Biodiversity Informatics, 18, 2024, pp. 1-12

THE IBDATA WEB SYSTEM FOR BIOLOGICAL COLLECTIONS: DESIGN FOCUSED ON USABILITY

Miguel Murguía-Romero[1*], Bernardo Serrano-Estrada[2], Gerardo A. Salazar[1], Gerardo E. Sánchez-González[3], Ubaldo Melo-Samper-Palacios[1], David S. Gernandt[1], Susana Magallón[1], Víctor Sánchez-Cordero[1]

[1]Instituto de Biología, Universidad Nacional Autónoma de México, Circuito Zona Deportiva s/n, Ciudad Universitaria, Coyoacán, 04510 Mexico City, Mexico.
[2]SERES Sistemas Especializados, Membrillos Mz. 120 Lt. 38. Col. Ojo de Agua, Tecámac, 55770 Estado de México, Mexico.
[3]Facultad de Ciencias, Universidad Nacional Autónoma de México.
**Corresponding author: Miguel Murguía-Romero, Email: miguel.murguia@ib.unam.mx*

Abstract. The software design process must put users at the core of the process to enable them to meet their specific objectives effectively, efficiently, and successfully. Thus, a software design for a computing system to consult biological collections guided by the concept of usability will result in an effective and efficient biodiversity informatics tool. Here, we introduce IBdata, a web system to consult biological collections, developed using a design approach based on the architecture of three layers: database, business rules, and user interface. The user interface design was guided by the concept of usability focused on four core concepts: simplicity, adaptability, guide the user through the journey, and feedback. The IBdata web system that we developed is composed of three modules (query, capture and editing, and administration), permitting it to query a database with about 1.7 million specimen records. Biodiversity data query systems must be effective and efficient and should meet the user's expectations. Software design methodologies play a central role in achieving these goals, and, in this context, interface design techniques that put the user at the core of development are valuable, as in the development of the IBdata web system.

Key words: biodiversity informatics, biological databases, GeoTax search, software design, user experience, UX design, UI design.

Biological collections document the biodiversity of our planet, and in many cases, they are the result of the efforts of many people over decades or centuries (Penn et al., 2018). In the face of the biodiversity crisis (Wilson, 1985; Sandor et al., 2022), biological collections may represent the only places where recently extinct species are found. Biological collections are the primary source of information for many types of research such as taxonomy, wildlife, floristics, and conservation biology studies, biodiversity analyses, phylogenetic and even phylogeographic analyses, as well as a variety of studies to investigate the evolutionary processes associated with the origin and maintenance of biological diversity (Castillo-Figueroa, 2018). For this reason, computer systems for consulting information in biological collections are central tools for conducting research and generating knowledge. These tools should be built by putting their users at the center of their design. In recent years, user interface design has benefited from the concept of usability (Bevan et al., 2016; ISO, 2018), understood as the degree to which specific users can use a system, product, or service to achieve specific objectives with effectiveness, efficiency, and satisfaction. Thus, the process of building web systems for biological collections may be guided by the concept of usability if the goal is to provide efficient systems for the user.

Many biological collections are framed in educational contexts, being central in the training of undergraduate and graduate students. In addition, they play an important role in raising awareness in society about conservation and biodiversity issues (Wen et al., 2015). For these reasons, it is important to make the associated data of the specimens available through the internet for universal access. In addition to the availability of biodiversity data online, it is important that the computer systems through which this is achieved ensure aspects of correctness and speed.

With computer and communication technologies such as databases, the World Wide Web and the Internet, biological collections are accessible to more users if they can be consulted virtually. A clear example is the Global Biodiversity Information Facility GBIF[1] (Gaiji et al., 2013), an intergovernmental effort to establish a global infrastructure for access to primary biodiversity data.

Initiatives aimed at digitizing biological collections around the world are very diverse. Although some of the largest collections are already available in public databases, for many collections especially from developing countries, whose biological diversity may be very high, these efforts are just beginning. For example, the William and Lynda Steere Herbarium at the New York Botanical Garden holds 7.8 million specimens, of which 4 million have been digitized[2], while in the herbarium of Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional (ENCB: Mexico's National School of Biological Sciences of the National Polytechnic Institute of Mexico), with about one million specimens (Thiers, 2018), only three families of flowering plants have been digitized[3].

The biological collections of the Instituto de Biología, Universidad Nacional Autónoma de México (IBUNAM: Institute of Biology, National Autonomous University of Mexico), which includes zoological, mycological, botanical, and ethnobiological specimens, represent one of the most important resources for the study of Mexico's biodiversity. The vast majority are national collections, received by the UNAM at the time of its foundation as part of the national heritage of Mexico. Two examples are the National Herbarium (MEXU), which, with over 1.5 million specimens, houses the most representative collection of plants collected in Mexico, and among the zoological collections, the National Mammal Collection, the most important of its kind in the country.

During the last 60 years, efforts have been made to digitize the biological collections of the IBUNAM (e.g., Scheinvar et al., 1967, 1968; Gómez-Pompa et al., 1975). In the 2000s, various databases were created, mainly within the framework of projects supported by the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad de México[4] (CONABIO: National Commission for the Knowledge and Use of Biodiversity). These efforts resulted in the contribution of about 200,000 records to the Sistema Nacional de Información sobre Biodiversidad database (SNIB: National Biodiversity Information System[5]). However, these efforts represented isolated initiatives and many specimens remained to be digitized. It was not until 2012 that the IBUNAM successfully undertook a comprehensive project aimed at digitizing most of its collections with the financial support from CONABIO and the Coordinación de la Investigación Científica, UNAM (CIC: Coordination for Scientific Research, UNAM; Gernandt et al., 2014; Sánchez-Cordero et al., 2021). To date, the total number of digitized records is approximately 1.7 million, which represents about 85% of the herbarium and most of the zoological collections, except for insects. The efforts of working groups in the construction of these systems have been great, since they have covered needs correctly (both in functionality and validity of the information) and with reasonably fast responses to online queries.

CONABIO has developed various computer platforms for promoting the development of knowledge on the biodiversity of Mexico, and the experience accumulated is valuable. Among the most important examples of their systems are the SNIB, the Red Mundial de Información sobre Biodiversidad (REMIB: World Network of Information on Biodiversity[6]); and the Biótica system (CONABIO, 2012) for the capture and editing of biodiversity data, for example, from the specimen labels of biological collections (Jiménez et al., 2016). The REMIB, along with North America Biodiversity Information Network (NABIN) from the Biodiversity Research Center at the University of Kansas, were important as examples for the subsequent development of GBIF (Soberón, 1999).

To our knowledge, there is no web system for accessing biodiversity records that explicitly incorporates usability concepts into its design process. Unfortunately, there is no documentation on the design process of any biodiversity web system. The only reference found is that of a system to build taxonomic identification keys on the web (Murguía-Romero et al., 2021), which reports the use of usability in its design process incorporating responsive web design technology. Therefore, the documentation of the de-

[1] https://www.gbif.org/
[2] https://sweetgum.nybg.org/science/
[3] https://www.encb.ipn.mx/herbario/
[4] https://www.biodiversidad.gob.mx/conabio/
[5] https://www.snib.mx/
[6] http://www.conabio.gob.mx/remib_ingles/doctos/remib_ing.html

sign process of a biodiversity system that incorporates usability concepts may benefit the community involved in the development of this type of tools.

Here we present the development process of a web system to consult biological collections guided by the concept of usability focused on four core concepts: simplicity, adaptability, guide the user through the journey, and feedback. The development of the web system for consulting the biological collections data presented here owes much to previous systems, among them, and very importantly, those developed by CONABIO (REMIB, SNIB, and Biótica). UNAM's Portal de Datos Abiertos[7] (Open Data Portal, UNAM) was also a key platform for the development of this system, mainly in the design of how to summarize the list of records that results from a search.

Software Engineering

The objective of software engineering is to propose and systematize techniques and tools for the development of computer systems that seek to optimally use human, economic, and time resources to achieve the construction of computer programs that meet the objectives for which they are created. A group of very important tools in this discipline are the "development methodologies," which emerged in the 1970s as a response to the "software crisis," a term that was used to designate the problem of the high proportion of failures in the generation of computer programs that did not meet their objectives, or that were never completed.

Building software is a complex process, and therefore, if it does not follow a formal development methodology, it is likely to fail. Although there are many differences in the estimates of the proportion of projects that are successful, approximately 50% have difficulty reaching their objectives and only 16% to 18% of software projects are successful, that is, they come to fruition (Glass, 2005).

Usability

In the last decade, ergonomics has gained importance in the development of computer systems. Specifically, it has been shown that evaluating the efficiency of user interfaces contributes to the success of the system, in that the system meets the tasks and objectives for which it was created.

Usability is the degree to which specific users can use a system, product, or service to achieve specific objectives with effectiveness, efficiency, and satisfaction in a specific context of use (Bevan et al., 2016; ISO, 2018). It is a relative measure because it provides an evaluation by comparing different designs; it is also subjective, since it depends on the user and on the information that she or he is interested in manipulating (Allanwood & Beare, 2014).

Various dimensions of usability have been defined; among the most important are:

- Learning ability: how easy it is for users to perform basic tasks the first time they interact with the system.
- Efficiency: how fast users can perform tasks.
- Memorability: when users return to the system after a period of inactivity, how easily they can remember how to use it.
- Errors: how many errors users make, how serious they are, and how easy it is to recover from them.
- Satisfaction: how satisfying the users find the system.
- Attractiveness: how much the interface invites the user to interact and how pleasant it is to use.

Usability tests determine if interactions through the interface meet the needs and objectives of the user and are an important element of the design methodology and interfaces (Allanwood & Beare, 2014; Jackson & Ciolek, 2017).

System development process

An "agile development methodology" was followed to build the IBdata web system (Beck et al., 2001; Abrahamsson et al., 2002). The development team was formed in March 2018, when work began formally. The documentation of the process was intended to be as agile and summarized as possible, basically consisting of minutes from meetings, drafts or layout of the interface design, executive presentations to the working group, user guides, analysis of the information to explore the best ways to process and present it in the interface, and suggestions for modifications to the system. Users of various profiles of the system were represented on this team.

The functionalities that the system should have were decided based on the experience of information needs of the participants in the project and by making comparisons with other systems that had already been built. The main functionalities included in the system design specification were classified into three types: query, update, and administration:

[7] https://datosabiertos.unam.mx/biodiversidad/

Query:

- Obtaining lists of records of specimens filtered by the most important fields of the database, such as taxonomic (family, genus, species), the geographic location of the collection (country, state, municipality, coordinates), the date of collection (year, month, day), or by administrative data (collection, catalog number).
- Viewing a summary at the level of the specimen record with the most relevant fields.
- Visualization of the image associated with the specimen when it exists.
- Summaries of the content of the database in three formats: graphs, lists (of taxa or geographic entities), total number of records by different types of grouping (e.g., by taxa or by geographic entities).

Update:

- Editing of existing records.
- Capture of new records.
- Deletion of records.

Administration:

- User registration in the system and password recovery.
- Updating by the user of their contact information and password.
- Specification of differential permissions per user.
- Reports on system usage statistics.

The limits of the desired functionalities were defined by specifying what is not expected of the system; for example, common or frequent functionalities that are desired from a system that stores information on biological collections but that can be covered by existing systems; or, one that is highly complex and its implementation would increase the required resources, putting the system at risk to come to fruition or simply because it departs from the central objectives for which the system is intended to be used.

Some of the limits that were established were as follows. (1) The system will not provide analysis tools; there are multiple packages that can be used for this with data exported from the system. (2) The system will not attempt to solve particular requirements of the academic staff (i.e. requirements that only solve the needs of one researcher and are not useful to others), as it is intended to be the institutional computer repository of the biological collections of the IBUNAM. (3) The system will be released incrementally, so that it was expected that the first versions would only provide some basic functionalities, and more functionalities would be added in later versions.

Integrating usability into the development process

Usability is directly related to the user experience, and therefore the user interface is one of the main components of the system where developers can participate, seeking the objectives of usability: that users can use the system to achieve their specific objectives with effectiveness, efficiency, and satisfaction. The way we achieved this was to start with the selection of four key usability concepts (simplicity, feedback, adaptability, and journey guidance). The second step was to define the way in which these four concepts could be represented in the user interface, taking into account the technological elements available in the development of interfaces (such as windows, types of controls, ways to navigate, colors, among others). The third step was to integrate the different components into a model (information architecture), arranging them with a congruent signature both visually and functionally. The fourth step was to implement that mockup into a functional system. The fifth and final step was to develop usability tests that guided adjustments to the user interface.

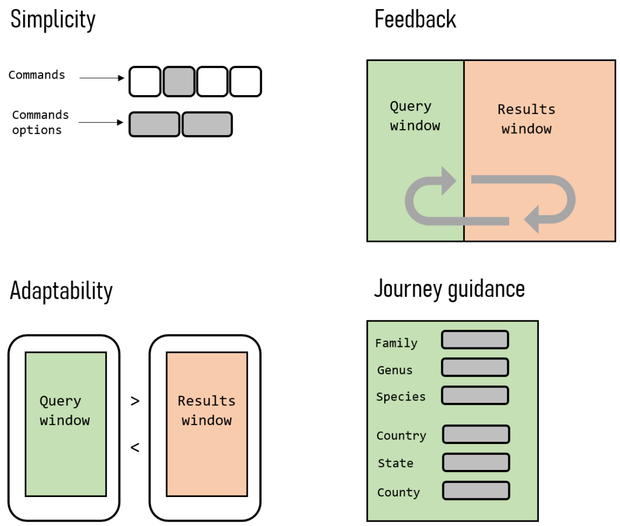
The process of building the user interface is explained below, which together with the database and business rules make up a three-layer system.

User interface design (UI design)

One of the central aspects for the success of a system is that it is used by the people for whom it is designed; that is, that the users use the system. There are other central aspects that cannot be ignored for a successful computer system, such as meeting the needs for which the system is built, and doing it correctly and reasonably quickly, i.e., correctness and speed. When the development team works with appropriate methodologies and those two aspects are taken care of, the question of the user interface is central to achieve the success of the development. Usability is precisely the concept that guides the application of design techniques to achieve a good user experience.

The following four axes were defined as usability criteria to achieve a good user experience in the system (Figure 1):

Figure 1. The four axes considered in the design of the system to achieve a high level of usability. Simplicity: the available commands to consult the database are visible. Feedback: visual and operative dialectic between the queries and their results. Adaptability: responsive interface changing its aspect according to the dimensions of the device. Journey guidance: the system informs the user where and how to perform a query of the database.



- Axis 1: Simple exploration. Commands available to the user with minimal hierarchy.
- Axis 2: Guide the user in their journey. Searches focused on geography and taxonomy as the main search criteria.
- Axis 3: Feedback in the queries (query feedback). Dialectic between the queries and their results.
- Axis 4: Adaptability to the user. Responsive and multilingual.

Axis 1 of usability: simple exploration (commands available to the user with minimal hierarchy). Every interface designer wants their product to be simple to use by all the targeted users. But the criteria for simplicity of use are relative, both to the user and to the state of the technology available at the time it is developed. Simplicity can be seen as the intelligent management of complexity (Allanwood & Beare, 2014). This conception is important especially if it is considered within the development team, because if the project management has that in mind, it will be able to recognize that the solutions that provide simplicity to the system are the result of a non-random and conscious analysis that shows the user in a simple way the intrinsic complexity of the system and the information it contains.

The specification of the system design was defined so that the simplicity focused on ease of access to available commands. This accessibility should be visual and as far as possible, immediate; that is, a panel should be provided that presents a complete, explicit, and organized picture as possible of the operations that could be requested from the system in the current state of the session. Thus, two of the criteria for the design of the interface were to a) avoid implicit commands, that is, ensure that they had a visual reference in the interface, and b) arrange the commands with minimal hierarchy.

The proposed solution was to present the application-specific commands to the user on an always-visible bar containing them. The commands for general use, which can be found in other applications, such as login, logout, registration, and "I forgot my password," among others, were placed in the usual way as a menu on the top bar and as a hamburger menu. Keeping the largest number of commands in a place of common use is something that provides simplicity in the interface, since the users find what they need where they expect it. But placing the specific commands of the application in such spaces would require a user-oriented search effort. For this reason, it was decided that they should be as explicit and visually obvious in the interface (Figure 1 "Simplicity" section). Therefore, its expression in the interface would be buttons with associated icons and tooltips. If the number of buttons required was excessive and difficult to accommodate in the interface, they could be ranked in no more than one level.

Axis 2 of usability: guide users in their journey; searches focused on geography and taxonomy. The interface should help the user to decide how to conduct a search, suggesting the types of filters that can be the simplest and most effective. There are different types of information that are customarily integrated in biodiversity databases, for example, administrative information (such as the name of the collection or the catalog number of the specimens), temporal information (e.g., the collection date), or description of the biology of the specimen (such as the phenological state or its size as measured by the collector), among other types. In this context, the interface is designed to interact with the user bearing in mind that *errare humanum est*, and that it is by means of a user-interface dialectic that the system can become, through successful and unsuccessful cycles of use, an efficient and effective tool.

Axis 3 of usability: feedback in the queries (dialectic between the queries and their results). By feedback in the queries, we mean the return of the system's output information to the query that generated it. The user-system feedback focused on the user's queries and the corresponding results reported by the system. There are other types of feedback in a user interface, such as the current state of the system (e.g., waiting, working, on, selected, among many others), the indication of consistency (or inconsistency) of the combination or sequence of commands selected by the user, the drill-down, warnings or additional notes that support or guide the user in performing tasks, among others.

Of all the types of feedback that are usually given in a user interface, it was decided to focus on the one that occurs between the user's requests (queries) and the corresponding results that the system provides. It was decided to implement the other types of feedback following the standards suggested in the design area, without analyzing or specifying in greater depth how to improve it. This feedback idea takes its heuristics from another system whose theoretical foundation is a "dynamic identification model" (Murguía-Romero, 1992; Murguía-Romero et al., 2021). The results can be queried; that is, it is possible to filter them so that the user can refine the query in a recursive, or spiral process.

In scientific databases, which are usually extensive and rich (that is, with many records and with many tables, fields, and relationships), the user's queries must be interpreted for their intention rather than taking them as a fixed structure, an immovable contract between the user and the system, from which a complete and correct answer is expected. The results of the queries should help the user to understand the content of the database and provide guidance to continue exploring the data (Kersten et al., 2011).

Axis 4 of usability: adaptability to the user (responsive and multilingual). By adaptability it is understood that the system recognizes the characteristics of the type of use that the user gives to the system, or the preferences that the user establishes on it. This adaptability can occur automatically or be directed by the user. In the design specification, two specific criteria of adaptability were established: that the system was responsive; that is, that the interface was adapted according to the dimensions of the screen of the device in which it is used, and that the user could choose the interface language with at least two options. The first specification is automatic, while the second, the language, is chosen by the user. Another specification of the design on adaptability that has not yet been implemented is the user profile; that is, the system stores the preferences of the environment indicated by the user. For example, when the user logs in, the system restores the environment according to the user's preferences, including the interface language, the collection to consult by default, the display of the most used search type, or the default taxonomic family searched, among others.

Usability testing

The design was incrementally implemented in functional prototypes. In each, usability tests were carried out, which guided the adaptation of the interface, its functionality, the graphic elements, and the associated texts, among many other elements. Some examples of the modifications are the adjustment of the location, shape, size, and color of each of the interface elements, as well as its structure and arrangement, among many others, seeking in each cycle to improve the user experience.

Figure 1 shows the general layout that guided the design and from which the interface was implemented in programming languages. The usability tests consisted of inviting 15 faculty and students from IBUNAM to use the IBdata web system. After an average of ten days, users were interviewed by the designers to record their observations, which were incorporated by making pertinent changes in the interface. This process was repeated three times incrementally, that is, inviting new users to join those who participated in the previous tests.

The usability measurement consisted of a measure relative to each user, verifying that they reported an increasingly better user experience. The context of each user was also taken into account, depending on the type of information or reports that interested them. In this sense, they were classified into three types of users: curator, taxonomist, and ecologist. Prior to the usability tests, three respective fictitious characters were created that guided the design (Allanwood & Beare, 2014). Finally, these characters were reflected in the data view of the interface in the form of a table, which shows a different set of fields depending on the chosen profile.

Information architecture

The design must define and plan the way in which the interface will present the flow of information or its subsets to the user, either through requests that they make or at the suggestion of the system. This specification, known as "information architecture", must differentiate between information that is relevant to the user and that which is secondary. It is in this design space that usability-oriented design specifications were concentrated. The resulting information architecture considered the integration of each of the four chosen usability criteria axes (Figure 2),

Figure 2. The design of the information architecture for IBdata. The main structure of the interface presented to the user (information architecture) is composed of 5 areas: three bands (Header, Tools, and Footer) and two windows (Query window and Results window).

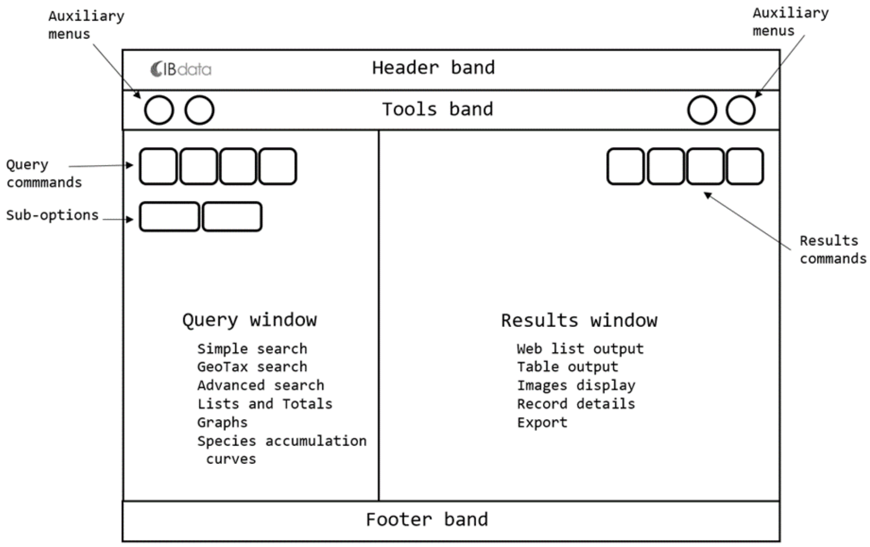
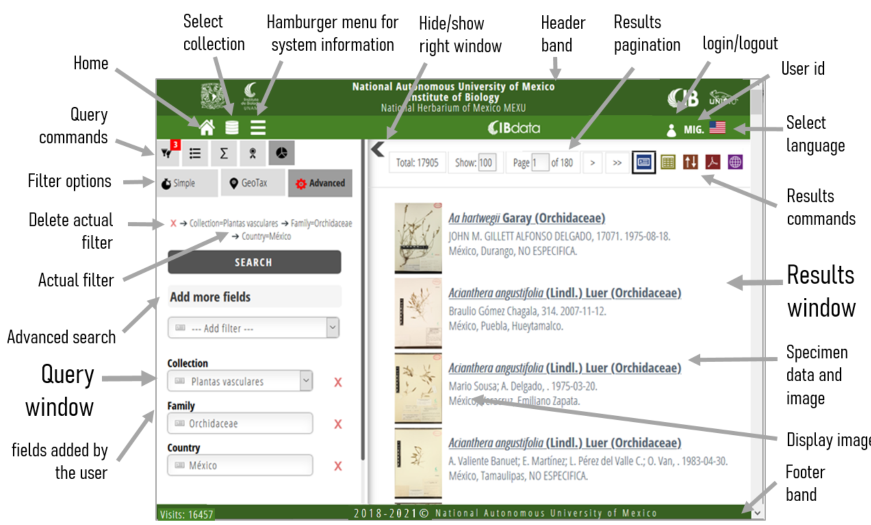


Figure 3. Components of the Ibdata bifenestra user interface. The result of the query "Specimens of the Orchidaceae of Mexico in the Vascular Plants collection" is shown through "Advanced Search" (left window) showing the results in the "Web view" (right window).



and its objective is to show, *grosso modo*, the desired appearance of the interface in its functional version. Institutional standards (UNAM, 2016) imposed demands on design, such as the location and size of the logos, copyright notice, and a visits counter. These were attended by locating the elements involved in an upper and a lower band (header and footer bands in Figure 2).

Implementation of designed components

Architecture of three layers

The web system, called IBdata[8], allows the consultation of a database with so far, 1.7 million specimen records of the biological collections of the IBUNAM (Murguía-Romero et al., 2023). It is made up of three modules: queries, capture and editing, and administration. The three layer architecture is implemented through a) a database with 70 tables, b) the program code in various programming languages, and c) the interface that the user accesses through the web.

User interface

The web interface (Figure 3) takes into account the specifications of the information architecture (Figure 2). Based on this architecture, the various previously defined elements were adapted, *mutatis*

[8] http://ibdata.ib.unam.mx/

mutandis, in a more specific way. It is made up of three bars, two above and one below; and two windows, one to the left and one to the right, which we called a "bifenestra" user interface (Figure 3). The first upper bar fulfils the institutional requirement that web pages must contain official names and logos. The second upper bar contains the menus and commands that web systems usually have, such as the user menu (with the login and logout commands, access to edit the account data, among others), an icon for language selection, and a hamburger menu with commands to view general system information and credits. In this second bar an icon of a drum was also included, which allows the user to select the collection to consult. This considers a display on a large screen, but because the system is responsive, these elements change in their arrangement and presentation on smaller screens.

Between the top and bottom bars, and occupying most of the general window, are the query window (left) and the results window (right). Each of these windows contains in its upper part a list of buttons with which commands are accessed, either to perform queries or to manipulate the results, respectively. The result of a search is presented in the right window, typically a list of specimen records in a summarized format, which has been called "web view". The format can be changed to "table view" by choosing the corresponding command in the results window. Clicking on any of the records displays the "Specimen data summary sheet." If the specimen has an associated image, it is shown as a thumbnail, which upon clicking on it, displays the high resolution image in a dialog where it can be explored by zooming in on the portions of interest to the user. Table 1 summarizes the most important characteristics of the web system that were imposed as system requirements in the design stage, always guided by the purpose of achieving a good user experience.

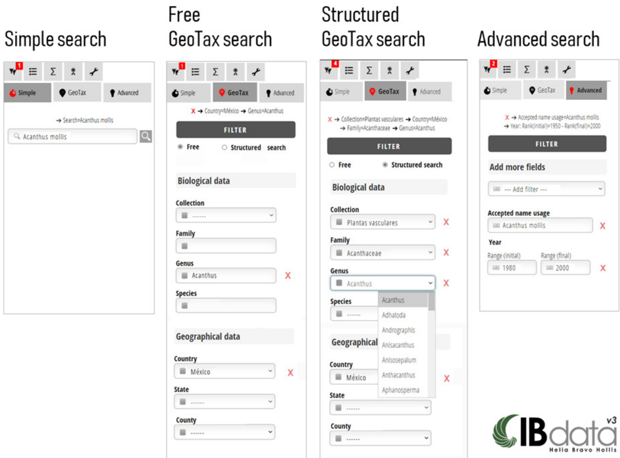
Query module

The interface reserves the left window as the space for the specification of user queries. Queries obtain lists of specimens filtered by different methods: simple search (filter by genus or species name), GeoTax search (family-genus-species and country-state-county hierarchy), and advanced search, in which the user can choose any set of fields to filter the specimens (Figure 4). It is in this space that it was decided to focus the guide of the user journey, because through text boxes the user is invited to make a successful query depending on the level of knowledge they have about his or her question or the available data.

Table 1. Characteristics of the IBdata web system.

	Feature	Description
1	Responsive system	The interface adapts to the size of the device screen
2	Compatibility with the most used operating systems	Windows, Mac OS, Linux Kernel, iOS, and Android
3	Compatibility with the most used internet browsers	Edge, Firefox, Google Chrome, Opera, and Safari
4	Level 1 navigation	Most of the functionalities are on the same page; it is not necessary to navigate outside of the current page
5	100% web	User does not need to install components
6	User login	Grants different access permissions and customizes some aspects of the interface
7	Multilingual	Spanish / English / Italian with the ability to add more languages easily
8	Multi-collection	Stores data from more than one collection
9	Darwin Core standard	Field names according to the Darwin Core standard
10	Query module	Advanced simple queries, lists, totals; export to PDF, Excel, and CSV
11	Administrative reports	Reports for the system administrator on usage and capture statistics
12	Managing users accounts	Assigns user permissions
13	Capture and edit module	Add, edit, and delete records

Figure 4. Query options in the IBdata web system.
Simple search: the user enters the name of a genus or species.
GeoTax free search: the user enters search terms in one or more text boxes out of the seven displayed (collection, family, genus, species, country, state, or county). GeoTax structured search: the user must specify hierarchically, choosing from lists.
Advanced search: the user can select the fields with which to perform the filter.



The simple search allows searching for specimens by species name. If users are taxonomists (or know the taxonomic names) and want information on the specimens of a particular species, the simple search indicates in the text box to type the name of the species (Figure 4, "Simple search" panel).

The GeoTax search is a recurring requirement of users when they want to display the records of specimens filtered by some category of political geography (country, state, or county) or by some taxonomic category (family, genus, or species). It has two variants or options, free and structured. The free search is oriented to situations in which the user does not require guidance in specifying the taxon name, or when she or he precisely know the name of the political geography entity for which they wish to apply a filter (Figure 4 "Free GeoTax search" panel).

If the user is uncertain about the names of the taxa or geographic entities to consult, the structured GeoTax search provides drop-down lists with which the user can hierarchically choose the name of the geographic entity (country, state, or county) or taxon (family, genus, and species) to be included in the query (Figure 4 "Structured GeoTax search" panel).

The advanced search allows users to choose the fields by which to filter the specimens (Figure 4 "Advanced search" panel). If the field chosen is of numeric type, for example the year of collection, the interface shows two boxes in which the lower and upper limits can be specified. If the field is of text type, for example for collection location, the interface interprets the text specified in the corresponding box as a substring of the location, that is, it selects those records that in the location field contain the text written by the user. The user can choose and add any number of fields to the query. When more than one field is included in the query, it is interpreted as a conjunction (as a logical "and"), requiring that all the conditions represented by the values indicated for each field must be met.

Edit and capture module

The IBdata edit and capture module can be used from any computer connected to the Internet. It was designed considering various features aimed at the best efficiency for users. The capture interface is not available to all users, only to those who have been granted the corresponding permissions. It contains basic validations to ensure its quality. In some cases, the values must be chosen from a list whose source is a catalog of the database, for example, the country or the state; for fields which the user can write in a free format, validations are made according to rules that ensure integrity and congruence.

Administrative module

The administrative functions are basically of two types: assigning permissions to users and their administration (delete, block, or add users) and display of usage reports and record capture. The administrator has access to the list of registered users in that interface, and by sorting by names, surnames, or some other field, can access the permissions assigned to edit them.

Lessons learned

Systems for consulting biodiversity data, such as specimen records of biological collections, are complex, and subject to failure risk. Therefore, the use of development methodologies arising from the field of software engineering are essential tools. In its planning, the architecture must be defined, among other aspects, to maintain the highest possible data independence. Likewise, the design of user experiences should benefit the greatest number of users, supporting them in efficiently obtaining the required information. The concept of usability is the central element to achieve better experiences. In this work it has been shown how the choice of four usability axes can guide the construction of a system that meets these goals.

CONCLUSIONS

In the development of systems for querying biodiversity data, such as those that allow the query and administration of biological collections, it is a *conditio sine qua non* to involve the user in all stages of development if such systems are to meet user expectations effectively and efficiently. Software design methodologies play a central role in achieving these goals, and in this context, interface design techniques that put the user at the center of development are valuable tools.

IBdata is an institutional system developed by the Instituto de Biología, UNAM, for consulting, capturing, and editing specimen information of the collections that it houses. Its creation followed a software development methodology focused on data independence, and the concept of usability was considered at all stages to meet the expectations of its users as much as possible.

ACKNOWLEDGMENTS

Many faculty members and students participated in the usability tests of the system, which contributed to the success of its development. The Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO project "KE002: Digitization and Systematization of the National Biological Collections of the Institute of Biology, UNAM") and the Coordinación de la Investigación Científica, UNAM financially supported the digitization of the collections of the Institute of Biology, UNAM. Javier Nori and three anonymous reviews made suggestions to the manuscript that improved it.

AUTHOR CONTRIBUTIONS

M.M.R. wrote the first draft with contributions from B.S.E. who developed the program code of the system; M.M.R. and B.S.E. designed the system database; all authors contributed to the design of the system and proof-read the manuscript.

Competing Interests

The authors have declared that no competing interests exist.

Literature cited

Glass, R.L. IT failure rates—70 percent or 10–15 percent? IEEE Software 22, 3 (May–June 2005).

Abrahamsson, P., Salo, O., Ronkainen, J. & Warsta, J. (2002) Agile software development methods: Review and analysis, VTT publication 478, Espoo, Finland, 107p.

Allanwood, G., and P. Beare. 2014. *Basics interactive design: User experience design: Creating designs users really love*. A&C Black.

Beck, K., M. Beedle, A. Bennekum van, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, J. Kern, B. Marick, R. Martin, S. Mellor, K. Schwaber, J. Sutherland and D. Thomas (2001). "Manifesto for Agile Software Development." 2002(22.3.2002) http://AgileManifesto.org.

Bevan, N., J. Carter, J. Earthy, T. Geis, and S. Harker. 2016. New ISO standards for usability, usability reports and usability measures. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9731*, 268–278. https://doi.org/10.1007/978-3-319-39510-4_25

Castillo-Figueroa, D. 2018. Beyond specimens: Linking biological collections, functional ecology and biodiversity conservation. Revista Peruana de Biología, 25(3), 343-348.

Coordinación de Colecciones Universitarias Digitales (CCUD). 2017. Manual de Datos Abiertos de Colecciones Universitarias Digitales. Secretaría de Desarrollo Institucional, Universidad Nacional Autónoma de México (SDI-UNAM). México. 117pp.

CONABIO. 2012. Sistema de Información Biótica. Manual de Usuario. Versión 5.0.3 México, D.F. 321pp. http://www.conabio.gob.mx/biotica/cms/Descargas/Biotica50/ManualBiotica50/Biotica50.pdf

Gaiji, S., V. Chavan, A. H. Ariño, J. Otegui, D. Hobern, R. Sood, and E. Robles. 2013. Content assessment of the primary biodiversity data published through GBIF network: Status, challenges and potentials. *Biodiversity Informatics*, 8(2).

Gernandt, D. S., V. Sánchez-Cordero, U. Melo Samper Palacios, O. J. Giménez-Héau, and G. A. Salazar. 2014. Digitalización del Herbario Nacional de México: Avances y retos del futuro. *Revista Digital Universitaria*, vol. 15, no. 4, pp. 1-13.

Grattarola, F., G. Botto, I. da Rosa, N. Gobel, E. M. González, J. González, J. Gonzállez, D. Hernández, G. Laufer, R. Maneyro, J. A. Martínez-Lanfranco, D. E. Naya, A. L. Rodales, L. Ziegler, and D. Pincheira-Donoso. 2019. Biodiversidata: An open-access biodiversity database for Uruguay. *Biodiversity Data Journal*, 7.

Glass, R. L. 2005. IT Failure rate- 70% or 10-15%? *IEEE Software* 22, 3 (May-June 2005)

Gómez-Pompa, A., J. A. Toledo, and M. Soto. 1975. Electronic data processing of herbarium specimens data for the Flora of Veracruz Program. In *Computers in Botanical Collections* (pp. 35-51). Springer, Boston, MA.

Gries, C., E. E. Gilbert, and N. M. Franz. 2014. Symbiota - A virtual platform for creating voucher-based biodiversity information communities. *Biodiversity Data Journal*, (2), e1114. doi:10.3897/BDJ.2.e1114

ISO. 2018. 9241-11: 2018 Ergonomics of human-system interaction—Part 11: Usability: definitions and concepts. *International Organization for Standardization,* https://kebs.

isolutions.iso.org/obp/ui#iso:std:iso:9241:-11:ed-2:v1:en, *9241*(11).

Jackson, C., and N. Ciolek. 2017. *Digital Design in Action: Creative Solutions for Designers*. AK Peters/CRC Press.

James, S. A., P. S. Soltis, L. Belbin, A. D. Chapman, G. Nelson, D. L. Paul, and M. Collins. 2018. Herbarium data: Global biodiversity and societal botanical needs for novel research. *Applications in Plant Sciences*, *6*(2), e1024

Jiménez, R., P. Koleff et al. 2016. La informática de la biodiversidad: Una herramienta para la toma de decisiones, en Capital Natural de México, vol. iv: Capacidades humanas e institucionales. Conabio, México, pp. 143-195.

Kersten, M. L., S. Idreos, S. Manegold, and E. Liarou. 2011. The researcher's guide to the data deluge: Querying a scientific database in just a few seconds. *Proceedings of the VLDB Endowment*, *4*(12), 1474-1477.

Morville, P., and L. Rosenfeld. 2006. *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites*. O'Reilly Media, Inc.

Murguía-Romero, M. 1992. *Métodos en la identificación biológica automatizada* (Master's Thesis). Facultad de Ciencias, Universidad Nacional Autónoma de México. México D.F. http://132.248.9.195/pmig2017/0187269/Index.html

Murguía-Romero, M., B. Serrano-Estrada, E. Ortiz, and J. L. Villaseñor. 2021. Taxonomic identification keys on the web: Tools for better knowledge of biodiversity. *Revista Mexicana de Biodiversidad*, 92: e923592.

Murguía-Romero, M., Serrano-Estrada, B., Salazar, G. A., Gernandt, D. S., Melo-Samper-Palacios, U., Sánchez-González, G. E., Sánchez-Cordero, V., and Magallón, S. 2023. IBdata versión 3 "Helia Bravo Hollis": Manual de uso. Instituto de Biología, Universidad Nacional Autónoma de México. https://www.ib.unam.mx/ib/programa-editorial/publicaciones-especiales/; doi.org/10.22201/ib.9786073073189e.2023

Nielsen, J. 2012. Usability 101: Introduction to Usability. Nielsen Norman Group. Available from: https://www.nngroup.com/articles/usability-101-introduction-to-usability/. Accessed on Sep 10, 2021.

Penn, M. G., Cafferty, S., and Carine, M. 2018. Mapping the history of botanical collectors: Spatial patterns, diversity, and uniqueness through time. *Systematics and Biodiversity*, 16, 1–13. https://doi.org/10.1080/14772000.2017.1355854

Sánchez-Cordero, V., S. Magallón, A. Contreras-Ramos, G. Salazar, D. S. Gernandt, E. González, U. Melo Samper, J. Giménez, D. Pérez, C. Reséndiz, and M. Murguía. 2021. *Digitalización y Sistematización de las Colecciones Biológicas Nacionales del Instituto de Biología, UNAM. Universidad Nacional Autónoma de México. Instituto de Biología. Informe Final SNIB-CONABIO Proyecto KE002.* Ciudad de México, México. http://www.conabio.gob.mx/institucion/proyectos/resultados/InfKE002.pdf

Sandor, M. E., Elphick, C. S., and Tingley, M. W. 2022. Extinction of biotic interactions due to habitat loss could accelerate the current biodiversity crisis. *Ecological Applications*, 32(6), e2608

Scheinvar, L., A. Gómez-Pompa, and L. Alonso. 1967. Sistema automático de recuperación de información para el Herbario Nacional del Instituto de Biología de la U.N.A.M. *Anales del Instituto de Biología de la Universidad Nacional Autónoma de México*, 38, Ser. Bot. (1): 203-250.

Scheinvar, L., L. Alonso, and A. Gómez-Pompa. 1968. Proyecto piloto de recuperación automática de información del Herbario Nacional de la UNAM. *Anales del Instituto de Biología de la Universidad Nacional Autónoma de México* 38 Ser. Bot.

Soberón, J. 1999. Linking biodiversity information sources. *Trends in Ecology and Evolution*, 14(7), 291.

The New York Botanical Garden. 2018. C. V. Starr Virtual Herbarium. http://sweetgum.nybg.org/science/vh/. Accessed on Agu/9/2018.

Thiers, B. M. 2018. The world's herbaria 2017: A summary report based on data from Index Herbariorum. William and Lynda Steere Herbarium, The New York Botanical Garden. Disponible en: http://sweetgum.nybg.org/science/docs/The_Worlds_Herbaria_2017_5_Jan_2018.pdf

UNAM. 2016. *Lineamientos para sitios web institucionales de la UNAM*. 22p. Available from: https://www.visibilidadweb.unam.mx/normateca/normaunam/%20Lineamientos_SitiosWebInstitucionales_CATIC_Octubre2016.pdf Accessed on Aug/9/2018.

Wen, J., S. M. Ickert-Bond, M. S. Appelhans, L. J. Dorr, and V. A. Funk. 2015. Collections-based systematics: Opportunities and outlook for 2050. *Journal of Systematics and Evolution*, 53(6), 477-488.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Doöring et al. 2012. Darwin Core: An evolving community-developed biodiversity data standard. *PLoS ONE* 7(1): e29715. doi:10.1371/journal.pone.0029715

Wilson, E. O. 1985. The biodiversity crisis: a challenge to science. *Issues in Science and Technology*, 2(1), 20-29.

Appendix 1: Glossary

Agile development methodology.—Computer systems development methodology that contrasts with classical or previous methodologies in which less importance is given to documentation and more to the development of functional prototypes before meeting all the requirements.

Architecture of three layers.—The way in which the elements of a computer system are classified and arranged into three large groups: user interface, business rules, and database.

Bifenestra user interface.—Two-window interface. Technical solution to the "Level 1 Navigation" feature, which seeks to avoid the user having to navigate through multiple submenus to reach the desired functionality and from having to leave the current page.

Business rules.—Specification of the restrictions and processes of the objective of a computer system.

Darwin Core.—Standard that specifies the nomenclature and meaning of the fields in biodiversity databases proposed by the Taxonomic Databases Working Group, now Biodiversity Information Standards (TDWG) and that is widely used in public biodiversity databases.

Development methodology.—Sequence of steps to build a computer system.

Drill-down.—Analytical capacity that allows users to change immediately from a general view of the data to a more detailed one by clicking on a metric on the dashboard or report in which said view shows the data.

GeoTax search.—Type of search in IBdata that considers two sets of fields, geographic and taxonomic, in a three-level hierarchy: country, state, county, and family, genus and species. Because it is a set of fields that is very common for the user to search, the interface makes it more efficient by displaying these fields so that the user simply enters the values by which they want to filter.

Information architecture.—The way in which the interface presents the flow of information to users, either by requests that they make or by suggestion of the system.

Navigation Level 1.—Characteristic of user interface consisting of presenting most of the options to the user in a single panel, without the need to navigate through submenus. In a technical way, in IBdata it is implemented through the bifenestra user interface, that is, a two-window interface.

Responsive system.—Characteristic of the interface design that consists of adapting to the size of the screen to show the amount and shape of elements that allow more efficient and pleasant communication with the user.

Software engineering.—Computer science branch that studies and proposes methods for the complex process of computer systems development.

Thumbnail.—Small version of an image that aims to show a preview to make computing more efficient.

Usability test.—Interface design technique to verify that the interface elements are adequately understood by users; this way, possible improvements, changes, or unnecessary elements are identified.

Usability.—Degree to which specific users can use a system, product, or service to achieve specific objectives with effectiveness, efficiency, and satisfaction in a specific context of use.

User experience design (UX design).—User interface design area that applies rules, techniques, and methods to create attractive and functional interfaces for users.

User interface.—Place where the user and the system interact and which considers the necessary elements depending on the type of user, as well as the social and technological moment.