Biodiv. Res. Conserv. 65: 49-67, 2022

BRC

www.brc.amu.edu.pl

DOI 10.2478/biorc-2022-0002

Submitted 10.02.2022, Accepted 31.03.2022

The Adam Mickiewicz University Nature Collections IT system (AMUNATCOLL): portal, mobile application and graphical interface

Maciej M. Nowak[1,3]*, Marcin Lawenda[2], Paweł Wolniewicz[2], Michał Urbaniak[2], Bogdan Jackowiak[1,4]

[1]Department of Systematic and Environmental Botany, Faculty of Biology of Adam Mickiewicz University in Poznań, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland; ORCID: BJ https://orcid.org/0000-0003-1684-7380; MMN https://orcid.org/0000-0003-0005-196X
[2]Poznań Supercomputing and Networking Center Jana Pawła II 10, 61-139 Poznań, Poland; ORCID: ML https://orcid.org/0000-0003-4844-3655; MU https://orcid.org/0000-0002-4493-293X; PW 0000-0003-2801-5303
[3]Biological Spatial Information Laboratory, Faculty of Biology of Adam Mickiewicz University in Poznań, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland
[4]Biodiversity Digitization Laboratory, Faculty of Biology of Adam Mickiewicz University in Poznań, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland

* corresponding author (e-mail: mcnowak@amu.edu.pl)

Abstract: The paper describes the interfaces implemented in the AMUNATCOLL IT system, which enable access to and exploration and manipulation of data available in the database containing unique natural collections from the Faculty of Biology of Adam Mickiewicz University in Poznań (FBAMU). Data can be accessed using the two available interfaces: graphical and programming application interfaces. The first is implemented in two forms: a portal, which is the main interface for accessing the data stored in the database, and a mobile application that complements functions related to field research and creating private collections. To deliver the required set of operations, the portal was equipped with simplified and advanced searching, statistical analysis and spatial processing (BioGIS). Data openness and the ability to collaborate with other solutions and systems are key elements in achieving synergies in conducting research on biodiversity. AMUNATCOLL IT offers an opportunity to respond to these challenges, enabling data export for independent processing with external tools related to portal functionality or giving access to data directly using an application programming interface. Graphical interfaces are subject to numerous requirements and restrictions reflected in the graphic design and accessibility issues related to the accommodation of disabled individuals. These interfaces must properly address both groups of target recipients, considering their different goals and level of knowledge, as well as adjusting the level of interaction due to the limitations of using the interface.

Key words: IT system on natural history collection, digitisation, biodiversity database, programming interface, graphical interface, API, REST

1. Introduction

The dynamic development of information technologies enabled the launch of the process of digitisation and the wide availability of large amounts of information on biodiversity. Much of this information has been systematically collected for several hundred years by several scientific institutions around the world. As a result, millions of specimens from all groups of organisms have become visible in every corner of the world that the internet has reached. The Faculty of Biology of Adam Mickiewicz the University in Poznań joined this global trend through the AMUNATCOLL IT project. The project is carried out in joint cooperation with PSNC supporting design and conceptual errands with a particular emphasis on IT tasks. The overall purpose of this endeavour is to meet the world's standards in digitisation and to share information on biodiversity (Jackowiak *et al.* 2022). This publication aims to present the AMUNATCOLL IT system with a special emphasis on the implemented interfaces and their purposes, types and characteristics.

Table 1. Comparison of functionalities offered by exemplary NHC database systems

Functionality	Global Biodiversity Information Facility	The herbarium at the Royal Botanic Gardens Kew	Muséum national d'Histoire naturelle in Paris	Tropicos® database	Natural History Museum
Browsing	+	-	-	+	+
Simple search	+	+	+	+	+
Advanced search	+	+	+	+	+
Specimen comparison	-	-	-	-	-
Georeference	GPS coordinates	Region or GPS coordinates	Region	+	Up to the region
GIS tools	Average	-	-	Average (geographic specimen search)	Average
Iconography	+	+	+	+	+
User owned database	+ (for testing reasons)	-	-	-	-
Mobile application	-	-	-	-	-
Publication reference	-	-	-	+	General reference to the type in the publication
Profile	Open access to data about all types of life on Earth	Wide range of plant specimens representing 98% of the genera in the world	Vascular plants	Comprehensive overview of species, taxonomy, images, references, DNA, etc.	Entomology, zoology, botany, palaeontology, mineralogy
Number of specimens	Almost 2 billion	Ca. 7 000 000 specimens	Over 5 500 000 specimens	Over 5 000 000 specimens	Almost 5 000 000 specimens
Notes	The biggest collection of Earth life specimens based upon external database contributions	Herbarium digital images of specimens: 609 197	Linked to GBIF, JSTOR and Tropicos	Over 930K digital images Linked to Plant List and International Plant Names Index	Linked to Biodiversity Heritage Library, Catalogue of Life, GBIF

AMUNATCOLL IT offers multiple ways of interacting with digitalised data through a portal, mobile application and an application programming interface (API 2021). These applications are based on the resources and the specificity of their description provided by Adam Mickiewicz University in the taxonomic database (stores specimen descriptions) and the iconographic database (contains multimedia materials).

Considering the graphical interaction perspective, the assumed form of the data presentation and processing includes both the portal and the mobile application. The portal is the primary means of access to data sources available in the database, while the mobile application is intended to provide supplementary functionality for facilitating fieldwork and the creation of user-owned databases. The design assumptions of the AMUNATCOLL IT project that provide unrestricted access to data based upon FAIR principle (FAIR 2021) complementary solutions were introduced. Within this strategy, exporting data in open format by means of portal functionality and providing direct access to data via programming interfaces constitute the fundamental range of services offered by the system. The implementation of the human interface is associated with taking into consideration the numerous conditions in the area of provided functionalities, transparency and facilities for people struggling with various physical and mental dysfunctions. This means that the way in which communication with users occurs should consider different levels of knowledge and intentions as well as generally

Smithsonian National Museum of Natural History	Integrated Digitized Biocollections	AMUNATCOLL
Quick browse of selected specimens	+	+
+	+	+
+	+	+
-	-	Graphical comparison
Region or GPS coordinates	GPS coordinates	GPS coordinates
-	Average	Advanced
+	+	+
-	-	+
-	-	+
-	-	+
Anthropology, botany, entomology, mineral sciences, paleobiology, invertebrate and vertebrate zoology	Mostly Animalia and Plantae; minority of Fungi Chromista, Protozoa, Eubacteria and Algae	Botanical and zoological specimens
10,691,864 specimen records	Nearly 130 million	Over 2 000 000 specimens
Only part of the Smithsonian NMNH collection is digitized. 1.5 million records georeferenced.	1704 datasets, ambition to integrate all U.S. biological collections	Linked to GBIF

accepted standards for the design of interfaces for people with disabilities.

2. Interfaces of online databases of biodiversity information: an overview of selected websites

Many systems on the internet related to the presentation of biodiversity information can be encountered. They differ in the purposes and scope of functionality offered to their users. As part of the work on determining the functional scope of the implemented interfaces, we learned the experience of renowned scientific institutions and IT databases, with a particular emphasis on the online databases, which are one of the most widespread within the scientific users.

The first example is the Global Biodiversity Information Facility (GBIF 2021), an international research network and infrastructure funded by governments around the world dedicated to providing open access to biodiversity data on Earth. The next example is the Kew Herbarium Catalogue of the Royal Botanic Gardens Kew (KEW 2021), which publishes an online digital herbarium catalogue containing scans of specimens and their metadata. Another example consists of the plant collections of the Muséum National d'Histoire Naturelle (Sonnerat) (MNP 2021; PVM 2021), which provides digital access to collections containing specimens from France and Europe, as well as from the former French colonies and the present overseas territories of France. In addition to herbarium collections, a digital botanical library of scientific publications is also made available online by Tropicos (TRO 2021), one of the largest online databases of plant information run by the Herbarium of the Missouri Botanical Garden. The Natural History Museum in London (NDP 2021) not only runs an exhibition but also offers a large set of collections (5 million specimens). Records are organised into five collections: entomology, zoology, botany, palaeontology, and mineralogy. The Smithsonian National Museum of Natural History (SMI 2021) has one of the most impressive collections of records originating from seven departments (anthropology, botany, entomology, mineral sciences, paleobiology, invertebrate and vertebrate zoology). Vertebrate zoology has special divisions for amphibians & reptiles, birds, fishes and mammals. In total, it constitutes almost 10.7 million specimen records currently available online. The overall goal of Integrated Digitised Biocollections (iDigBio 2021) is to develop infrastructure that supports the advanced digitisation process of biodiversity collections. Its goal is to integrate all U.S. biological collections by enabling digitisation to facilitate the discovery and increase the understanding of biological diversity. This process is combined with the use of edge standards and practices in the field of the digitisation process itself and building a cloud environment for storing and sharing the collections. Currently, iDigBio hosts nearly 130 million specimens and more than 43 million media records organised in 1700 datasets. Most of the specimens originate from kingdoms such as Animalia (48%) and Plantae (45.9%), with many fewer from the Fungi (5.3%), Chromista (0.5%), Protozoa kingdoms and Eubacteria and Algae domains. Data collected in iDigBio can be explored natively through the portal interface as well as using an API to directly access the data and utilise functions that are not available through the portal.

The creation of the AMUNATCOLL IT interfaces was preceded by a comparison of the website interfaces of the above databases and considered the following categories (Table 1):

- Browsing – simple browsing of the database content without using a search function; it can be either the usual functionality that allows users to freely navigate through the collections or driven by specific categories.
- Simple search – search using one field solution.
- Advanced search – search with multiple fields defined; it is advantageous if logical operations are able to be implemented.
- Specimen comparison – ability to compare two or more specific specimens at one page; the comparison can be implemented either by a generation collation table where specimen data are presented (metadata comparison) or by the juxtaposition of pictures (graphical comparison).
- Coordinates of specimen locations – information about geographical referencing along with precision data; data may be given with varying accuracy (continent, country, region, descriptive locality, GPS coordinate).
- Spatial analysis (GIS) tools – if a database has a functionality to show up specimens on a map, the functionality of the spatial analysis tools can be divided into three categories: elementary (just map representation), average (simple manipulation of data represented on a map by filtering), advanced (advanced search and data manipulation with GIS functions).
- Iconography – if a specimen description is accompanied by multimedia data.
- User-owned database – the ability to create datasets (databases) owned by the user.
- Mobile application – if the system is facilitated by a mobile application.
- Publication reference – whether there is an association between the specific specimen and the publication(s) where it is mentioned.
- Collection profile – information regarding whether the collection is specialised with specimens belonging only to a specific group or if it is a multicategory collection. The exemplary categories include entomology, zoology, botany, palaeontology, mineralogy, anthropology, mineral sciences, palaeobiology, invertebrate zoology, amphibians and reptiles, birds, fishes, mammals. etc.
- Number of specimens – information about the number of records stored in the database.

The comparison of those systems is a very cumbersome process, mainly due to different approaches to the offered functionalities resulting from different expectations and needs of target users as well as the level of technological advancement and IT solutions used at the date of their implementation.

Basic offerings, such as browsing and searching in both simple and advanced ways, are provided by almost all databases, which is natural and expected. The comparison of at least two specimens at one page is a rather unique feature. This operation seems to be particularly useful when a detailed analysis is carried out to detect differences between the digitised objects. Georeferenced information is commonly provided across investigated systems but is subject to the accuracy of the information provided. In some cases, only the region where the object was harvested is provided, while in other solutions, more accurate GPS coordinates are mentioned. However, in all cases, we cannot transfer these coordinates to the map; these tools also differ in the possibilities of available data manipulations while presented on the map. Iconography (scans and pictures) is omnipresent, and the only difference is the quality of the photos offered. The presence of such services as the possibility of building a private database based on one's own observations made using a mobile application is now extremely rare. In our comparison, this is only offered by the AMUNATCOLL IT system, which we consider important, as this type of functionality potentially increases social commitment and sensitivity to biodiversity issues. The next functionality relates to the publication references, which is quite useful when one wants to read more about the specimen presented in the database. It is sporadically encountered in the analysed systems and presented in two ways: the references refer either to a given specimen or generally to a given species in publications.

In summary, it is worth mentioning that many databases closely cooperate with other systems, which is very beneficial due to the completeness of the offered data.

3. Portal

The portal is the first-choice tool to obtain access to the database for users of all kinds. Essentially, the functionality of the portal includes data search, viewing the list of specimens, displaying information about the specimen, presenting aggregate information about specimens, and viewing iconographic data. In addition, the portal has integrated tools for coordinators that enable data editing and checking usage statistics, such as number of visits, number of users, and number of data downloads.

The user interface allows efficient execution of typical scenarios of working with the portal. These are obtaining general information about AMUNATCOLL and the natural history collections database, retrieving information from the database, applying analytical tools, obtaining more information about the portal, obtaining assistance in using the portal, and contacting the authors. The main portal menu reflects these typical usage scenarios. Profiling the access methods and displaying data is aimed at adjusting the scope and the functionality to the requirements of target groups, e.g., state and local government units, state services and officials, teachers, students, research workers and doctoral students, nongovernmental organisations, and all other nature fans.

3.1. Data exploration

The portal allows searching by a variety of criteria and using an interface tailored to different needs and capabilities of the target audience. A simple, general search engine is provided on the main page of the portal, is available to all users, and is unregistered. It searches for elementary information on specimens based on scientific names of species and genera of plants, fungi and animals, offered in Polish or Latin.

3.1.1. Specimen and taxa browsers

The form consists of a single field solution with a value that is used to search for specimens by the following fields: species, genus and higher taxa. The result of the search is a list of specimens, which contains up to several thousands of items. The list of retrieved specimens is presented in a few tabs: (1) Genera – retrieved specimens grouped by genus; (2) Species – retrieved specimens grouped by species; (3) Specimens – a list of individual specimens; (4) Iconography – scans of specimens identified taxonomically; (5) Multimedia – other images of nature that are unrelated to the stored specimens.

The results are presented as a paged list. The number of items per page is the user's choice. Detailed information about each specimen in the list can be displayed (Fig. 1). Information about the retrieved specimen is grouped under several tabs. When a photo is associated with a specimen stored in a collection, it is displayed on the left side of the screen. The picture can be enlarged or displayed in full-screen mode. If there are more scans or photos of a specimen, they can be viewed one at a time. The right side of the screen presents information about the specimen. There is also a preview of the location on a map where the specimen was harvested and the information about citations in scientific papers. The list of specimens can be downloaded as a file in XLS, XLSX or CSV format with the functionality that is enabled for log-in users only.

Complementary to the simple search engine, specialised engines are also accessible for logged users and offer three search perspectives: (1) Simple – selected number of fields with the possibility of defining the search area on the map; (2) Extended – enables the selection of a number of fields of user interest where fields can be defined as inclusive and exclusive (Fig. 2); (3) Advanced – allows the user to freely define the searched fields and combine them into logical AND, OR and NOT formulas.

In the specialised search, only Latin names of plants, fungi and animals and higher systematic units are examined. The specialised engine results are

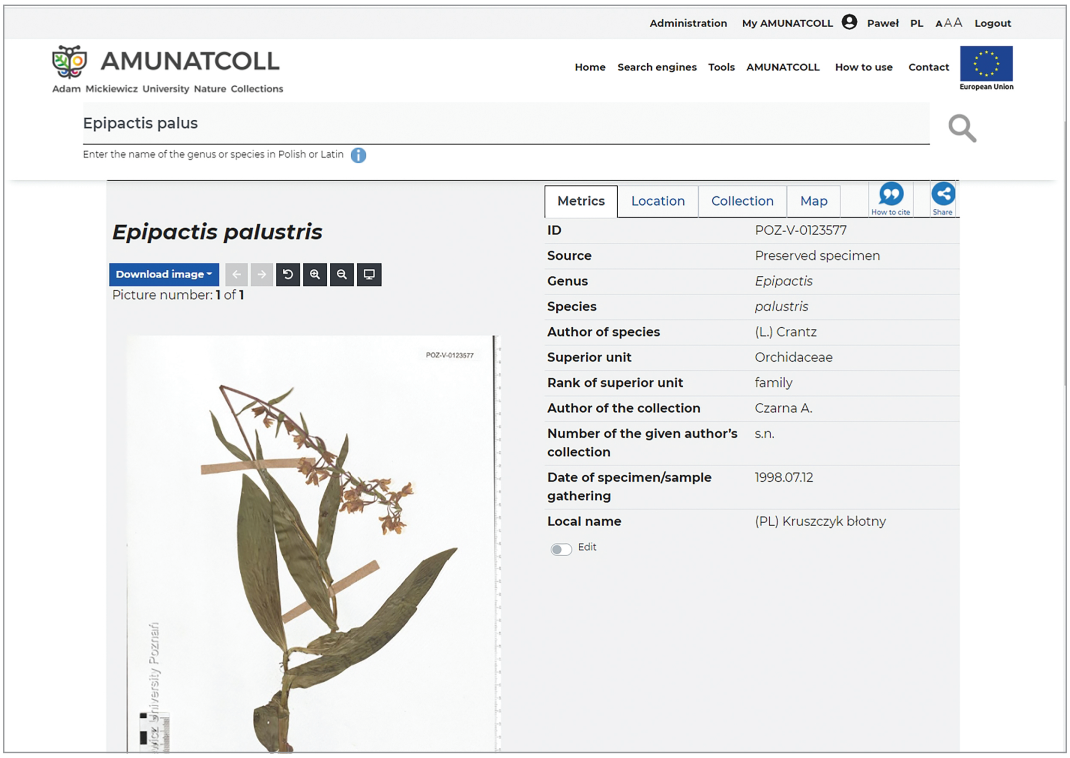

Fig. 1. The specimen details view

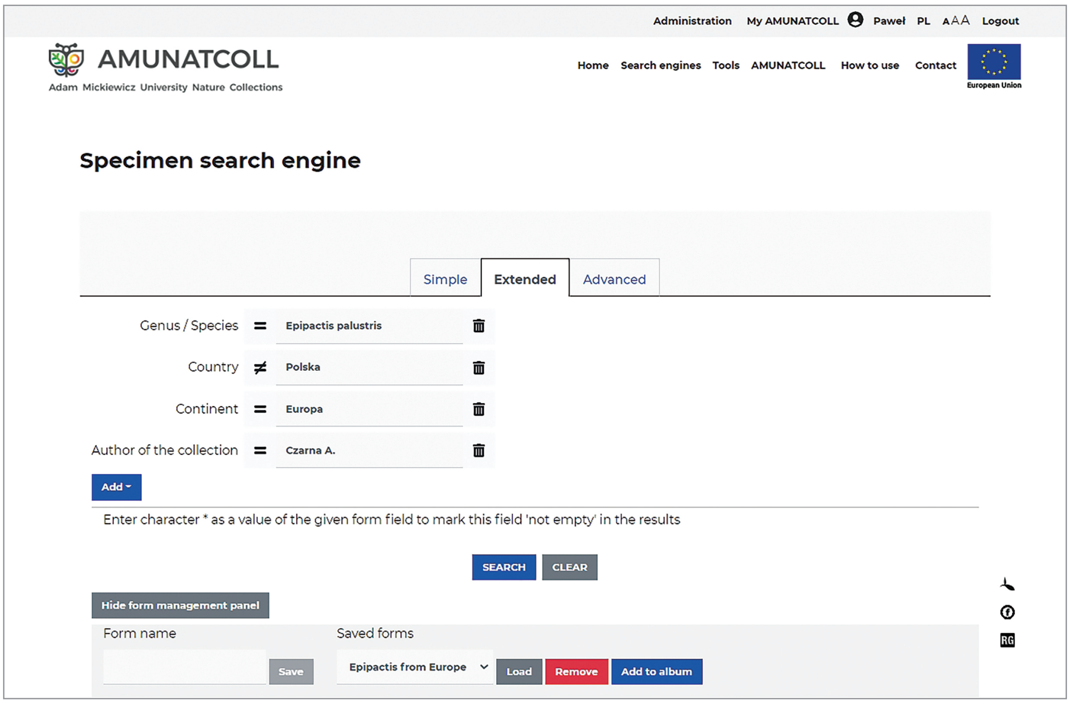

Fig. 2. Extended search engine view

similar to those of the general search results, but two additional tabs are presented: (1) Map – cartographic presentation of the found specimens on a world map; (2) Report – presentation of search results in aggregate form.

Another way of searching for data is to browse the database by systematic classification or by collections; sets of specimens are organised by systematic groups, region of origin, or collector's name by the FBAMU. The systematics view presents a tree of taxonomic units along with the number of specimens of a given unit in the database. This allows easy orientation in the content of the database and access to the specimens in the selected branch of the systematic tree. The collection view is organised in a very similar way; after selecting the type and kind of a collection, one can browse the specimens that are part of it.

3.1.2. Multimedia browser

Multimedia data in the form of images, films, maps or sound recordings may be associated with the deposited specimens (e.g., a scan of a herbarium page) or may be independent of the specimens (e.g., a photo of a habitat) and form a multimedia database complementary to the deposited specimens. Searching for multimedia materials is possible using the following fields: title, author, photo number, country, protected area, and date. The photos are presented in the form of thumbnails, and once selected, the large photo is shown on the right side of the screen with related information below (Fig. 3). The photo can be zoomed in to view the details. All attributes of the media material are presented next to the photo. It is possible to select any pair of photos from the database and compare them by displaying them side by side.

3.1.3. Remaining search engines

In addition to specimens and multimedia files, the collection database also includes information on samples (e.g., soil samples), bibliography, and educational presentations.

Soil is the habitat of many different species of animals that live both in the ground and on the ground. The result of soil sample testing is the identification of the specimens about which the information goes to the taxonomic database. There are still unidentified specimens in the soil samples that may be tested in the future. The advanced search engine is used to search such soil samples, which enables the selection and specification of the selected fields.

The AMUNATCOLL IT database enables the storage of information on bibliographic sources in which data on specimens from the FBAMU collection were cited. A dedicated search engine was prepared for searching this part of the database. The search can be carried out according to the following criteria: citation, authors and editors, year of publication, title of publication, keywords and publication type.

A very interesting solution used in AMUNATCOLL IT is the possibility of generating educational

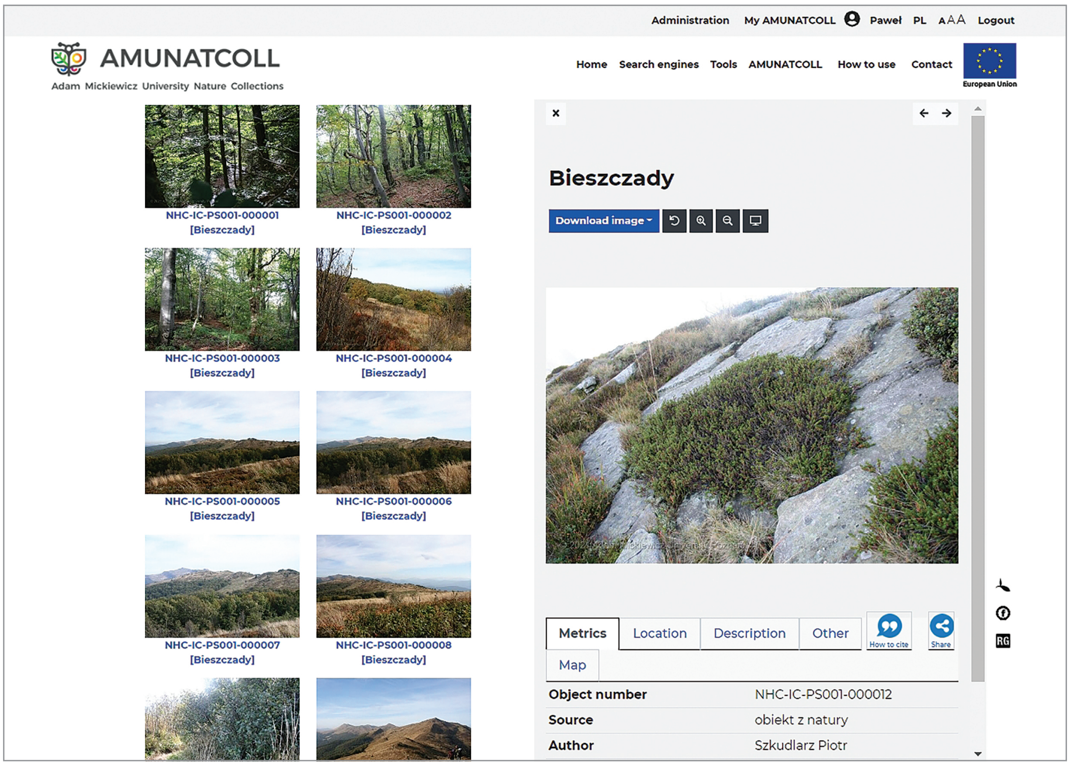

Fig. 3. The result of the multimedia search view

presentations. They are created on the basis of one or more specimens from the database with the possibility of adding a related description. Educational presentations explain the biological aspects of a given group of organisms and are intended for different groups of users, including students, officials, and nature lovers. The links to the presentation are gathered on one or more pages, enabling browsing and selecting the appropriate one.

3.2. BioGIS tools

Based on the textual description of the specimen locations, all records have been geotagged; i.e., their geographic coordinates are determined (Nowak *et al.* 2021). This enables spatial analysis using GIS tools built into the portal. Seven BioGIS tools (from the Geographical Information System of Biodiversity) have been implemented in the AMUNATCOLL portal. This solution enables users to perform online geoprocessing. Each devised tool considers the specimen location represented by a point and brings the outcome map.

The map functionality is supported by a toolbar consisting of the following functionalities: (1) a specimen search engine as a required step for map preparation, (2) map settings enabling the personalisation of the map style and the use of labels for specimen localities on the map, (3) map description (i.e., map title and author). Moreover, apart from records of the AMUNATCOLL IT database, the user is welcome to show their own observations collected in the field experiments with the AMUNATCOLL IT mobile application on the map. This functionality is available in the separate part of the portal called "My observations." Furthermore, the "Thematic Layers" toolbar functionality makes it possible to display predefined layers, such as a UTM grid, countries, and nature conservation areas. The "User layer settings" functionality is dedicated to personalising the style of point, polyline and polygon layers, which can be added on user demand as an additional space for creating new features on the map. All layers created by the user can be exported in the GeoJSON format. The user is allowed to import his or her own spatial data through KML (2021) format and add new Web Map Services (WMS 2021). These two functionalities are called "Save layers/Add data." When all steps leading the outcome map are applied, the "Save Map" template functionality can be used. The last functionality, "Manage map tools," enables the display or hiding of useful features, such as map scale windows, zoom slider tools, basemap managers, specimen labels, and map exports for filing or drawing tools.

Several types of map preparation can be distinguished depending on the purpose they are used for. Therefore, a few methods were proposed for visualising the AMUNATCOLL IT database data and user observation data. The first approach to map

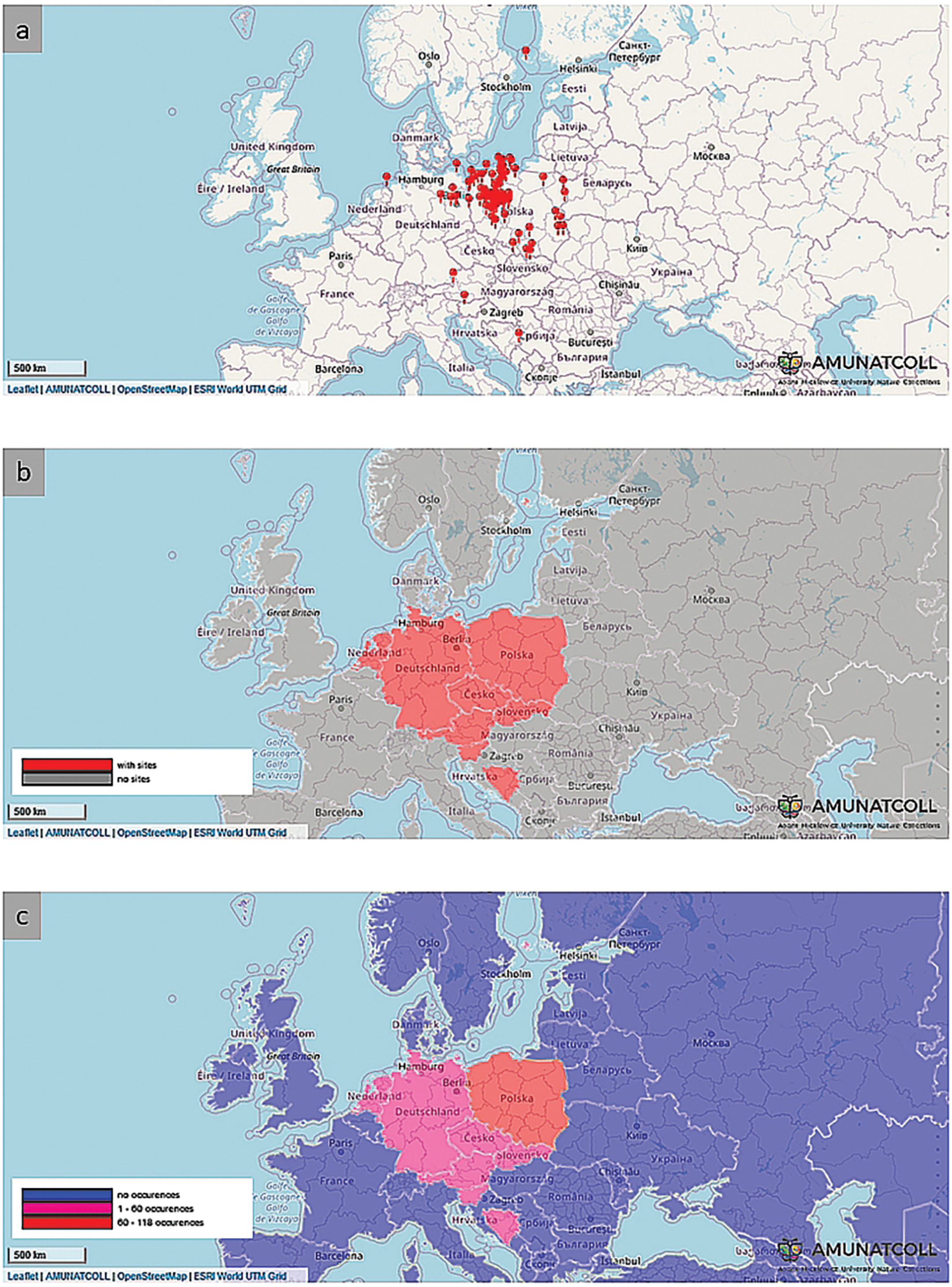

Fig. 4. Maps of the *Epipactis* genus performed by BioGIS tools

Explanations: a – dot distribution map, b – area class map, c – choropleth map

preparation, dot distribution, allows the visualisation of the spatial distribution of each specimen location regardless of the scale level used (Fig. 4a). The second tool, the area class map, allows users to present the distribution of specimen locations using the approach of the presence/absence data within several types of base areas, such as the UTM grid or countries (Fig. 4b). The third tool is the choropleth map, which enables users to prepare the distribution maps where a set of predefined areas is coloured in proportion to a statistical variable that represents an aggregate summary of a geographic characteristic within each area, e.g., the number of specimen individuals (Fig. 4c). A more extensive version of the choropleth map presents the diagram map as the next example of BioGIS tools. This tool makes it possible to use the pie charts within the

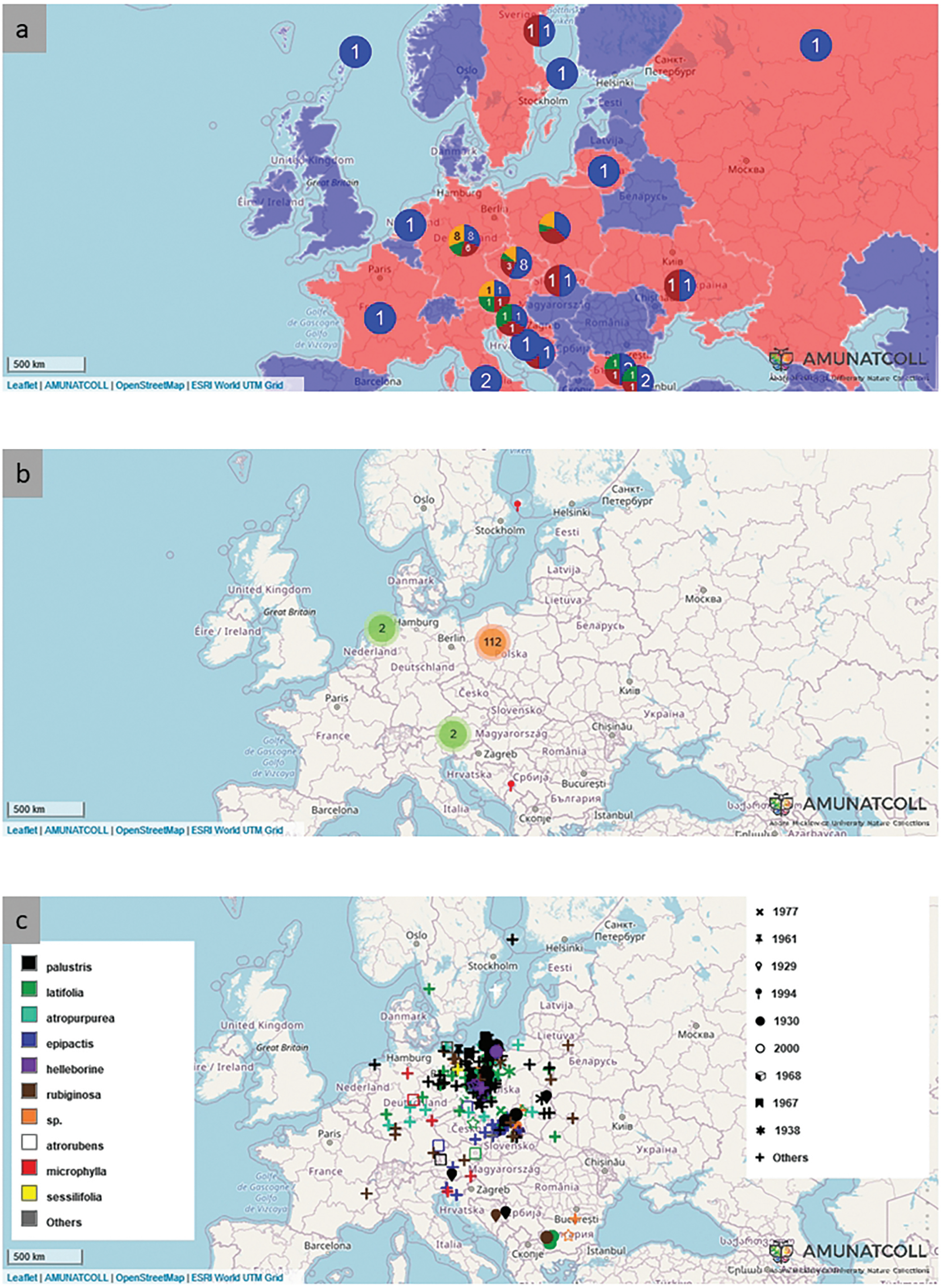

Fig. 5. Maps of the *Epipactis* genus performed by BioGIS tools
Explanations: a – diagram map, b – cluster map, c – attribute grouping map

choropleth map (Fig. 5a). The consecutive tool, the cluster map, enables a presentation of specimen localities as clusters of points. The size and the number of clusters depend on the abundance and spatial scale applied to the map (Fig. 5b). The opportunity to visualise the distribution of specimens using their attributes is possible because of the tool calling the attribute grouping map (Fig. 5c).

3.3. Statistical analysis

The statistical analysis tool enables users to present their search results in an aggregated form considering that the report aggregation can be made by any defined field. The result is presented as a table showing the percentage distribution of specimens against the selected field. It is allowed to define multiple aggregation fields.

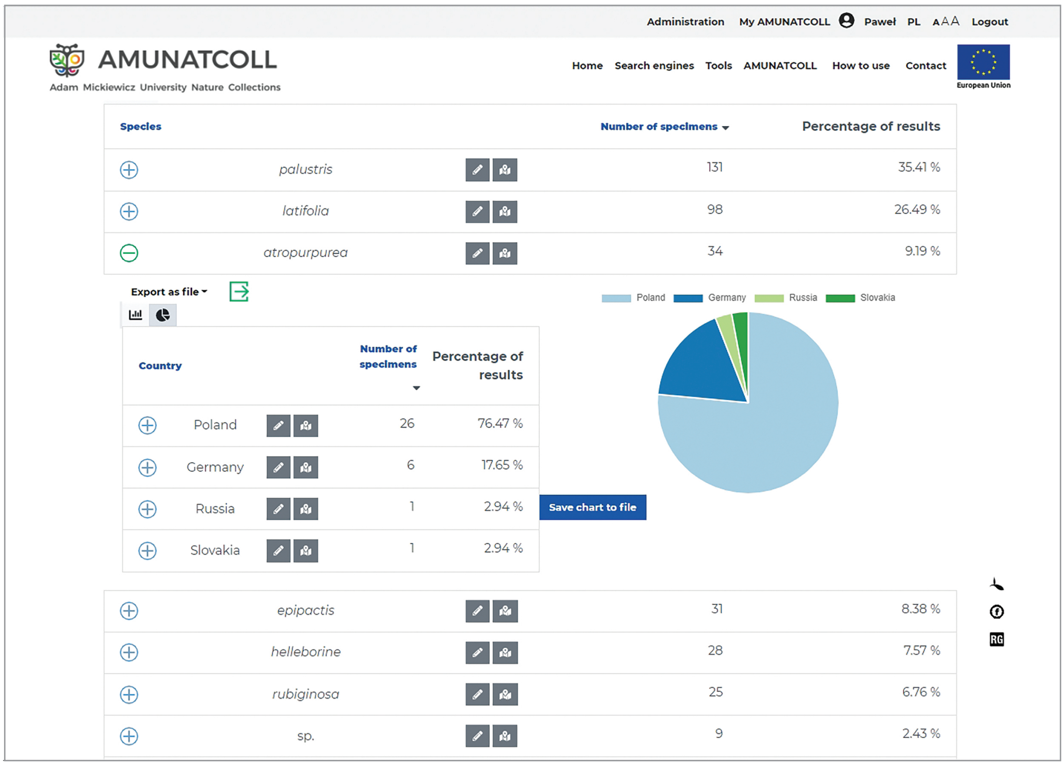

Fig. 6. Statistical analysis results for the genus *Epipactis*

The report against the next aggregation field is shown as a consequence of selecting the downstream aggregation row. It is possible to present the results as a pie chart or a bar chart (Fig. 6).

3.4. Exporting data

The portal data are made available on an open access basis. Users have access to up-to-date, open, searchable data, with the exception of data that are related to the physical storage organisation of specimens. Wherever required by law, the exact location of the collection site is also restricted (e.g., for protected specimens). It is worth noting that data are made available using an open data format for registered users only. The data that can be retrieved by users using the portal including the following: (1) a list of taxa resulting from the search in CSV, XLS or XLSX format. For performance reasons, data export is limited to 10000 records; (2) a list of taxa from aggregated reports in CSV, XLS, or XLSX format; (3) prepared maps in PNG format; (4) user layer points on the map in KML or GeoJSON format; (5) photographs in JPG format; (6) other multimedia materials, e.g., audio or video files; (7) album-based presentations in PDF format.

The ability to download data in an open format enables processing using other, unrelated software of user preference. This may be highly desirable when the user uses a different application on a daily basis or when it is required by the nature of the work.

3.5. Data modification

The AMUNATCOLL IT portal integrates functionalities for users and for data administrators. Users with the data editor role are able to use the portal to generate new specimen records and modify them. It is possible to add a new record or to import a set of records from exported spreadsheet files. The form of specimen details in view mode can turn into edit mode, where a modification of attribute values is feasible of a single record straight in the portal (see edit switch in Fig. 1). Edit mode is also available in the statistical results view, where the editor can correct attributes for many records at the same time. Another option for a mass update is a spreadsheet import tool used in update mode. All specimens adding and modifying operations go through a validation procedure to verify that the data are compliant with the metadata specification.

3.6. Portal administration

The user with the highest privileges is allowed to use the administration interface of the AMUNATCOLL IT portal. This includes access to statistics about the use of the portal, tools for managing user roles, and tools for importing and editing data. For all data types in the AMUNATCOLL IT database, including specimens, samples, multimedia materials, and bibliography, it is possible to edit any field describing a given record. For the sake of correcting values for multiple records, one is able to use the collective editing mode. In the "Analytical Reports" tool, the offered functionality enables changing in the search results the corresponding value of the aggregation criterion. It permits a specific value of many records, for example, a typo in the name of the author, to be corrected at a single step.

4. Mobile application

The idea of the mobile application was devised based on the analysis of the needs of target groups, where the overriding value was to increase the involvement of target groups in the interaction with the AMUNATCOLL IT system. The developed concept allows you to create and save observations conducted in the field and export them to the AMUNATCOLL IT database. This information becomes available from the portal level, which in turn allows for their editing and the use of existing analytical and georeferencing tools. Using this approach allows increasing the involvement of target groups, as well as reaches external users who utilise the AMUNATCOLL IT system by creating their own collections of specimens and other natural observations.

Data collected in the field, e.g., information on the locations of plants, fungi, and animals, are described with a birth certificate and documented with a photo. After conducting the export procedure, the data are stored in the AMUNATCOLL IT system in the user's private space, which ensures their separation from the main database containing only verified information about specimens and natural objects deposited at the FBAMU. Data from a private database are available only to their owner; however, after consent is given and the appropriate verification procedure is completed, they can be included in the AMUNATCOLL IT database. This is possible because the structure of the data collected in the mobile application is consistent with the target AMUNATCOLL IT metadata structure.

4.1. Creating a project

The mobile application uses the principle of a project that allows you to group your observations within one set, which favours better organisation of the collected information. The project itself is equivalent to a directory in the operating system. As part of the created project, the application user adds further observations. The scope of the fields to be completed is defined at the initial stage of creating the project. The field can be a simple text input field, but it can also be a more complex field, e.g., an editable selection list or a field defining a location of the object.

The first operation the user needs to do is to create an empty project, name it, select the number of applicable fields, which they can edit later on when gathering information on the observed object, as well as the optional default location for the observations in the project (Fig. 7a). The fields for a detailed description of the object

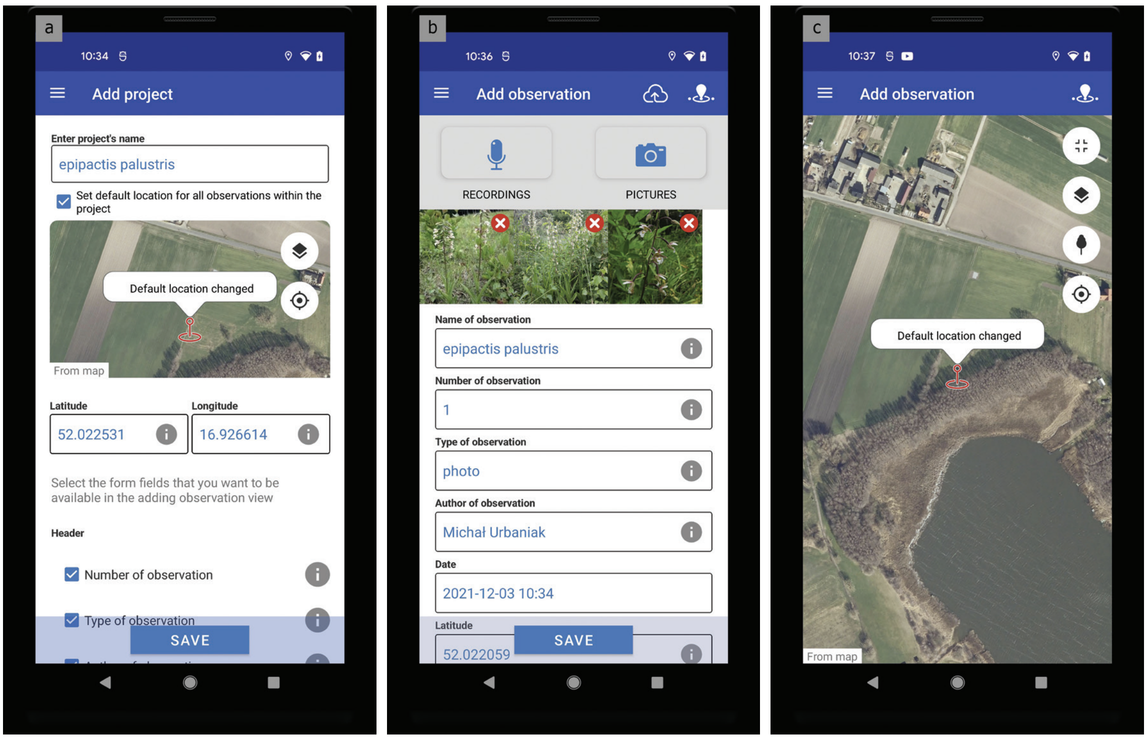

Fig. 7. The process of creating field observations using the AMUNATCOLL IT mobile application

Explanations: a – project configuration as an initial stage of creating observations, b – creating an observation with attached photos, c – using the map component to set the exact location of the object

Table 2. Parameters and tooltips of the observation form in a mobile application

Field name	Tooltip
Observation name	Name of observation assigned by author
Observation id	Observation id assigned by author
Type of observation	Collected specimen / photo / note
Author of observation	Possibility of adding successive persons
Date	Date of observation
Longitude	Longitude in decimal notation
Latitude	Latitude in decimal notation
Altitude	Height above sea level in meters
Location	e.g. name of locality, street; distance from road, lake
Habitat	e.g. deciduous forest; pine forest; moist meadow, xerothermic sward
Area of observation	e.g. area determined on the basis of standard measurement or derived from tool measurement “measure area”
Plant/vegetation coverage	e.g. percentage coverage of tree, bush layer, layer of herbs as well mosses and lichens
Remark	Working area of author of observations
Species name	e.g. common pine, *Pinus sylvestris*
Quantity	e.g. 10% or 2.2 in Braun-Blanquet scale
Gender	e.g. female, male individual, monoecious
Age	e.g. 5 years
Phase	e.g. nymph, larva, seedling
Dimensions/Measurement	e.g. 33 cm
Natural habitat	e.g. Natura 2000 habitat with code
Other objects	e.g. brickyard, church
Persons	Names of persons shown on photograph
Description	e.g. photograph title, key words describing content of photo
Sample number	A figure used to show the position of a sample in a series
Location/plot	Exact place where samples were collected
Type of sample	Qualities and features of the sample
Method of collection	The way in which samples were collected
Merocenosis	A part of a biocenosis consisting of a group of organisms inhabiting a certain section of the natural environment
Type of merocenosis	Qualities and features of merocenosis
Merocenosis host	An animal species whose body and habitat are a merocenosis inhabited by other organisms
Merocenosis position in relation to the ground	The distance and positioning of an object relative to the ground surface
Merocenosis type of bird nest	The characteristics of the bird nest
Dead wood merocenosis: extent of decay	The degree to which dead wood has already rottened
Merocenosis diameter of tree trunk	The measurement of the tree trunk in the centre
Merocenosis length of tree trunk	The measurement of the tree trunk from one end to the other
Merocenosis species of tree	The typical characteristics of the tree
Sample humidity	The amount of water in a sample
Additional information	Supplementary description

are divided into 4 sections: common fields (aka header), taxonomic fields, fields related to habitats, landscapes, phenomena or natural processes and research trial fields. A corresponding tooltip is prepared for each field to guide the user to properly fill in a certain field (Table 2).

4.2. Gathering information on observed objects

The interface of the mobile app has been designed to meet the user’s basic expectations working in the field. In the AMUNATCOLL IT mobile app, the user is requested to fill up a minimal, desired number of attributes (Fig. 7b). Some fields of the form may be completed automatically, e.g., date and author of the observation, observation ID or GPS position. It is also possible to automatically rewrite data repeated in subsequent observations.

The app provides customised controls or entire views to store even more complex data types in the easiest way. To exemplify how location coordinates of observed objects are delivered to the system, let us follow regular user procedures. As mentioned above, the position is one of the properties automatically detected and saved. Sometimes we face the situation when the current position that the user (mobile device) is located in does not correspond to the position of the observed object. In this case, the user needs to point the location on the map using the full screen map component. There are two options supported by the app to find a solution: use the default project location if it was set while the project was created and read the position from the EXIF metadata of the photos. It is a user's choice where one of the location sources is most accurate for a given case. The third attainable solution is to tap on the map and mark the location where the object was observed (Fig. 7c). If a more advanced GPS source (e.g., an external device) is available rather than a receiver included in the mobile device, there is a possibility to easily enter the latitude and longitude into proper fields of the form. Two notations are acceptable: decimal or classic with degrees, minutes and seconds. The app will translate the units under the hood. Furthermore, to improve the orientation on the field, the app offers extra WMS overlays, which can be simply added to the base map.

Another feature offered by the application is tracking the path that the user follows for their work. Once the user turns this option on in the project configuration, the location change is stored and shown as a route on the preview map of the list of observations and on the map component of the object under information collection. The user of the AMUNATCOLL IT mobile app is able to attach multiple pictures to the observation. Another possibility for taking notes is a voice recorder. It is worth mentioning that all the media created in the mobile app could be easily exported or shared with other apps. An example is the use of an external plant recognition application based on photos imported from the AMUNATCOLL mobile app.

4.3. Data export from the mobile application to the portal and their processing

Data collected in the field using the mobile application can be analysed using the tools implemented in the AMUNATCOLL IT portal. The data export procedure allows you to transfer both individual observations and entire projects. This process is completely under the control of the user (Fig. 8).

The first step is to create a project on server (1). In response, the application receives a remote project identifier (2). The program then performs another observation in the project, sends the data from the form (3), and obtains the remote observation ID (4). Then, all media files are uploaded to the server (5). Assuming the result (6) of uploading the last media file in the observation is successful, the observation is available for further work in the "My AMUNATCOLL" section of the IT AMUNATCOLL portal. In the event of a technical error during the upload process, the work chain is saved, and the background engine tries to continue the process later.

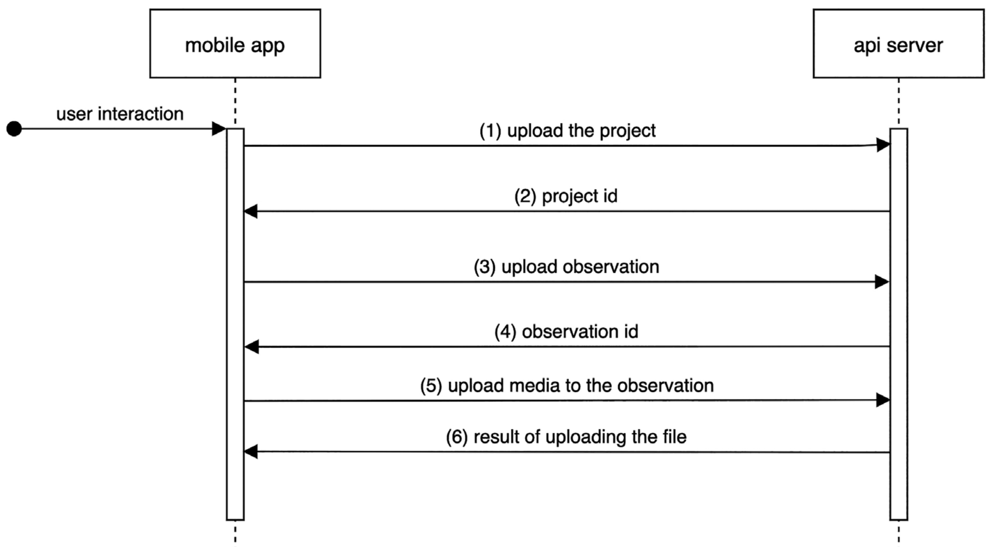

Fig. 8. Sequence diagram showing the process of uploading data from the mobile application to the portal

4.4. Mobile data in My AMUNATCOLL

"My AMUNATCOLL" represents the part of the portal functionality that allows users to create and manage their own, private space and populate them with the data harvested in the field research. It is also a specific place on the portal where the user may store albums and maps designed upon both AMUNATCOLL and private data. This part of the portal is available and operated exclusively to registered users. Field observations made on the mobile device are transferred to a separate database, which is available through the portal in the section "My Observations". This information is independent of the data that make up the AMU natural history collection database and is only reachable to a user and authorised persons. The "My Observations" view contains a list of projects and their constituent field observations added using the mobile application. The user of the portal has the possibility to compare their own observations with images originating from the AMUNATCOLL IT database and present them on a map. It enables conduct analysis where the user's own observations are shown against the distribution of specimens from the main database.

On "My Albums", users can create their own datasets. They can include specimens taken from the taxonomic database, multimedia materials coming from the iconographic database, and user observations (Fig. 9). Such a personalised assemblage permits quick access to interesting content and makes the preparation of educational presentations possible.

5. Graphical and WCAG aspects

5.1. Graphical project

The graphical design of the AMUNATCOLL IT portal meets a number of requirements regarding usability, appearance and accessibility. The work on the graphical interface began with adopting the assumptions of the portal goals, content architecture, defining the audience, and then creating a functional model and graphic design. A very important task was to understand who the user is, what their expectations are and how the website is used.

Users originate from different groups: academics, teachers and students, representatives of the state and local administration involved in nature protection, state services and officials responsible for the protection of species, representatives of nongovernmental organisations and simply nature fans. Their needs and experiences in using IT solutions differ; therefore, it was necessary to design an appropriate UX (user experience) and UI (user interface) so that each user could find exactly the content they expected.

According to the UX rules, it was necessary to anticipate the steps taken by users to gain access to functionalities or information, also known as the user flow. The content of the website is available for people who are not logged in, as well as for users who need additional information and functionality available after logging in. Both of these groups have a defined path to access information. Most of the pages are available

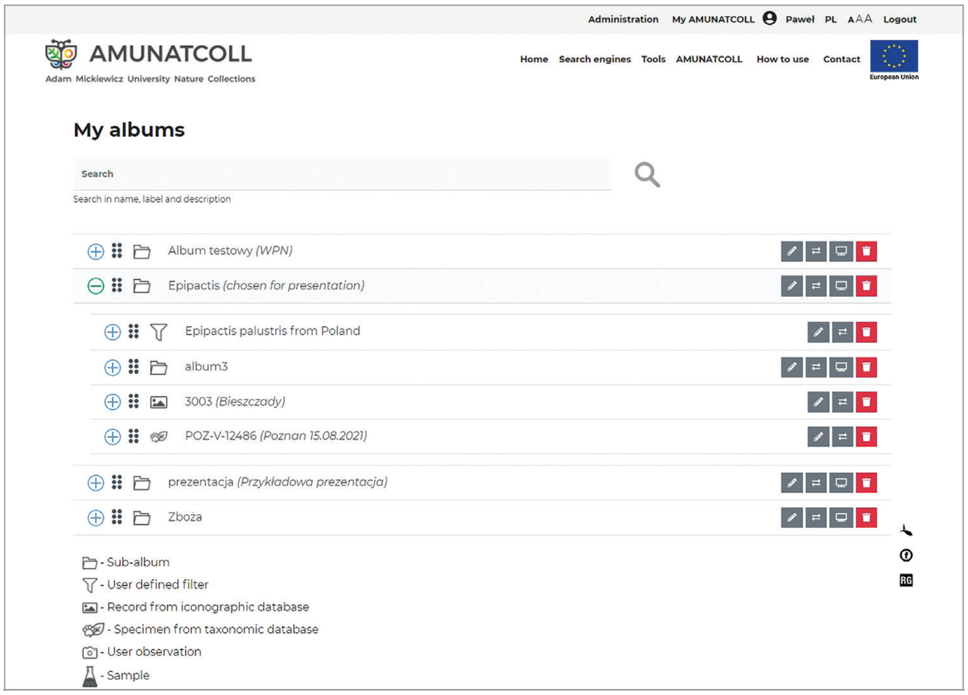

Fig. 9. View of 'My albums' presenting the user's own datasets

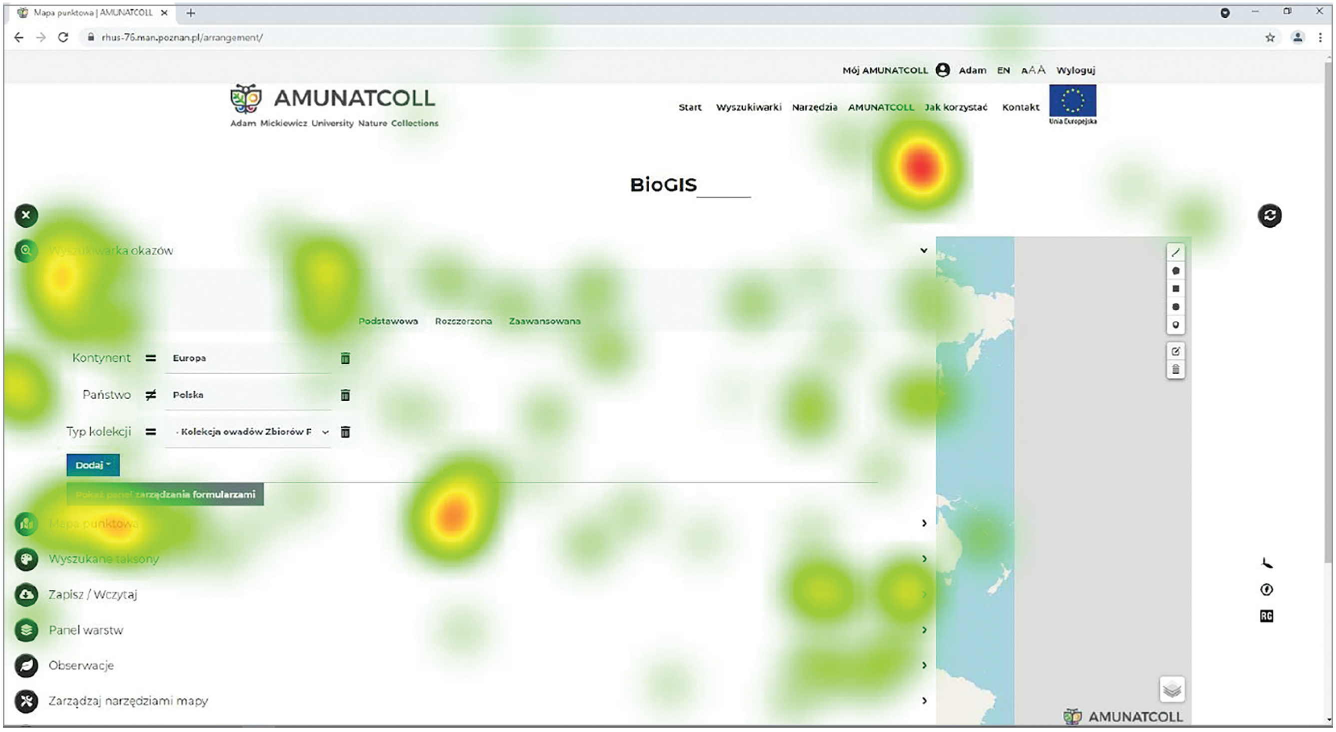

Fig. 10. Heatmap for the specimen search on the map using the BioGIS form for defining parameters

from the top menu, whose structure was intuitively implemented.

Another principle made in the AMUNATCOLL IT portal concerns appearance. It must be considered that the portal serves both its content provider and recipient. This goal is achieved by using an appropriate colour palette. A well-chosen colour palette enhances a positive impression of using the website. Complementary colours create balance and overall harmony. Using contrasting colours for the text and the background makes it easier to read. The website uses the whitespace principle – "*white space*". Whitespace is a space that separates the elements of a website. Empty spaces have a very positive impact on the reception of content on the AMUNATCOLL IT portal. First, they give the impression of spaciousness on the site, and thus, eliminate the feeling of tightness between the individual elements of its content. Because we use additional spaces, they are much more readable. Whitespace highlights the most important parts of the page and allows the reader to focus their attention on them. One such element consists of the call-to-action buttons on the home page in each module. Forms and advanced search engines are located in empty fields. This increases the chances that they are noticed and filled.

The graphic design of the subpages is prepared in accordance with the expectations of the Google search engine and the expectations of the user. The developed subpage template includes the header with the page title, the lead text introducing the page content and the actual content, where the text blocks are divided with graphics. In addition, when designing the pages, we consider SEO requirements (SEO 2021).

The AMUNATCOLL IT portal is designed in accordance with the responsive web design rule, so it has been designed and coded in such a way that it works and looks satisfactory regardless of the screen resolution from which it is viewed. The portal is prepared for proper operation on a large monitor, tablet or smartphone. Responsive web design makes the user find and easily read all the information they are looking for without the need to reduce or enlarge individual elements of the portal.

5.2. Portal usability tests

To verify that all the defined design rules are correctly implemented in accordance with their goals, usability tests were carried out. The purpose of the usability test was to check the intuitiveness of using the user interface of the amunatcoll.pl portal. The course of the study was divided into the following stages: providing information about the course of the study, completing the first part of the form, performing tasks at the computer, and completing the second part of the form. The picture presented below depicts an exemplary visualisation of the heatmap for the selected interface operation. It shows the areas that the user observed (focused their eyes on) the longest period of time (Fig. 10).

It can be concluded from the test outcome that respondents:

- Quickly found the search form and the required functions (e.g., “search”).
- Expected the button to close the side menu on the right side of the modal window with the search engine due to the dynamics of the form, i.e., drop-down selection lists.
- Used the specimen and multimedia search engines, which shows how easy the portal mechanisms are to learn (before that, users got to know specimen search engines), very well.
- Spent a relatively long time getting to know all the options of the search form for statistical reports. Its operation was not difficult for the respondents. This finding drives us to the suggestion that guides explaining available functionality are essential.

5.3. WCAG

Meeting the Web Content Accessibility Guidelines (WCAG 2021) requirements in the project concerns the access interfaces on which the user operates. They meet the requirements of the WCAG, which are visibility, functionality, comprehensibility and reliability. The content on the portal is prepared and available to people who have various limitations but still want to use the knowledge and the tools offered by the portal.

The functionality increasing the availability of data is implemented in both applications constituting the interface with the user: the portal and the mobile application. However, depending on the access interface, various solutions are applied to increase the accessibility of the content and the services offered to people with disabilities. One of the examples is to deliver information presented on the photo, although the user cannot see or use the mouse but can only use the keyboard, enlarge the view of the pages, change their colours to be able to see better content, and change the browser settings to make the content more readable.

On the portal, the user can find the declaration of the availability of the AMUNATCOLL IT portal, from which users can determine to what extent the website complies with the requirements of the act on digital accessibility of websites and mobile applications of public entities. With respect to the portal, special attention is given to the following:

- Visual adaptation – the ability to increase the size of letters and change the contrast of the page.
- Content structure – a clear structure that clearly reflects the relationship between its most important elements of the website, which makes it easier to understand the content presented.
- Keyboard accessibility – the offered functionalities are available using the keyboard, which allows people who have difficulties operating them in using, for example, a mouse.
- The dynamics of the website content – limiting the dynamic elements on the website is favourable for people, e.g., with impaired cognitive functions.
- Hints in forms – easier understanding of the scope of the content entered by the user.
- Descriptions of graphic elements – textual description of the graphics makes them easier to read by a computer program, e.g., for visually impaired people.

In the mobile application, by operating with the system settings, the user enables: adjusting the size of the user interface (scaling), adjusting the font size in the user interface, colour correction and its reversal, screen-reader interface compliance and zooming in on the selected area of the application.

6. API interface

API defines a set of rules for how data are exchanged between applications. It enables separation of the data logic layer from the presentation layer. Access to information stored in the AMUNATCOLL IT database takes place via the access interface implemented in the HTTP technology. The same interface is used by the website and the mobile application. Most of the offered functionality is available only to the logged in user; therefore, access to individual interface methods is secured with the use of JWT tokens.

6.1. Backend API

The portal implementation consists of two parts (Fig. 11). The server part (backend) was prepared in the Python language (PYT 2021) with the use of the Django programming platform (DJA 2021) with the Pandas (PAN 2021) and Gunicorn libraries (GUN 2021). The browser part of the portal (frontend) was prepared in the SPA (Single Page Application) model in the JavaScript language using the ReactJS library (REA 2021). The BioGIS tools were prepared using the GeoJSON library (GEO 2021) on the database side and the Leaflet (LEA 2021) and Pixi.js libraries (PIX 2021) on the frontend side. Communication between the frontend and backend is performed using HTTP API calls. Query authentication uses JWT token technology (JWT 2021).

The API provides methods for accessing data. There are four groups of API endpoints in AMUNATCOLL IT:

- Access to data describing specimens, samples and bibliography API makes it possible to retrieve information about a single specimen, a sample or a bibliography and to search records on the basis of the given parameters.

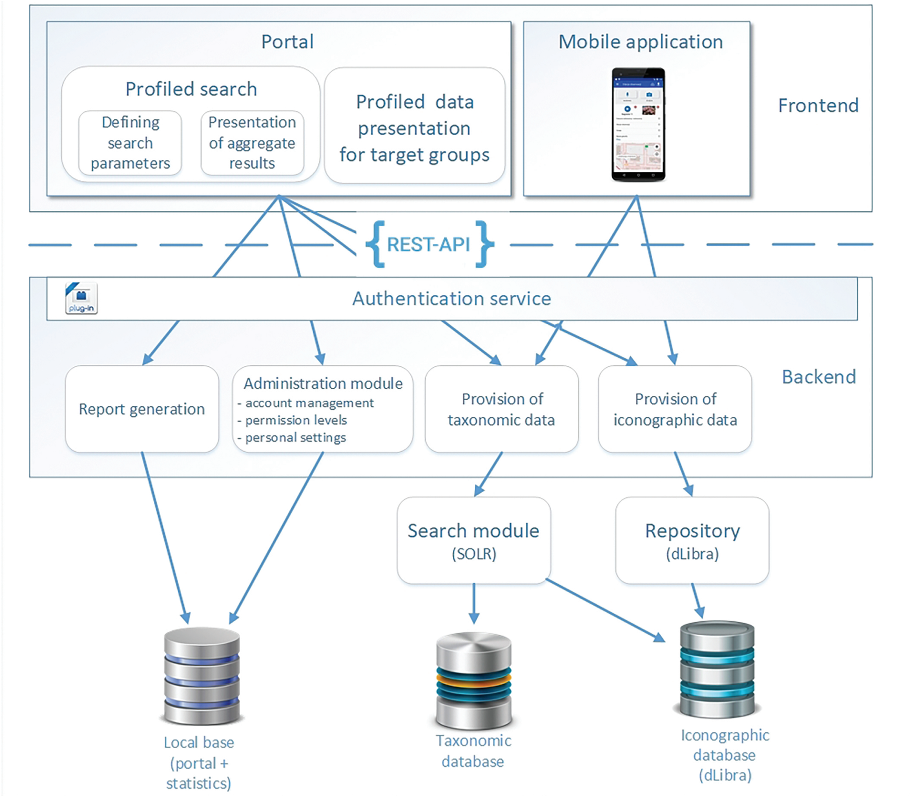

Fig. 11. Architecture of the presentation system in the AMUNATCOLL IT project

- Access to multimedia data API allows us to retrieve information about a single record describing iconographic or multimedia materials and to search records based on given parameters.
- Access to data created by the user on the portal API allows us to create and retrieve information about observations, albums, maps, search filters and reports.
- The user authentication API allows the user to log in, log out and refresh tokens. It is possible to log in using the user and password, as well as using external authentication services, such as Facebook, Google and Apple.

The API supports the GET and POST requests and returns results in the JSON format.

6.2. Interaction among databases

Another API interface implemented in the AMUNATCOLL IT system is the one that provides open access to selected taxonomic data stored in the database that are made commonly available to external entities. For this purpose, "BioCASe Provider Software" (BPS) was used (Fig. 12), a middleware compatible with the "Biological Collection Access Service" (BioCASe 2021). The BioCASe Biological Collections Access

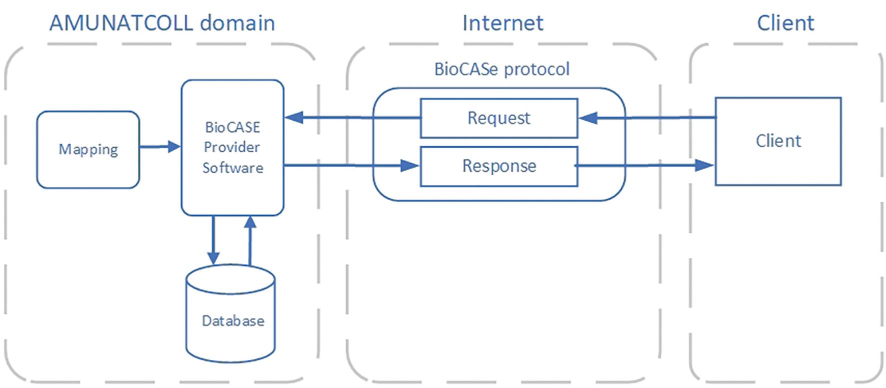

Fig. 12. Communication of the AMUNATCOLL IT system with external databases

Service is a transnational network of biodiversity repositories. It combines specimen data from wildlife collections, botanical/zoological gardens and research institutions around the world with information from large observation databases. Mapping that is developed for BioCASe services assures that structures used by both sites (BioCASe and AMUNATCOLL IT) are compliant.

The Global Biodiversity Information Facility is an exemplary biodiversity database where the records will be exported to. It is an international network and research infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth. Currently, approximately 2 billion specimens are deposited there by contributors from all over the world.

7. Conclusions

In this paper, the overall idea of the AMUNATCOLL IT system interfaces is presented. Compared with other systems available on the internet of a similar domain, the AMUNATCOLL IT solution is found to be competitive with respect to the scope of stored data and functionality offered for exploring and processing biodiversity information.

The portal, which is the main interface for accessing the data stored in the database, offers a number of additional capabilities that go beyond the simple browsing of records. Simplified and advanced searching, statistical analysis, BioGIS processing, records modification directly from the portal, extensive administrative functions and finally the ability to export data to process them independently with external tools involved are the main parts of the possibilities offered by this interface.

A mobile application is a valuable supplement to the portal's functionality. It supports effective field work, which is, after all, an inherent part of a biologist's work. The easy creation of projects according to a specific pattern, grouping of observations, various support for georeferencing, the possibility of attaching photos and voice notes and numerous functions supporting filling up the fields define the scope of the possibilities of this solution.

Importantly, the implementation of interfaces with which users come into contact is subject to numerous requirements and restrictions, which are reflected both in the graphic design itself and in accessibility issues related to the accommodation of disabled individuals. A supplement to the considerations on interfaces is the discussion of solutions supporting software communication in the field of data exchange between existing databases. There are two solutions of this type in the project: the backend API providing functionalities for interfaces, such as portal and mobile application, and an openly available API allowing for free data exchange with external systems.

Implemented to the AMUNATCOLL IT system operational and programming procedures we prepare this solution to accept data from other NHCs. Moreover, the AMUNATCOLL IT system, through the resources offered, substantially contributes to the global biodiversity information network, enriching information about the world's biodiversity.

Acknowledgements. This work has been supported by the AMUNATCOLL project and has been partly funded by the European Union and Ministry of Digital Affairs from the European Regional Development Fund as part of the Digital Poland Operational Program under grant agreement number: POPC.02.03.01-00-0043/18. This paper expresses the opinions of the authors and not necessarily those of the European Commission and Ministry of Digital Affairs. The European Commission and Ministry of Digital Affairs are not liable for any use that may be made of the information contained in this paper.

Author Contributions. All authors contributed to the study design, material preparation, data collection and analysis. The concept of the publication and the first draft of the manuscript were presented by Maciej Nowak and Marcin Lawenda. All authors participated in the development of subsequent versions. Maciej Nowak, Marcin Lawenda and Bogdan Jackowiak read and approved the final manuscript.

References

API 2021. Application Programming Interface. [Online] Available at: https://en.wikipedia.org/wiki/API

BioCASe 2021. Biological Collection Access Service. [Online] Available at: https://www.biocase.org/

DJA 2021. Django framework. [Online] Available at: https://www.djangoproject.com/

FAIR 2021. FAIR Principles. [Online] Available at: https://www.go-fair.org/fair-principles/

GBIF 2021. The Global Biodiversity Information Facility. [Online] Available at: https://www.gbif.org/what-is-gbif

GEO 2021. GeoJSON geographic data structures encoding format. [Online] Available at: https://geojson.org/

GUN 2021. Gunicorn – Python Web Server. [Online] Available at: https://gunicorn.org/

iDigBio 2021. Integrated Digitized Biocollections. [Online] Available at: https://www.idigbio.org/

Jackowiak B., Błoszyk J., Dylewska M., Nowak M. M., Szkudlarz P., Lawenda M. & Meyer N. 2022. Digitization and online access to data on natural history collections of Adam Mickiewicz University in Poznan: assumptions and implementation of the AMUNATCOLL project. Biodiv. Res. Conserv. 65: 23-34.

JWT 2021. JWT – JSON Web Tokens. [Online] Available at: https://jwt.io/

KEW 2021. The herbarium at the Royal Botanic Gardens Kew. [Online] Available at: http://apps.kew.org/herbcat/navigator.do

KML 2021. Keyhole Markup Language. [Online] Available at: https://en.wikipedia.org/wiki/Keyhole_Markup_Language

LEA 2021. Leaflet – library for mobile-friendly interactive maps. [Online] Available at: https://leafletjs.com/

MNP 2021. Muséum national d'Histoire naturelle in Paris – collections. [Online] Available at: https://science.mnhn.fr/institution/mnhn/item/search

NDP 2021. Natural History Museum Data Portal. [Online] Available at: https://data.nhm.ac.uk/about

Nowak M. M., Słupecka K. & Jackowiak B. 2021. Geotagging of natural history collections for reuse in environmental research. Ecol Indic 131: 108131.

PAN 2021. Pandas – data analysis tool. [Online] Available at: https://pandas.pydata.org/

PIX 2021. PixiJS – advanced text rendering graphic. [Online] Available at: https://pixijs.com/

PVM 2021. Plantes vasculaires at Muséum national d'Histoire naturelle. [Online] Available at: https://www.mnhn.fr/fr/collections/ensembles-collections/botanique/plantes-vasculaires

PYT 2021. Python language. [Online] Available at: https://www.python.org/

REA 2021. React – JavaScript library. [Online] Available at: https://reactjs.org/

SEO 2021. Search engine optimization (SEO). [Online] Available at: https://en.wikipedia.org/wiki/Search_engine_optimization

SMI 2021. Smithsonian National Museum of Natural History. [Online] Available at: https://collections.nmnh.si.edu/search/

TRO 2021. Tropicos database. [Online] Available at: http://www.tropicos.org/

WCAG 2021. Web Content Accessibility Guidelines. [Online] Available at: https://en.wikipedia.org/wiki/Web_Content_Accessibility_Guidelines

WMS 2021. Web Map Service. [Online] Available at: https://en.wikipedia.org/wiki/Web_Map_Service