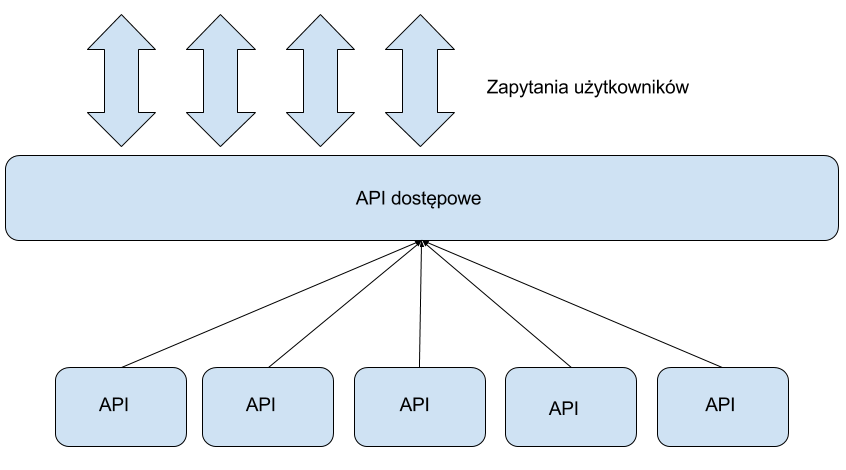


Diagram 17. Functional diagram of the access API

Creating request addresses

Example address: /api/authorities.json?limit=20&sinceId=2

.json - specifies the response format

? - the symbol preceding request parameters, e.g. id, date of creation

& - the symbol separating individual parameters

Request parameters

id

Record identifiers, natural numbers, with the possibility of specifying them after a coma. May have the prefix “a” and a check digit.

The system is required to offer the following functions:

1. API for bibliographic records
2. API for statistics
3. API for authority records
4. API for location
5. API for the full-text search
6. Structure-returning APIs
7. API for the logged in user

###### API for location

The API for location, in line with the requirement for the importer mechanism and converters, shall be able to provide information on the location of the source from which data of a single copy originate. The Contractor shall ensure that at the data import stage, each source may be provided with information about physical location of a given collection. For example, data from the National Library may be accompanied by information about the location of the National Library:

*al. Niepodległości 213*

*02-086 Warszawa*

As the basic function, this mechanism has to find a location that is nearest to the user and offers the title selected by the user.

###### API for the full-text search

This response from the API may be obtained as a separate query and a separate API terminal or as part of the API for the bibliographic record. It shall be accessible only on the user’s demand when the full-text search option is selected. The syntax of the query could be similar that applied in the query about the bibliographic record with the “fulltext” information added to it:

?query=treść\_zapytania&fulltext=1

If this is the case, the system has to identify it as a request of a different type and send the query to the full-text search engine.

As a response, the user will receive a list of documents containing the searched word of phrase together with fragments of texts where the desired word or phrase has been found.

###### Structure-returning APIs

The API mechanism has also to be able to return information about the data structure and record’s location in such a structure. As described in section 3.6, the presentation of results is divided into at least three levels. The API should return information about each of them. Consequently, on the user’s request, there will be available information about other issues of a given object and their copies as well as about all related records.

The record structure also defines data returned from the semantic engine LOD and such data should be placed in the API in this category. Similarly to case 3.10 “One access API”, it is assumed that in this category may be located API terminals from multiple semantic mechanisms that may operate completely independently.

###### API for the logged in user

The API for a logged-in user shall allow authorizing the user and providing access to data saved in his/her account. In case of the multibrowser, these will mainly be previously saved search queries. The Contractor shall ensure that all functions necessary for providing this functionality are ready for operation without using the GUI but just via the API.

###### API for statistics

It is required that the statistics system be of a possibly simple and flat structure that returns purely quantitative data for the defined entities. Quantitative information should reflect a general view of what is present inside the metadata Hub, i.e. the total number or records, their subject categorization, sources or mode of access. The API for statistics has to be able to present the overall number of the Hub user accounts and the overall number of saved search queries.

An example of API’s response to a query about statistics as received from the Polona.pl:

**GET** /api/stats/

**HTTP 200 OK**

**Allow:** GET, OPTIONS

**Content-Type:** application/json

**Vary:** Accept

{

"entities": {

"total": 1634548,

"public": 658087,

"by\_category": {

"maps": 5498,

"serials": 951026,

"scores": 5001,

"manuscripts": 5992,

"press": 549532,

"prints": 15629,

"photographs": 22054,

"postcards": 16272,

"ephemerals": 34692,

"books": 73460

}

}

}

###### Example of API for bibliographic records

Access to bibliographic data via the API is based on the assumption that the system has to provide access to a single object through the GET parameter and entering the unique ID. /entities/QWERTY/.

The API shall also make it possible to search for a group of records, which has to be possible using the basic parameters of the query. In case of such a query, the system shall also be able to return random results, if a seed-type parameter has been specified. Furthermore, an option has to exist, whereby it is possible to select sorting options such as in the predetermined order, within a defined range or according to identifiers.

Example:

*sort [optional] (default: score desc) - required format is field type, allowed fields are: score, title, creator, date, allowed types are: asc, desc*

*from [optional] (default: 0) - integer, indicates from which record results should be returned.*

*seed [required] - random seed, integer 1-99. Used with conjunction with form and size.*

Basic search parameters:

*query - string eg. "mickiewicz".*

*Advanced query*

*query - eg. field:value OR field:value AND (field:value OR field:value). Allowed fields are: title, author, keywords, publish\_place, publisher, frequency, sources, signatures, entire\_description, content*

The full list of allowable parameters shall be determined by the Contracting Authority in cooperation with the Contractor within the cooperation the result of which should be a detailed OMNIS model as it determines the target set of input data. One may assume that the list of parameters will include a slightly changed parameter list of the current API of the service data.bn.org.pl The Contractor shall offer access to data at least in the unified format OMNIS and formats MARC21, DublinCore, ONIX, BibText, EndNote and Zotero.

An example of API’s response to a query about a single UID as received from the Polona.pl:

**GET** **GET** /api/entities/uid/

**HTTP 200 OK**

**Allow:** GET, OPTIONS

**Content-Type:** application/json

**Vary:** Accept

{

"id": "ODQ0NjUz",

"slug": "najnowsza-kuchnia-domowa-najnowsza-polska-ksiazka-kucharska-zawierajaca-latwe-wskazowki",

"title": "Najnowsza kuchnia domowa : najnowsza polska książka kucharska zawierająca łatwe wskazówki w gotowaniu i przyrządzaniu używanych w życiu polskim zup, warzyw, sosów, potraw mięsnych i rybich, klusek, sałatek etc. etc.",

"alternative": null,

…

"creator\_corp": null,

"contributor": null,

"contributor\_corp": null,

"date": "1913-01-01T01:25:00",

"date\_descriptive": "1913",

"country": "Polska",

"publisher": "Księgarnia Popularna",

"publish\_place": "Warszawa",

"publisher\_number": null,

"imprint": "Warszawa : Księgarnia Popularna, 1913 ([s.l.] : \"Drukarnia Krajowa\")",

"frequency": null,

"edition": null,

"series": null,

"copublished": null,

"chronological\_term": null,

"physical\_description": [

"79 s. ; 20 cm"

],

"file\_characteristics": null,

"categories": [

"books"

],

"metatypes": [],

"udc": null,

"material\_type": "a",

"bibliographic\_level": "m",

"isbn": null,

"issn": null,

"ismn": null,

"call\_no": [

"I 1.576.792"

],

"oclc\_no": null,

"national\_bibliography": null,

"academica\_id": 844653,

"year\_of\_edition": null,

"tome\_of\_edition": null,

"number\_of\_edition": null,

"release": null,

"press\_title": null,

"languages": [

"polski"

],

"sources": [

"Biblioteka Narodowa"

],

"projects": [],

"provenances": [],

"subject": [

"Kuchnia polska"

],

"keywords": [

"Kuchnia polska"

],

"resources": [

{

"mime": "application/xml",

"hidden": false,

"url": "https://polona.pl/metaexport?uid=844653&mapping=polona\_src\_out&lang=pl"

}

],

"notes": null,

"literature": null,

"contents": [],

"has\_subseries": null,

"musical\_presentation\_statement": null,

"cartographic\_math\_data": null,

"rights": [

"Domena Publiczna. Wolno zwielokrotniać, zmieniać i rozpowszechniać oraz wykonywać utwór, nawet w celach komercyjnych, bez konieczności pytania o zgodę."

],

"is\_public": true,

"digital\_copy\_by": [

"Zakład Reprografii i Digitalizacji Biblioteki Narodowej"

],

"links": {

"item\_url": "//polona.pl/api/entities/ODQ0NjUz/",

"academica\_url": "https://academica.edu.pl/reading/readSingle?uid=844653",

"catalogue\_url": "//katalogi.bn.org.pl/iii/encore/record/C\_\_Rb1657161"

},

"has\_text\_content": false,

"aggregations": {

"user": null,

"global": {

"views\_count": 0,

"bookmarks\_count": 5,

"collections\_containing\_count": 1,

"favorites\_count": 194,

"notes\_count": 10

}

}

}

###### Example of API for authority records

Similarly to the bibliographic records, the API has also to offer the possibility to make a query about a specific record or range of records or make it possible to carry out a search using defined indices. Thus, this functionality should also be similar with the assumption that as a response a record of different type will be delivered and consequently some indices specific for entities will not be applicable in this case.

The full list of allowable parameters shall be determined by the Contracting Authority in cooperation with the Contractor within the cooperation the result of which should be a detailed OMNIS model as it determines the target set of input data. One may assume that the list of parameters will include attributes of authority records and will slightly deviate from the parameter list of the current API of the service data.bn.org.pl

An example of the API’s response for an authority record as received from the Polona.pl:

**GET** /api/authorities/uid

**HTTP 200 OK**

**Allow:** GET, OPTIONS

**Content-Type:** application/json

**Vary:** Accept

"data": "Kik, Kazimierz"

},

{

"code": "d",

"data": "(1947- )"

}

]

}

},

{

"tag": "667",

"ind": " ",

"data": {

"subfields": [

{

"code": "a",

"data": "osobowe"

}

]

}

},

{

"tag": "667",

"ind": " ",

"data": {

"subfields": [

{

"code": "a",

"data": "Nauki historyczne, nauki o polityce: integracja Europy, politologia, współczesne aspekty polityczne Europy."

}

]

}

},

{

"tag": "670",

"ind": " ",

"data": {

"subfields": [

{

"code": "a",

"data": "Komunistyczna Partia Hiszpanii / Kazimierz Kik. - Warszawa, 1986."

}

]

}

},

{

"tag": "670",

"ind": " ",

"data": {

"subfields": [

{

"code": "a",

"data": "Kto jest kim w Kielcach. T. 1 / Zbigniew Judycki, Józef Siwek. - Toruń, 2002."

}

]

}

},

{

"tag": "670",

"ind": " ",

"data": {

"subfields": [

{

"code": "a",

"data": "Nauka Polska"

}

]

}

},

{

"tag": "670",

"ind": " ",

"data": {

"subfields": [

{

"code": "a",

"data": "Wikipedia"

}

]

}

},

{

"tag": "024",

"ind": "7 ",

"data": {

"subfields": [

{

"code": "a",

"data": "https://viaf.org/viaf/59505931"

},

{

"code": "2",

"data": "viaf"

}

]

}

},

…

{

"fieldTag": "y",

"marcTag": "040",

"ind1": " ",

"ind2": " ",

"subfields": [

{

"tag": "a",

"content": "WA N"

},

{

"tag": "c",

"content": "WA N"

}

]

},

{

"fieldTag": "y",

"marcTag": "999",

"ind1": " ",

"ind2": " ",

"subfields": [

{

"tag": "a",

"content": "pej"

},

{

"tag": "d",

"content": "zkd"

},

{

"tag": "b",

"content": "bw"

}

]

},

{

"fieldTag": "\_",

"content": "00000nz a2200061n 4500"

}

]

}

]

}

# AUTHORIZATION

The adopted system user authentication method is of a key importance for access to the system functionalities. In particular, in case of access to the Publishing Repository (A2B), the following approach is required:

1. Authentication with a trusted ePUAP profile
2. Authentication with a user account established in the system (based on the user name, password, and authentication certificate).

The system is required to offer the possibility of authentication the user with a trusted profile via the ePUAP platform, using the SSO mechanism (Single Sign On). Availability of this mechanism makes the system more friendly (the user who already has a trusted profile does not need to remember another password). In case of services for publishers, this method of authentication allows one to enhance the level of integration with already offered services of the public administration.

Apart from the method based on the trusted ePUAP profile, there must be available an authentication method based on the user name, password and authentication certificate, whereby users without a trusted ePUAP profile may access the publishing repository of the e-service Omnis.

The following, required access modes have been adopted for the Integrated Browser Omnis:

1. Anonymous without logging in
2. Authentication with a trusted ePUAP profile
3. Authentication with a user account established in the system (based on the standard authentication method with the user name and password).
4. Authentication using existing information from a major social networking service such as Facebook, Twitter, Google.

Anonymous access shall be the leading mode of access to the Integrated Browser Omnis. After logging in, the user shall be granted access to personalization-related functionalities.

All ICT systems under the project shall be implemented in line with interoperability requirements specified in the Regulation of the Council of Ministers of April 12, 2012, on the National Interoperability Framework, the minimum requirements for public registry and electronic information exchange as well as the minimum requirements for ITC systems (the KRI).

# ADMINISTRATIVE INTERFACE

Apart from the interface described in section Multibrowser, the Contractor shall provide an interface for the system administrative functions. Such interfaces should allow one to configure additional importers, add conversion tables (e.g. as xml files) and link them to specific sources, trace operations of applications, data collection queues and their statistics as well as to stop and resume operations. The administrative interface shall be a tool for administrators responsible for routine operations of the system and shall allow one to configure the system on current basis. Considering the assumptions, it must be possible for the Contracting Authority to carry out these task using the described interface.

The Contractor shall ensure the following administrative functions:

1. Displaying a list of all active importers together with their status
2. Displaying activities currently under way in a single importer
3. Displaying information on possible errors in indexing/downloading data at the levels of all importers and a single importer
4. Stopping/starting a single importer
5. Listing all conversion tables linked to a given source with highlighting the main conversion table for the Hub indices
6. Replacing any conversion table
7. Downloading any conversion table
8. Switching a selected importer from the sandbox mode to the production mode
9. Checking statistics supported by a given importer (number of records, volume of the index)
10. Configuring and running a new importer by providing required information about the data source and specifying relevant conversion tables
11. Configuring, for selected sources, a frequency and preferred times (range of hours) of their updates
12. Setting authorization keys for applications that use and support APIs
13. Managing keys issued for applications that use Hub services
14. Managing user list: creating, deleting and inactivating users
15. Stopping the whole application or its part and displaying a message about a service break

All the listed functions are obligatory and the Contractor shall provide the Contracting Authority with such tools in the course of work on the project. A graphical interface for managing the specified 15 functions is one of the criteria for offers evaluation.

# TESTS

In tests of applications, the Contractor shall take into account all remarks provided by the Contracting Authority. In the intervals specified in the Schedule, the Contractor shall carry out its own tests of deliverables of the project. This task shall include the carrying out of tests and functional start-up of the metadata Hub OMNIS that is to cover the following areas:

1. Preliminary safety tests (including penetration tests)
2. Function performance and system performance tests, making modifications, and verification based on real search scenarios for operation of all functionalities planned and prepared by the Contractor. Test results may affect further work related to, for example, enhancing formal, subject and bibliographic authority records for the descriptor model in the BN, matching facets, or references to content of documents in the linked access services (Academica, Polona).
3. Cooperation in safety tests
4. Functional and performance tests

The final goal is ensuring a proper quality of the run and produced products as specified below:

1. Installation and launch of production versions
2. Developing procedures of connecting data sources for each type
3. Connecting the target number of external data sources - library systems
4. Launching the API of the Omnis Hub

# SCHEDULE

The Schedule of planned work, as part of the Description of the Object of the Contract that directly enters into the Agreement on providing the described work, is presented in detain in Attachment 1 of the Description of the Object of the Contract.

The constructing of the Hub and its implementation involve a series of analytical tasks that are designed to provide the Contractor with specialist support. These tasks are not covered by this contract and pertain to developing assumptions for conversion tables, data modeling, and testing the system on the side of the user. However, the Contractor shall take them into account in its schedule of work because of products created within such tasks, which will call for adequate reaction in the configuration of the metadata Hub. It is also assumed that the hub will be connected to further data sources right from the very beginning of the implementation and after its completion. The general time perspective for implementation of the object of the contract is as follows:

**Implementation of the metadata Hub - purchase of metadata Hub with licenses**

**from October, 2017 to 31 September, 2018**

Within the above period, the Contractor shall take into account the following products from the tasks meant as technical and information support for the object of the contract:

1. Analysis of metadata from the Libraries’ services in the course of which further data sources are identified and their data models and tables of conversion to the OMNIS format are prepared after which the Contractor’s shall implement relevant configuration and link identified sources to the Hub. This task will be carried out **by March, 2018**, but the possibility of linking the Hub to additional sources should continue for the whole duration of the project and even after its completion and such a possibility has to be ensured by the Contractor.
2. Linking library systems to the Hub in the course of which the Contractor wants to actively participate in the process of linking the Hub to additional data sources specified in the above-mentioned task. The Contracting Authority shall appoint a person(s) responsible for this task and the Contractor shall provide them with any and all information necessary for configuring and managing further data sources so that, after completing the programming work, it is possible for such persons to carry out on their own the process of connecting new data sources, including the Sandbox stage, testing, and production deployment. The task has been in progress **since October, 2017,** or from the very beginning of the programming work, and will continue **until August, 2019.**
3. System tests, for which the Contracting Authority shall ensure a team to carry out functional tests of the system and the Contractor, before commencement of the tests, shall provide a developer version of the system and then take into account, in consultation with the Contracting Authority, remarks and reports regarding the operation of the prepared tool. This task will continue **by April, 2018**, and by this time the Contractor shall deliver a developer version of the system for test purposes.
4. Design of GUIs for the Multibrowser and its mobile version PROXIMA, under which a graphical design and UX assumptions for the Multibrowser shall be developed. The Contractor shall schedule the work so that the graphical project can be applied to the developed software within the specified deadline. Scheduled task results delivery date: **by January, 2018**

Additionally, the Contractor shall comply with the following schedule at the stage of planning execution work for individual elements of the system.

1. Selection of the subset of functions
2. Detailed project analysis
3. Implementation of selected functionalities
4. Testing the prototype with the participation of the final users
5. Evaluating and defining the functional scope of successive iterations

It complies with the project methodology described in section 2.3. In each successive iteration of the task within the work on the metadata Hub and multibrowser, the Contractor shall closely cooperate with the Contracting Authority so that the Contracting Authority is fully informed about completed tasks and the work schedule for the next stage.

# DOCUMENTATION

Following every stage, the Contractor shall deliver a design documentation for tools created under the project together with the full right to use it or a Creative Common Attribution License. It is obligatory for the Contractor to provide documentation that is useful and comprehensible for a broad spectrum of potential users. Within the final stage, the complete provided documentation should undergo evaluation in order to be adapted for the final application design, including all changes and corrections of which the Contracting Authority shall have notified the Contractor.

Documentation shall contain at least a description of the following elements:

1. Installation and configuration of the system and its individual elements
2. Management of individual system tools
3. The source codes of the system and its individual elements with description of software settings and comments of coded elements
4. User documentation of the system API and its individual elements
5. Context diagrams
6. Actor diagrams including main cases of use for delivered modules
7. Architecture of the solution and aspects of integration with other systems
8. Diagram of application services and deployment diagram including specification and all instance level deployment diagrams
9. Service functionalities
10. List of views of the multibrowser system in the form of mockups with description

# GUARANTEE AND WARRANTY

The Contractor shall grant the Contracting Authority a statutory warranty for physical and legal defects and a guarantee for physical defects. The warranty and guarantee period starts from the final, unqualified acceptance of the Object of the Contract and lasts for 24 monts.

The Contracting Authority requests the Contractor to provide warranty services within the warranty period. Warranty services shall be available from 8.00 a.m. to 4.00 p.m., on weekdays, from Monday to Friday, except for statutory holidays in Poland. In order to provide warranty services, the Contractor shall arrange an electronic communication channel with the Contracting Authority whereby possible claims may be lodged.

The warranty covers:

a. Operation of the Software in line with the Description of the Object of the Contract found in Attachment no. 2 of the Contract, in particular section 8. DOCUMENTATION

b. The Contractor’s obligation to provide free of charge software updates and technical support for the maintenance of the software, including integration of new data sources and modifications of conversion tables over the electronic communication channel

c. Under the warranty, the Contractor shall identify and correct faults in the delivered software. A fault is understood as operation non-compliant with the documentation.

The Contractor, within 14 days of being notified of the fault by the Contracting Authority, shall correct the fault together with all its consequences that arouse within this period (records in the database that corrupt data)

The Contractor shall not be subject to warranty liability or liability for damages for consequences of using the provided Software in a computer system operation of which is degraded due to hardware defects

Within the Warranty period, the Contractor shall submit, up to the 7th day of each month, a monthly report with information about work that has been carried out and statistics on operation of the Software (number of indices, monthly increment, disk space usage, processor and memory usage)

# COPYRIGHTS

The Contractor shall ensure that providing the Contracting Authority with the Object of the Contract does not infringe copyrights or other intellectual property rights of any third parties. The transfer of copyrights to the Object of the Contract in the part not covered by a non-exclusive license (if it exists), including each part of it, takes place in all fields of exploitation known on the Contract date, as presented in detail in § 5 of the Contract.

# ESTIMATED DEMAND FOR THE SERVICE OMNIS

In order to estimate demand for the e-service Omnis, data analysis has been carried out and its summary results are presented in the table below.

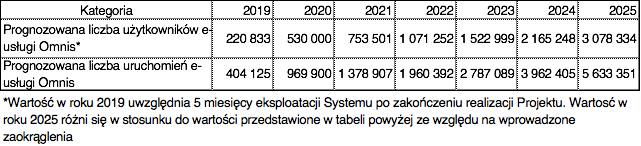


Diagram 18. Estimated demand for the service

The direct benchmark for the analysis in question are the results obtained in the study on the Omnis e-service potential that was carried out by the National Library in 2016 (study titled: Potential of the new e-service dedicated to individual users). In the study, 38% of the respondents declared strong interest in using the e-service. It should be pointed out that the respondents were only persons who had read, bought or borrowed at least one book within last year. Consequently, it was possible to link the obtained rate to the part of the country that reads books. Taking as a basis the preliminary results of the readership study in Poland, as carried out by the National Library, which shows that 36% of people of 15 or more years old had read at least one book in the year preceding the study (1-6 read books: 27.2%, 7 or more read books: 8.4%) the number of readers in Poland is 11,646,345. Before calculating what part of this group will use the e-service Omnis, necessary corrections were made. As only people with access to the internet will be able use the e-service, the number of people without such access was deducted from the total number of readers in Poland. The readership study results (BN 2015) show that as many as one fourth of the respondents has no access to the internet. This result is confirmed by the Central Statistical Office according to which the internet access indicator for Polish households is 74.5% (GUS 2014). This result was reduced by deducting those readers who have access to the internet but do not use it. According to the readership study (BN 2015), there are about 7% of such persons. The level of 38% of the study respondents who strongly declared their interest in using the proposed solution is equivalent to 3,078,632 persons in the whole population.

One should note that rather conservative assumptions were employed. In fact, in the study on the potential of the e-service Omnis as many as 42% of the respondents declared that they would probably use the proposed tool. If the logical approach of the above calculations is applied to this group, the result is nearly 3.5 million persons. This is considerable potential and the planned promotional activities will make it possible to take advantage of it.

The forecasted numbers of users of the e-service and the service launches are shown in the table below.



As a basis for estimating the number of users of the e-service in the years following its activation (August, 2019), statistics on the users of the Polona.pl were taken into account with the assumption that this category of persons is the same as the project target group and falls within the immediate range of promotional activities. As the number of users for the years 2019 and 2020 the number of the current users of the Polona.pl portal was assumed. In 2015, this number amounted to 530,132 (the data from the Google Analytics tool for tracing traffic on web pages, used for monitoring statistics of the Polona.pl portal with the final result adjusted with so-called bounce ratio). As the basis for the number of launches of the e-service a value of 1.83 was used, that is a mean number of sessions per a single user of the Polona.pl portal in 2015.

The number of users will rise every year by additional 45% of users to achieve the estimated level of demand in 2025. This assumption is justified by the scale of promotional activities that owing to connecting the System with libraries all over Poland will have a far-reaching impact.

In their work on the data Hub, the Contractors have to take into account the above data on the estimated number of users.

# PERFORMANCE AND SCALABILITY

The contractor will ensure the performance and scalability of the implemented system, reflecting the provisions of section 11 ESTIMATED DEMAND FOR THE SERVICE OMNIS, provided that at the end of the OMNIS project it is expected to contain 20 million records and the response time for a single query of 10 characters for bibliographic data (metadata) should not exceed 3 seconds.

The contractor will also ensure that the failure or unavailability of any subsystem in the primary infrastructure cannot cause the entire system to stop and in the critical situation the functioning of the Multibrowser module on the basis of replicated data in the backup center is possible.

System scalability is the ability to handle increasing load by attaching more resources (computers, processors, memory). The contractor will ensure that the system can be supplied with extended resources such as physical, virtual and storage servers and memory resources without the need for software modifications.

# GLOSSARY

**Academica** - Digital interlibrary loans system Academica; https://academica.edu.pl

**AtomPub** - The Atom Publishing Protocol; https://bitworking.org/projects/atom/rfc5023.html

digital library polona.pl - Polona is a modern portal that gives network access to collections of the National Library; <https://polona.pl>

**BIBFRAME Ontology** - RDF vocabulary that can be used broadly in the library and cultural heritage environment, initially developed by the Library of Congress to enable experimentation with the BIBFRAME model, further developed by LD4P http://www.loc.gov/bibframe/docs/

**BibTex** - the tool used for formatting bibliographies in line with specified criteria; http://www.bibtex.org

**DBN (National Library Descriptors (Deskryptory Biblioteki Narodowej))** - the indexing and retrieval language used for subject and formal descriptions based on the vocabulary of the previously used subject headings language of the National Library (JHP BN); http://www.bn.org.pl/dla-bibliotekarzy/dbn

**DBPedia Ontology** - the DBpedia Ontology is a shallow, cross-domain ontology, which has been manually created based on the most commonly used infoboxes within Wikipedia; <http://wiki.dbpedia.org/services-resources/ontology>

**dcterms** - DCMI term declarations represented in RDF schema language, http://dublincore.org/documents/dcmi-terms/

**Dublin Core** **(DC)** – the general metadata standard adopted as the standard ISO 15836-2003; http://dublincore.org/

**e-ISBN** - the service designed for quick assignment and management of ISBN numbers and related data; <https://e-isbn.pl>

**EDM OWL Ontology** - Europeana Data Model Ontology, which is accessible through content negotiation but it is also directly available. The Europeana Data Model (EDM) is aimed at being an integration medium for collecting, connecting and enriching the descriptions provided by Europeana data providers. https://github.com/europeana/corelib/blob/master/corelib-edm-definitions/src/main/resources/eu/rdf/edm.owl

**EndNote** - EndNote is a commercial reference management software package, used to manage bibliographies and references when writing essays and articles; http://endnote.com/

**ePUAP** - Electronic Platform of Public Administration Services

**FOAF** - a machine-readable ontology describing persons, their activities and their relations to other people and objects. FOAF is a descriptive vocabulary expressed using the Resource Description Framework (RDF) and the Web Ontology Language (OWL) http://www.foaf-project.org/

**Linked Open Data (LOD)** - is Linked Data which is released under an open license, which does not impede its reuse for free.

**MARC 21** - MAchine-Readable Cataloging) is the format for recording and exchanging data in accordance with the standard ISO 2709:1996 (PN-ISO 2709:1998) used in catalog and bibliographic databases. Data are arranged and uniformed so that they may be properly interpreted by properly designed software.

**OAI-PMH** - Open Archives Initiative Protocol for Metadata Harvesting is a protocol used for downloading and storing (exporting and importing) bibliographic descriptions of records from multiple archives

**ONIX (ONline Information eXchange)** - currently refers to any of three XML standard metadata formats for use primarily within the book trade.

**SRU** **Protocol** – a protocol for retrieving internet search queries using CQL (Common Query Language)

**RDF Store** - A triplestore or RDF store is a purpose-built database for the storage and retrieval of triples[1] through semantic queries.

**SPAQRL endpoint** - A SPARQL endpoint is a conformant SPARQL protocol service as defined in the SPROT specification. A SPARQL endpoint enables users (human or other) to query a knowledge base via the SPARQL language.

**SSO (Single Sign On)** - the option of a single logging in to the web service and getting access to all authorized resources compliant with the service.

**BN systems** - a collective name of the systems Sierra/Millennium, Academica, Polona, e-ISBN

**WCAG 2.0** - WCAG 2.0 (Web Content Accessibility Guidelines)

**Browser OMNIS “Proxima”** - a mobile version of the Omnis browser with the focus placed on finding the nearest copy of the desired book.

**Z39.50** – a protocol for searching and retrieving information from various databases, often remote from the target computer. The protocol is defined in line with the standards ANSI/NISO Z39.50 and ISO 23950. It is widely used by libraries around the world to exchange bibliographic data.

**external data sources** – such systems as Alma (ExLibris), Aleph (ExLibris), Voyager (ExLibris), Prolib (Grupa Sygnity), Mak+ (Instytut Książki), Virtua (Innovative Interfces Inc.), Horizon (SyrsiDynix), Patron (MOL), Libra (MOL), Sowa (Sokrates Software), Mateusz, Koha, Co-Liber and other.

**Integrated Browser, MultiBrowser or Browser OMNIS** - an element of the metadata hub that is responsible for direct communication with the user over the browser interface and allowing one to use simple and complex searching strategies.

**Zotero** - a free and open-source software for managing bibliographies and bibliographic references and organizing data sources for writing scientific publications ; <https://www.zotero.org/>

**other Polish digital libraries** – Polish digital libraries working with DLibra system

# Diagram translations

Str. 12

|  |  |
| --- | --- |
| Biblioteka MARC21 | Library MARC21 |
| Biblioteka DC | Library DC |
| Wydawca ONIX | Publisher ONIX |
| Indeks danych semantycznych LOD | Semantic data index LOD |
| Główny indeks koncentratora danych OMNIS | Main index of metadata hub OMNIS |
| Indeks danych pełnotekstowych FULLTXT | Full-text data index FULLTXT |

Str.15

|  |  |
| --- | --- |
| Pobieranie danych | Data download |
| Zewnętrzne instytucje | External institutions |
| Mapowanie i konwersja | Mapping and conversion |
| Mediator metadanych | Metadata mediator |
| Przechowywanie | Storing |
| Określanie tożsamości i ujednolicenie | Identifying and uniforming |
| Atomizacja danych | Data atomization |
| Moduł wydobywania wiedzy | Knowledge retrieval module |
| Instytucja | Institution |
| Miejsca geograficzne | Geographical location |
| Autor | Author |
| Publikacja | Publication |
| Ujednolicona baza wiedzy | Uniformed knowledge database |

Str. 19

|  |  |
| --- | --- |
| Źródło danych | Data source |
| Kopia pliku mapowania | Mapping file copy |
| Kopia źródła | Source copy |
| Konwerter | Converter |
| Plik | File |
| Konfigurowalne tablice... | Configurable conversion tables allowing one to define data processing method for a newly identified source to an output format, e.g. the universal OMNIS format or other freely defined format |

Str. 22

|  |  |
| --- | --- |
| Przekazywanie danych | Data transfer |
| Konwerter | Converter |
| Tablica konwersji | Conversion table |
| Przekazywanie danych po konwersji | Transfer of converted data |
| Wymiana tablicy konwersji | Exchange of conversion table |
| Testy użytkowników | User test |

# Attachements

Attachement 1: Description current of Polona.pl API page 72 of main Specification

Attachement 2. Description of current data.bn.org.pl (Sierra API) API page 89 of main Specification

Attachement 3. Description of current repozytorium.bn.org.pl API page152 of main Specification

Attachement 4. English translation page

Str. 24

|  |  |
| --- | --- |
| Pobierz metadane | Download metadata |
| czy istnieje powiązanie | Is there a link? |
| tak | yes |
| nie | no |
| dodaj do istniejącej struktury | add to the existing structure |
| utwórz strukturę | create a structure |
| czy struktura ma dodatkowe informacje | Has the structure got additional information? |
| Przekaż info o nowym rekordzie | Transmit information on the new record |
| Uzupełnij dane | Complete data |
| Pobierz nowy rekord | Download a new record |
| czy mam dodatkowe informacje | Have I got additional information? |

str. 27

|  |  |
| --- | --- |
| Wyszukiwanie danych | Data search |
| Biblioteka lokalna | Local library |
| BN | BN |
| Internet | Internet |
| Czytelnik | Reader |
| Czy odnaleziono obiekt poza bibliotekami | Object found outside libraries? |
| Wyszukiwanie | Searching |
| Tak | Yes |
| Nie | No |
| Wytypowanie biblioteki | Selecting a library |
| Czy szukam w kolejnej bibliotece? | Am I to search another library? |
| Zakończenie wyszukiwania | End of search |
| Szukanie w BN? | Searching the BN? |
| Przeszukanie katalogu | Searching a catalog |
| Zasoby sieci Internet | Internet resources |
| Czy znaleziono pozycje? | Position found? |
| Katalog Sierra | Sierra catalog |
| Inne katalogi | Other catalogs |
| Udostępnianie | Making available |
| Czy obiekt zdygitalizowany? | A digitized object? |
| Czy są prawa autorskie? | Are there copyrights? |
| Udostępnianie w Polona.pl | Available on polona.pl |
| Udostępnianie w czytelni BN | Available in the BN reading-room |
| Polona | Polona |
| Czy obiekt jest w Academica? | Is the object in Academica? |
| Czy jest zasób cyfrowy? | Is this a digital resource? |
| Rezerwacja | Booking |
| Udostępnianie w terminalu AC | Available on an AC terminal |
| Udostępnianie w czytelni | Available at a reading-room |
| Wykorzystanie | Using |
| Pobranie treści | Content download |
| Odczyt zdalny (bezpośrednio u czytelnika) | Remote reading (directly at the reader) |
| Podróż do BN (1-500 km) | Travel to BN (1-500 km) |
| Podróż do BL Academica (1-50 km) | Travel to BL Academica (1-50 km) |
| Odczyt w bibliotece | Reading at a library |
| Podróż do BL (1- 1000 km) | Travel to BL (1-1000 km) |

str. 28

|  |  |
| --- | --- |
| Przeszukanie katalogu | Catalog search |
| System źródłowy | Source system |
| Użytkownik | User |
| Utworzenie lub modyfikacja zapytania | Query creation or modification |
| Tak | Yes |
| Nie | No |
| Analiza pozycji z listy | Analysis of an item from the list |
| Wykonanie zapytania | Execution of the query |
| Akceptacja pozycji | Acceptance of the item |
| Analiza dostępnych metadanych | Analysis of available metadata |
| Czy pozycja jednoznacznie poprawna? | Is the item clearly correct? |
| Czy metadane jednoznacznie poprawne? | Are metadata clearly correct? |
| Czy jest dostępny podgląd? | Is a view available? |
| Czy treść zgodna z intencją zapytania? | Is the content in line with the query intent? |
| Podgląd treści | Content view |
| Czy kończę pracę ze źródłem? | End of work with the source? |
|  |  |

|  |  |
| --- | --- |
| Wytypowanie biblioteki | Selecting a library |
| Otoczenie | Environment |
| Użytkownik | User |
| Stworzenie lub modyfikacja zapytania | Query creation or modification |
| Wyszukanie najbliższej biblioteki | Identifying the nearest library |
| Konsultacja bezpośrednia | Direct consultation |
| Udzielenie informacji | Providing information |
| Czy informacja satysfakcjonująca? | Is the information satisfactory? |
|  |  |

Str. 33

|  |  |
| --- | --- |
| Szukaj | Search |
| Okno wyszukiwarki powinno... | The browser window should offer the option of switching to the full-text search |
| Zaloguj | Log in |
|  |  |
| Podpowiedź | Prompt |
| Podpowiedzi dla... | Prompts for items and headings in pattern list (descriptors) that match the entry |
| Po zalogowaniu informacja... | After logging-in, information on new results in previously saved search queries presented as a total number of modified search queries. |
| Podpowiedź hasło/deskryptor | Prompt entry/descriptor |

Str. 34

|  |  |
| --- | --- |
| Szukaj | Search |
| Zaloguj | Log in |
| Autor | Author |
| Tytuł | Title |
| Język | Language |
| Lista wyników | Result list |
| Tytuł dzieła | Title of work |
| Tytuł wydania | Release title |
| Oznaczenie wydania | Release identification |
| Wyróżnienie... | Highlighting of available digital version |
| Pełny opis | Full description |
| Dane powiązane... | Linked data from the semantic engine/LOD |

Str. 35

|  |  |
| --- | --- |
| Szukaj | Search |
| Zaloguj | Log in |
| Autor | Author |
| Tytuł | Title |
| Język | Language |
| Lista wyników | Result list |
| Tytuł wydania | Release title |
| Oznaczenie wydania | Release identification |
| Data wydania | Release date |
| Pełny opis | Full description |
| Tytuł egzemplarza | Copy title |
| Oznaczenie egzemplarza cyfrowego | Identification of digital copy |
| Dane powiązane... | Linked data from the semantic engine/LOD |

Str. 36

|  |  |
| --- | --- |
| Szukaj | Search |
| Zaloguj | Log in |
| Autor | Author |
| Tytuł | Title |
| Język | Language |
| Egzemplarz | Copy |
| Tytuł egzemplarza | Copy title |
| Dane powiązane... | Linked data from the semantic engine/LOD |
| Źródło egzemplarza | Copy source |
| Lokalizacja – dostępność | Location – availability |
| Oznaczenie egzemplarza cyfrowego | Identification of digital copy |
| Informacje charakterystyczne dla danego egzemplarza | Information specific for the given copy |

Str. 37

|  |  |
| --- | --- |
| Szukaj | Search |
| Zaloguj | Log in |
| Egzemplarz | Copy |
| Lokalizacja egzemplarza | Copy location |

Str. 38

|  |  |
| --- | --- |
| Szukaj | Search |
| Hasło deskryptora | Descriptor heading |
| Dane powiązane... | Linked data from the semantic engine/LOD |
| Źródło egzemplarza | Copy source |
| Informacje charakterystyczne dla danego egzemplarza | Information specific for the given copy |
| Autor | Author |
| Tytuł | Title |
| Język | Language |

Str. 42

|  |  |
| --- | --- |
| Moduł główny | Main module |
| Moduł pełnotekstowy | Full-text module |
| Tablice konwersji | Conversion tables |
| Konwersja do formatu RDF | Conversion to the RDF format |
| Ekstrapolacja encji | Extrapolation of entities |
| Pliki RDF | RDF files |
| LD interfejs | LD interface |

Str. 43

|  |  |
| --- | --- |
| Zapytania użytkowników | Users’ queries |
| API dostępowe | Access APIs |

Str. 56

|  |  |  |
| --- | --- | --- |
| Item | Category | Value |
| A | Polish population in total (GUS (Central Statistical Office), 2014) | 38,478,602 |
| B | Population of 15 years old and over (GUS, 2014) | 32,714,451 |
| C | Percentage of persons who read at least one book a year (Readership Study, BN, 2015) | 36% |
| D | Number of book readers (B x C) | 11,646,345 |
| E | Access to the internet in households (D x E) | 8,711,466 |
| F | Number of readers with access to the internet (D x E) | 8,711,466 |
| G | Percentage of persons with access to the Internet who don’t use it (Readership Study, BN, 2014) | 7% |
| H | Number of readers with access to internet who use it (F-(FxG)) | 8,101,663 |
| I | Respondents of the study on the e-service potential strongly interested in using offered tools (Study on potential of e-service Omnis, BN, 2016) | 38% |
| J | Forecasted number of users of e-service Omnis (HxI) in the last year of the reference period | 3,078,632 |

str. 57

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Category | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 |
| Forecasted number of users of e‑service Omnis\* | 220,833 | 530,000 | 753,501 | 1,071,252 | 1,522,999 | 2,165,248 | 3,078,334 |
| Forecasted number of launches of e‑service Omnis | 404,125 | 969,900 | 1,378,907 | 1,960,392 | 2,787,089 | 3,962,405 | 5,633,351 |

\* The value for 2019 takes into account 5-month period of operation of the System after completion of the Project. The value for 2025 deviates from the values presented in the above tables because of rounding.