User-Centered Design and Evaluation of the Neotoma Paleoecology Open Software Ecosystem

Jonathan K. Nelson[1,3], Jessica Blois[2], Simon Goring[1], John W. Williams[1,3], Socorro Dominguez[4]

[1]Department of Geography, University of Wisconsin-Madison, Science Hall, 550 N. Park Street, Madison, WI 53706

[2]Department of Life and Environmental Sciences, University of California - Merced, 5200 N. Lake Rd., Merced, CA 95343

[3]Center for Climatic Research, University of Wisconsin-Madison, Science Hall, 550 N. Park Street, Madison, WI 53706

[4]Hari-Toto Data Services (Contractor for University of Wisconsin-Madison)

Please address all correspondence to:
Jonathan K. Nelson, PhD
University of Wisconsin-Madison
Department of Geography
Science Hall, 550 N. Park Street
Madison, WI 53706
jknelson3@wisc.edu

Abstract

Making paleoecological data accessible, usable, and useful to a global and diverse community of researchers, educators, and students across Earth sciences is an essential yet challenging task. Multiple data access and discovery tools must co-evolve with ever-changing user needs through an iterative, open-ecosystem software development approach. We employ a user-centered design study to evaluate one such ecosystem, the Neotoma Paleoecology Database, whose mission is to advance understanding of global-change processes by providing an expert-curated data resource. Neotoma contains over 11 million observations from 30,000+ datasets across 20,000+ sites, representing proxies of environmental and ecological conditions of the past. Neotoma comprises three interrelated software components – *an interactive web mapping application, statistical programming package, and API* – that provide different levels of functionality to different audiences. Although development efforts of these software components have involved the user community, Neotoma lacks a systematic evaluation process, which is important for informing the gap between perceived potential and actual utility. We address this knowledge gap using scenario-based design exercises and a usability/utility assessment. Major contributions include actionable insights for enhancing Neotoma's user experience and a user-centered design model that can be adapted to advance the development of other Earth and environmental science software ecosystems.

Keywords

user-centered design; scenario-based design; usability & utility assessment; open software ecosystems; geoinformatics; ecoinformatics

Introduction and Background

The Neotoma Paleoecology Database is an open software ecosystem whose mission is to advance understanding of global-change processes through a community-curated approach to paleoecological and paleoenvironmental data access and management (Williams et al., 2018, Goring et al., 2018). By consolidating many disparate paleoecological data sources into one centralized resource, Neotoma lowers costs of community-wide paleodata management while offering paleoecologists a high-quality data resource that enables a wide variety of macro-scale scientific research projects. For example, Neotoma data resources are being used to study global-scale biodiversity dynamics over geologic to human time scales (e.g., Blois et al., 2013; Nieto-Lugilde et al., 2021; Mottl et al., 2021; Lacourse & Adeleye, 2022; Wang et al., 2023; Stegner & Spanbauer, 2023), improve ecological forecasting (e.g., Veloz et al., 2012; Nogués-Bravo et al., 2016; Allen et al., 2019; Rollinson et al., 2021), and reconstruct past environments and landscapes (e.g., Dawson et al., 2016; Kauffman et al., 2020; Lézine et al., 2021).

Neotoma has grown significantly in data volume, proxy data types supported, and scientific impact since its launch in 2009. Neotoma holds over 11 million observations from 30,000+ datasets, 20,000+ sites, and 8,000+ publications. 2,094 new datasets were uploaded to the platform in 2023. Neotoma stores primary paleoecological data, representing biotic and abiotic proxies of environmental and ecological conditions in the past, including pollen, ostracodes, diatoms, testate amoebae, charcoal, stable isotopes, vertebrates, and geochronological data. Data mobilization campaigns over the past few years have sought to expand the Neotoma global footprint by incorporating the African Pollen, Latin American, and Indo-Pacific Pollen Databases. As a result of sustained growth in high-quality paleoecological data, Neotoma has become a prominent international resource, widely used by paleoecologists, biogeographers, and global change scientists to support research, teaching, and outreach initiatives. Neotoma-linked publications had an H-Index of 137 with 46,165 total citations in Web of Science between 1995 and June 2024 (Fig. 1). Neotoma data are of particular interest to early career scientists, who report that access to data is a primary barrier to scientific and career progress (Koch et al., 2018). Neotoma data and services are also widely used by college-level instructors in computer labs for data-driven teaching (Oliver & Graham, 2015; Goring et al., 2018).

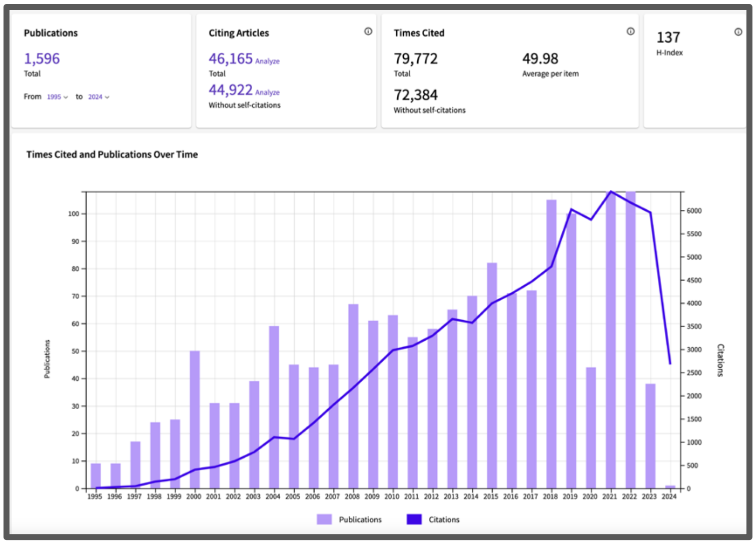


Figure 1. Distribution of Neotoma-linked publications and citations in Web of Science between 1995 and June 2024.

Making data in Neotoma accessible, usable, and useful to a global and diverse community of researchers, educators, and students is challenging and requires an open ecosystem approach to software development, in which multiple data access and discovery tools co-evolve in response to the changing needs of different audiences (Goring et al., 2022). Figure 2 depicts a schematic diagram of Neotoma's software ecosystem. Paleodata enter the ecosystem via Tilia, which is a password-protected data management tool that enables data stewards to prepare and upload data into the database. Once validated, data are stored in the Neotoma relational database, which is a PostgreSQL implementation. Neotoma data are then exposed via Neotoma's application programming interfaces (APIs) (Goring & Stryker, 2023), which support Explorer, the *neotoma2* R package (Dominguez & Goring, 2023), and a variety of third-party web applications. Explorer is an interactive web mapping application that serves a broad audience in finding, downloading, and visualizing paleodata. The *neotoma2* R package extends beyond the capabilities of Explorer by providing an interactive, programmatic experience between the user and the database, allowing for more advanced paleodata analysis and visualization. For a detailed discussion of the technical specifications of the software ecosystem see Williams et al. (2018) and Goring and Grimm (2023).

Figure 2. Schematic diagram of the Neotoma software ecosystem.

Neotoma's software ecosystem has evolved organically based on semi-structured feedback provided by the user community via Neotoma's communication channels (Slack, github, and google group), at workshops hosted by the software development team and/or expert disciplinary stewards, and through discussion with Neotoma's leadership bodies (the Neotoma Leadership Council and its Executive Working Group). Since the COVID-19 pandemic, Neotoma stakeholders have also made a concerted effort to reach communities that often cannot attend big conferences via virtual technical workshops, while also translating English presentation materials into other national languages such as Spanish. Thus, Neotoma community input is solicited from multiple angles and using diverse channels. Requested features and functionality are prioritized based on the extent to which they serve the broader user community, alignment with Neotoma's mission, and implementation feasibility. Large multi-year initiatives are scoped, parsed into manageable components, and iteratively developed with input from key stakeholders at all stages of the development process. However, although these development efforts have closely involved the user community, Neotoma lacks a process for systematic evaluation, particularly of its three core *data access and discovery components* (Neotoma Explorer, the *neotoma2* R package, and the API). Systematic evaluation is important for not just guiding future development efforts, but also gaining insight into the gap between perceived potential and reality of the open software ecosystem (Kassen 2013; Zhu & Freeman, 2019).

Here we present a two-part, user-centered design (UCD) study that aims to advance understanding of the extent to which Neotoma's three data access and delivery components meet the needs of its global user community and identify directions for future development. A UCD philosophy reflects an iterative process, in which the intended user of an interface, product, or service is involved in all stages of development from ideation to final implementation with the goal of creating a highly accessible, usable, and useful design solution (Norman 2013). Our UCD model consists of scenario-based design techniques and usability/utility assessment. In the following sections, we first present three representative use-case scenarios, one for each of Neotoma's core data access and discovery components, each of which is followed by a complementary claims analysis (Rosson & Carroll 2002) to identify critical features and functionality needed to support tasks identified in each scenario. We then present findings from a usability and utility assessment based on a structured survey completed by Neotoma users. We conclude with a synthesis and discussion of a) productive avenues for future development informed by the two-part user centered design study and b) the broader applicability of the evaluation approach and findings to other open community data resources in the Earth and environmental sciences.

User-Centered Design and Evaluation

We deployed a two-part user-centered design study to systematically evaluate Neotoma's core *data access and discovery* components including: the Neotoma Explorer user interface for exploring, finding, downloading, and visualizing data (Explorer, https://apps.neotomadb.org/explorer/); the *neotoma2* R statistical computing package (*neotoma2*, https://github.com/NeotomaDB/neotoma2); and the Neotoma application programming interface (API, https://api.neotomadb.org/api-docs/) that supports Explorer, the R package, and other custom web applications. The aims of this study were to 1) better understand the extent to which these tools are supporting research and education, and 2) identify opportunities for future development. The first part of the study employed scenario-based design techniques (Rosson & Carroll, 2002) to formatively assess Neotoma user feedback that was solicited through workshop activities, focus groups, slack correspondence, and email exchanges over a multi-year software ecosystem development cycle. Insights generated from these activities were then integrated into hypothetical use case scenarios for each of the three data access and discovery components and a supporting claims analysis to characterize user needs, exemplify design challenges, and synthesize knowledge of software use gleaned from the development process. The second part of the study consisted of a user survey distributed in March, 2023 to summatively assess the usability and utility of Neotoma Explorer, the *neotoma2* R package, and API. Figure 3 illustrates our UCD approach; methodology and results for both parts of the study are reported in the following subsections.

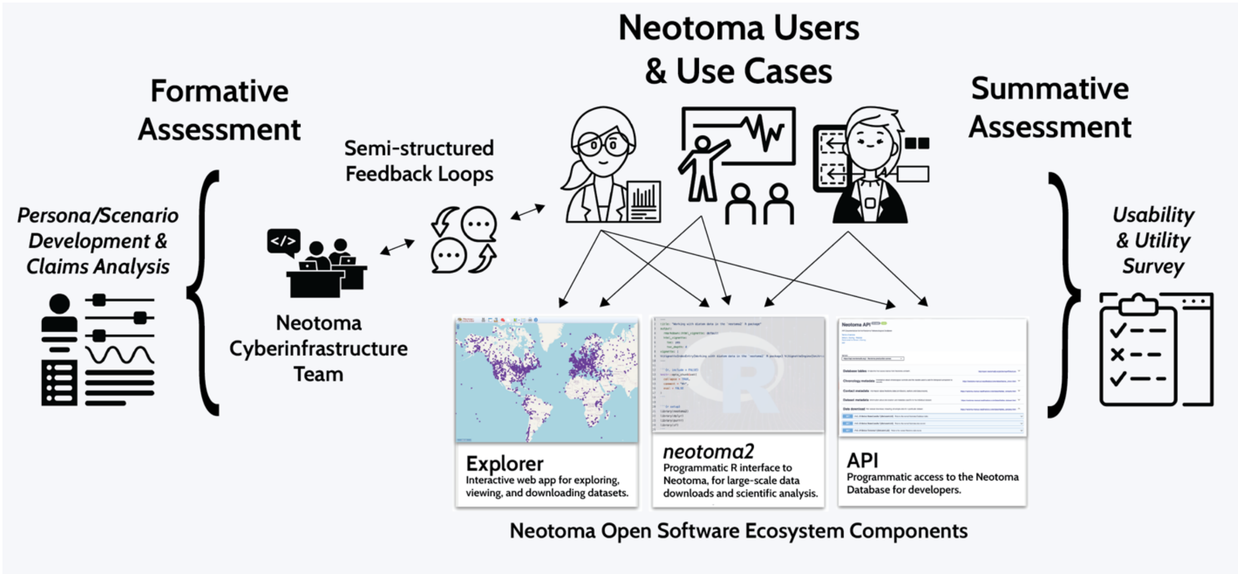


Figure 3. A conceptual illustration of our two-part UCD approach, which considers a wide range of Neotoma users (e.g., researchers, educators, students, and data scientists) and their respective use cases with Neotoma's open software ecosystem. The two-part approach consists of 1) scenario-based design techniques to formatively assess semi-structured Neotoma user feedback solicited over a multi-year software ecosystem development cycle and 2) a user survey to summatively assess the usability and utility of Neotoma Explorer, the *neotoma2* R package, and API.

Part I: Scenario-Based Design

The goal of scenario-based design (SBD) is to abstract the domain context or problem into essential tasks using a written *scenario* together with a complementary *claims analysis* that distills positive and negative attributes about the features and functionality available to support critical tasks identified in the scenario (Rosson & Carroll, 2002). SBD methods have been used by geovisualization researchers to guide and validate UCD strategies for developing geodata portals (Aditya & Kraak, 2005), as well as geovisual analytics applications that enable situational awareness during crisis events (MacEachren et al., 2011), facilitate exploration of volunteered geographic information in Open Street Map (Quinn & MacEachren, 2018), and support urban planners in making sense of large amounts of bicycling data (Nelson & MacEachren, 2020). SBD emphasizes *how* people use a system rather than documenting its features and functionality. While task-based scenarios are commonly created to guide early stages of the system development life cycle (e.g., establishing a project vision, gathering user requirements, and rationalizing a design strategy), they can also be used to support functionality specification, formative/summative assessment, and documentation at later stages of development life cycles (Rosson & Carroll, 2002; Nelson & MacEachren, 2020). The goal of SBD claims analysis is to distill a balanced view on the extent to which critical features and functionality are serving the needs of the intended audience. In this case, we developed scenarios of envisioned use to formatively assess project progress toward design goals based on semi-structured feedback provided by Neotoma's user community via workshops, focus group sessions with data stewards, slack correspondence, and email exchange. Additionally, these scenarios reflect task-based user documentation that serves as an additional help resource for a growing user community. In the following subsections, we present a “sketch of use” for each of Neotoma’s three core data access and discovery components (Explorer, *neotoma2*, API) followed by a complementary claims analysis that distills both positive and negative aspects of each software component within the context of fulfilling the user needs illustrated through the scenarios.

Neotoma Explorer Scenario

Aaron is an assistant professor of Earth science at a public land-grant university in the western United States. Aaron is currently teaching an undergraduate course on past climates and environments and wants to find an engaging way to introduce his class of 24 students to hands-on work with real-world paleodata. One of Aaron’s colleagues recently shared a link to Neotoma teaching modules available through Carleton College’s Science Education Resource Center (https://serc.carleton.edu/neotoma). Within these resources, Aaron discovers references to Neotoma Explorer and follows the link to the interactive mapping application. Aaron is initially unsure where to begin given the blank interactive map but quickly discovers the search window and is pleased to see that users can discover data based on a variety of dataset type, taxa, time, space, and metadata queries. Aaron initiates a search for “pollen” datasets in the “United States” (Fig. 4).

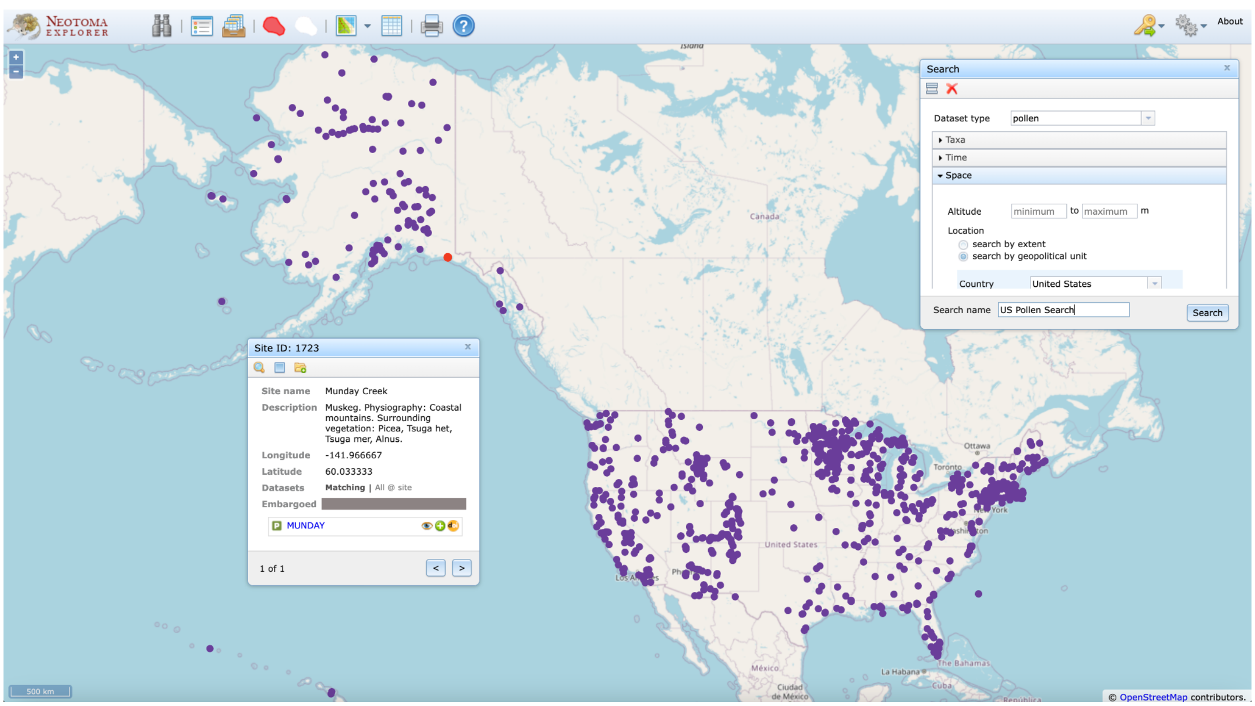


Figure 4. Neotoma Explorer search results for pollen datasets in the United States with the Munday Creek site selected.

Aaron clicks on a point in southeastern Alaska to learn more about the “Munday Creek” site, then drills further down into the dataset attributes by clicking the eye icon. Aaron really wants to be able to show his students how the abundances of plant taxa have changed over time so is particularly interested in data for a single site. The “Samples” tab includes helpful attributes but the scrolling nature of the table makes it difficult to extract any broader insights about how plant groups at Munday Creek have changed over time. Some fields lack units (e.g. depth, thickness). Fortunately, Aaron notices there is an option to export the table to a csv format and realizes there is also a much more useful visual representation of this data in the “Diagram” tab (Fig. 5).

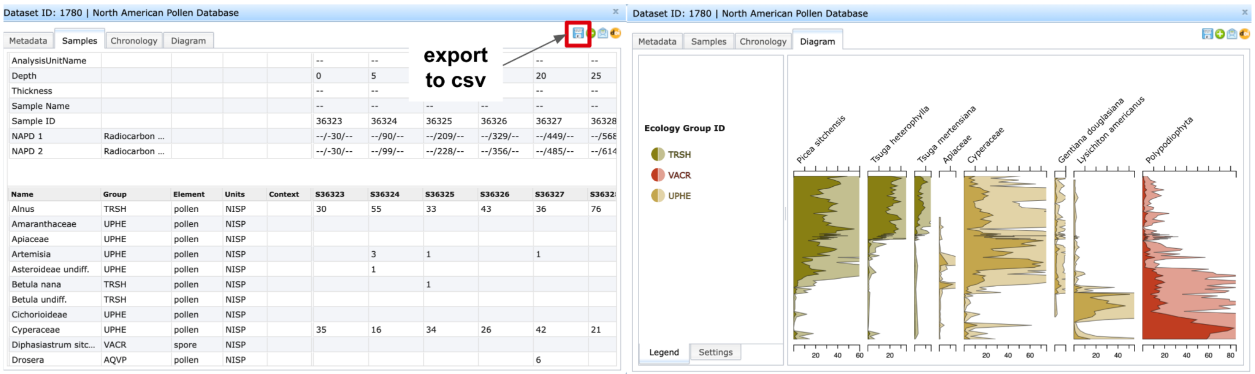


Figure 5. Neotoma Explorer Samples tab (left) and Diagram tab (right) for the Munday Creek site.

Aaron is pleased with the level of information available for pollen datasets but wants to explore another dataset type so that his students can visualize how different kinds of species responded to the effects of past environmental change. Aaron enters search parameters of dataset type = “vertebrate fauna” with a taxon specification of “Neotoma albigula” and an age range of “20,000 to 5,000 yr BP”. Aaron then clicks the search button and the user interface alerts that 11 sites have been found that meet that criteria and the map automatically zooms to the spatial bounds of the sites. Aaron clicks on a site in the vicinity of Ciudad Acuña, Mexico and learns from its metadata description that the “Hinds Cave” reflects a “Rock Shelter” depositional environment and is part of the “FAUNMAP” database. Aaron also discovers a DOI icon, which links to a landing page for the dataset and will be a great resource for teaching his students about proper data citation (Fig. 6). Lastly, Aaron explores the “Modern Ranges” layer, adding the modern range boundary for “Neotoma albigula” to help contextualize how the species has responded to past environmental change. Aaron quickly discovers that 6 of the 11 sites no longer reside within the modern range boundary, which will be a productive example to share with his class (Fig. 6).

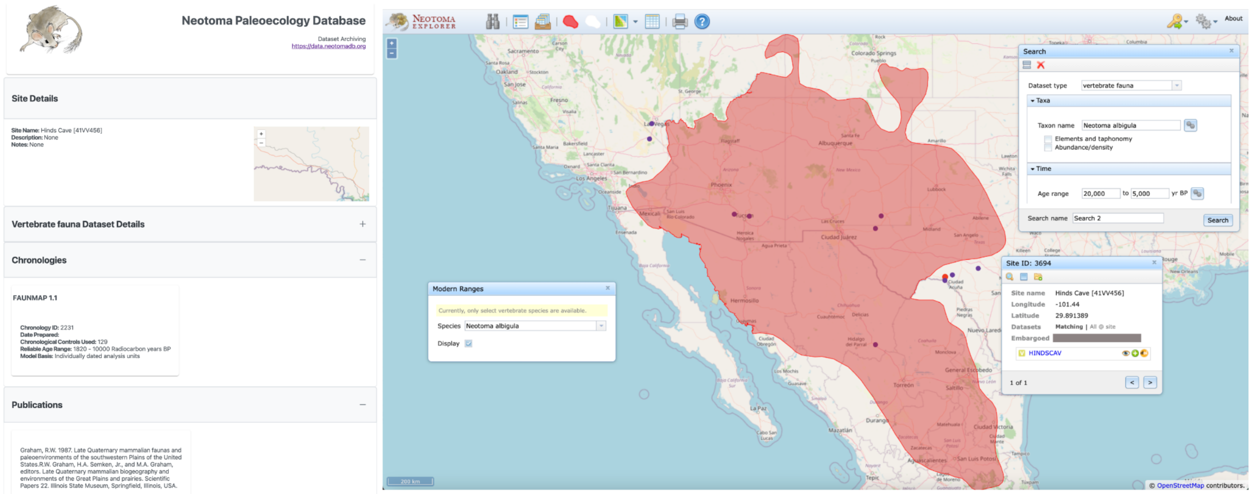


Figure 6. DOI landing page for the Hinds Cave site (left) and Explorer map of Neotoma albigula modern range boundary and past sites (right).

Neotoma Explorer Claims Analysis

Building upon the scenario presented above, we conducted a claims analysis to synthesize positive and negative aspects of Explorer's features that have important consequences for users (Table 1).

Table 1: SBD Claims Analysis for Neotoma Explorer

Component	Feature Followed by Claims
Neotoma Explorer	***Search Form*** + allows users to discover data using over 10 combinable query parameters across five categories + zooms map to bounds of search result + allows users to name, merge, and reference search layers in tabular views - forces users to perform a search prior to seeing any data on the map - contains some search categories (e.g., taxa) that require domain knowledge to use ***Map View*** + includes three base layer options and two contextual layers (vertebrate contemporary ranges and glacial boundaries) + supports selection of one or more sites via click events + allows users to modify the color and/or shape of site symbols + displays results for multiple searches simultaneously - does not encode data attributes via symbology - does not support hover events

Tabular Views

\+ enable users to view site, dataset, and sample level data attributes
\+ allow users to download sample level data
\+ link to DOI data landing pages for additional citation and publication details
\+ enable users to zoom to site bounding box on map
\- do not support downloading site and metadata attributes
\- do not expose all data attributes (e.g., specimen isotope data)
\- require users to scroll to view all sample data
\- can only view one dataset at a time
\- do not include explanatory information about acronyms in the tables (e.g. for group, units)

Stratigraphic Diagram

\+ provides effective visual summary of the changes in abundance over time at a sample level for different taxon groups
\+ allows users to parameterize the primary axis, variable unit, and chart type
\+ supports hover events to reveal details at the sample level
\- is not intuitive for certain data types (e.g., vertebrate fauna)
\- requires users to scroll to view all sample data
\- potential mismatch between chronology presented in stratigraphic v. tabular views

Chronology Graph

\+ allows users to view one or more chronologies associated with a single dataset
\+ provides notes and citation associated with each available chronology
\- has potential to plot a misalignment between age controls and age models when dates are stored in different formats
\- does not visualize for all chronologies

neotoma2 R Package Scenario

Sierra is a postdoctoral scholar at a European university working in the fields of biogeography and global change ecology. Her research interests focus on understanding ecosystem responses to abrupt climate change and she is currently examining late Holocene signals of aridity using fossil pollen data. She regularly uses the programming language R for her research analyses. Sierra decides to use the *neotoma2* R package to search for and examine data from Neotoma, so that she can more easily merge her database searches with analyses and functionality available in other R packages.

Sierra first searches for sites in the Czech Republic that have fossil pollen samples. She uses the *get_sites()* function which returns a table of 25 sites. She is puzzled because she knows that there are more sites with pollen data in the Czech Republic, but then she realizes the default is to return only 25 sites, and she has to specify *all_data = TRUE* to see all sites. Sierra re-runs the query and gets a list of 106 sites. When Sierra was looking at the function *help*, she also noticed that she could add a filter associated with age. So, she re-runs her search for all pollen data, adding an age constraint to capture sites just from the late Holocene (*ageold = 2000*). This search returns a smaller list of 58 sites that have at least one sample younger than 2,000 years before present. Sierra examines a summary of the sites and notes an apparent data inconsistency, because the summary returns a table with 68 rows, not 58 rows (Fig. 7). She checks and sees that several sites are repeated in the table, but does not yet understand why they are duplicated. Sierra also notices that the sites summary table indicates that none of the data have an associated chronology, which is confusing because the added constraint of *ageold* implies a time constraint that acted to filter the sites, indicating that there are chronologies available for all sites.

```
> length(pollen_sites)
[1] 58
> dim(summary(pollen_sites))
[1] 68  6
> summary(pollen_sites)
   siteid             sitename collectionunit chronologies datasets  types
1    1399            Kameničky          KAMEN            0        1 pollen
2    3021                Bláto         BLATO1            0        1 pollen
3    3090           Dvůr Anšov       DVURANSO            0        1 pollen
4    3170        Červené blato         JC-3-A            0        1 pollen
5    3171     Borkovická blata         JC-5-A            0        1 pollen
6    3171     Borkovická blata         JC-5-D            0        1 pollen
7    3172               Branná         JC-6-A            0        1 pollen
8    3173              Barbora         JC-6-B            0        1 pollen
9    3174          Švarcenberk       SVARCEN3            0        1 pollen
10   3174          Švarcenberk       SVARCENB            0        1 pollen
```

Figure 7. Result of the *get_sites()* function from the *neotoma2* R package. Note that the number of sites is 58, whereas the table has 68 rows. No chronologies are indicated and one pollen dataset per site is shown per row.

She consults the *neotoma2* R *help* function again and notices that another filter allows users to set the *datasettype*, and that there are both pollen and geochronologic dataset types. A single site can have multiple datasets, which explains why some sites were duplicated in the original search: there were 68 pollen datasets across 58 sites. Sierra also learns that there is a vignette available for the *neotoma2* package at https://cran.r-project.org/web/packages/neotoma2/vignettes/neotoma2-package.html. After starting to read through the vignette, she learns that the *get_datasets()* function returns a fuller set of chronology information for each site than the *get_sites()* function. To examine all possible datasets available at a site, including geochronology datasets, she needs to use the *get_datasets()* function. She performs the *get_datasets()* search using the list of sites that have pollen data (but relaxing the constraint *datasettype = pollen*). By examining the underlying structure of one of the sites within the returned datasets object, she sees that there are more site and dataset-level details available from this search. However, Sierra remains perplexed because the datasets summary still indicates there are 0 chronologies for each site, and for this search, only pollen datasets are returned, no geochronologic datasets. This is especially odd because the vignette indicated that additional chronological information would be available in the *get_datasets()* return. Through reading the full vignette, Sierra learns that the *neotoma2* functions are designed to return different levels of data and metadata. The *get_sites()* function is intended as a quick scan for relevant sites that returns minimal metadata, *get_datasets()* returns a medium amount of site- and dataset-level information, and a third function, which Sierra has not tried yet, *get_downloads()*, returns the fullest set of data for all datasets at a site. Sierra then downloads the full dataset information for all Holocene pollen sites in the Czech Republic. With this latest *get_downloads()* search, Sierra is now able to access a fuller set of the site metadata and she sees that there are now chronologies associated with sites.

Before investing more effort into analyzing the data from all 58 sites, Sierra first wants to examine the detailed data available for sites using a site that she knows well, Stará Boleslav. This is one of the sites within the *downloads object*, but she decides to re-download it as a separate object for her detailed explorations. Once she does that, she explores the chronologies at the site using the *chronologies()* function. Sierra sees that three chronologies have been defined for this site. Sierra then examines the taxa at Stará Boleslav using the built-in *taxa()* function and sees that there are 151 different taxa at the site. Sierra next uses the *samples()* function to return data from all samples for the focal site, Stará Boleslav. This function returns a large dataframe (Fig. 8): 2,464 individual observations, which record count data for all 151 taxa at each of the different sampled depths from the core at Stará Boleslav. The format is difficult to digest, because there is a lot of repeated information. It is challenging to discern which columns within the dataframe contain the most relevant sample data, and it requires substantial knowledge of the Neotoma data model (in particular, relationships among the *units*, *element*, and *taxonid* columns) to convert these data to a typical age x taxa dataframe. The samples table, however, also appears to provide Sierra what she needs for subsequent analyses (determining pollen sums and relative abundances, which she needs to create a pollen diagram for each site using the third-party package *rioja*), so she saves this as a new object. Overall, Sierra feels confident that she will be able to use the *neotoma2* package to find and then examine data from sites in order to investigate regional signals of Holocene aridity in the Czech Republic.

```
> samples.stara[1:5,]
  age                      agetype ageolder ageyounger chronologyid chronologyname units value context     element taxonid symmetry
1 385 Calibrated radiocarbon years BP       NA         NA        14591         PALYCZ  NISP     1    <NA>   ascospore   10006       NA
2 385 Calibrated radiocarbon years BP       NA         NA        14591         PALYCZ  NISP     1    <NA>   ascospore   33251       NA
3 385 Calibrated radiocarbon years BP       NA         NA        14591         PALYCZ  NISP     1    <NA> palynomorph    5391       NA
4 385 Calibrated radiocarbon years BP       NA         NA        14591         PALYCZ  NISP     1    <NA>      pollen     300       NA
5 385 Calibrated radiocarbon years BP       NA         NA        14591         PALYCZ  NISP     1    <NA>      pollen     318       NA
      taxongroup elementtype              variablename ecologicalgroup analysisunitid sampleanalyst sampleid depth thickness samplename
1          Fungi   ascospore                Ascomycota            FUNG         194585  Břízová, Eva   240212     5        NA       <NA>
2          Fungi   ascospore Microthyrium microscopicum            FUNG         194585  Břízová, Eva   240212     5        NA       <NA>
3          Algae palynomorph              Zygnema-type            ALGA         194585  Břízová, Eva   240212     5        NA       <NA>
4 Vascular plants      pollen                Thalictrum            UPHE         194585  Břízová, Eva   240212     5        NA       <NA>
5 Vascular plants      pollen                 Valeriana            UPHE         194585  Břízová, Eva   240212     5        NA       <NA>
  datasetid                 database datasettype age_range_old age_range_young                             datasetnotes siteid       sitename
1     24238 European Pollen Database      pollen          2020             447 Data contributed by PALYCZ via Kunes Petr.  15771 Stará Boleslav
2     24238 European Pollen Database      pollen          2020             447 Data contributed by PALYCZ via Kunes Petr.  15771 Stará Boleslav
3     24238 European Pollen Database      pollen          2020             447 Data contributed by PALYCZ via Kunes Petr.  15771 Stará Boleslav
4     24238 European Pollen Database      pollen          2020             447 Data contributed by PALYCZ via Kunes Petr.  15771 Stará Boleslav
5     24238 European Pollen Database      pollen          2020             447 Data contributed by PALYCZ via Kunes Petr.  15771 Stará Boleslav
       lat     long area                             sitenotes
1 50.19804 14.66948   NA Paleomeander of lower floodplain level
2 50.19804 14.66948   NA Paleomeander of lower floodplain level
3 50.19804 14.66948   NA Paleomeander of lower floodplain level
4 50.19804 14.66948   NA Paleomeander of lower floodplain level
5 50.19804 14.66948   NA Paleomeander of lower floodplain level
                                                                    description elev collunitid
1 Filled-in oxbow lake. Physiography: Flat plain. Surrounding vegetation: Cultivated fields.  167      17758
2 Filled-in oxbow lake. Physiography: Flat plain. Surrounding vegetation: Cultivated fields.  167      17758
3 Filled-in oxbow lake. Physiography: Flat plain. Surrounding vegetation: Cultivated fields.  167      17758
4 Filled-in oxbow lake. Physiography: Flat plain. Surrounding vegetation: Cultivated fields.  167      17758
5 Filled-in oxbow lake. Physiography: Flat plain. Surrounding vegetation: Cultivated fields.  167      17758
```

Figure 8. Detailed information is returned from the *neotoma2 samples()* function, showing the first five rows (observations) of the samples table, from a dataframe of 2,464 individual observations.

neotoma2 R Package Claims Analysis

Based on the scenario presented above, we again conducted a claims analysis to synthesize positive and negative aspects of the *neotoma2* R package's features that have important consequences for users (Table 2).

Table 2: SBD Claims Analysis for the *neotoma2* R Package

Component	**Feature Followed by Claims**
***neotoma2* R package**	**'get_' functions (*get_sites, get_datasets, get_downloads*)** + support extensive filtering + allow users to return searches with different amounts of information for each dataset + provide a richer set of data and metadata for each site than available through Explorer + allow data to be ported to other third-party R packages (e.g. *bchron*, *rioja*) for further analysis - difficult for new users to understand 'nested' relationship between the different 'get_' functions - help resources are split between the R function documentation, the R vignette, and GitHub, which is not clear to new users

	- not all of the search filters are documented in the help functions; some are only documented in the vignette or in help resources on GitHub
***neotoma2* R package**	**data exploration functions (e.g., *chronologies, taxa, and samples)*** + provide access to a very full set of data for each site + provide users the core data that most will want for each site - difficult to manipulate data within the sites objects without using built-in helper functions - require advanced R knowledge to manipulate the data - require some level of understanding of the Neotoma data model to contextualize the information that is returned

Neotoma API Scenario

Esther is a spatial data scientist who works for an environmental consulting firm that is part of an initiative to highlight efforts by oil producers to maintain and restore ecosystems in the region. The firm has been tasked with developing a web service that allows users to search for instances of particular species in space and in time within the spatial bounds of the Athabasca Tar Sands, and to pair that with restoration work and narrative text about restoration planning. Esther's goal is to identify the locations of the species occurrences and provide proper citations to the primary literature for users who may be interested in understanding more about the historical and paleoecological context of the region. In addition, she wishes to link species names to ecological information for those species and modern distributional data from services like the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/) and the Integrated Taxonomic Information System (ITIS; https://www.usgs.gov/tools/integrated-taxonomic-information-system-itis).

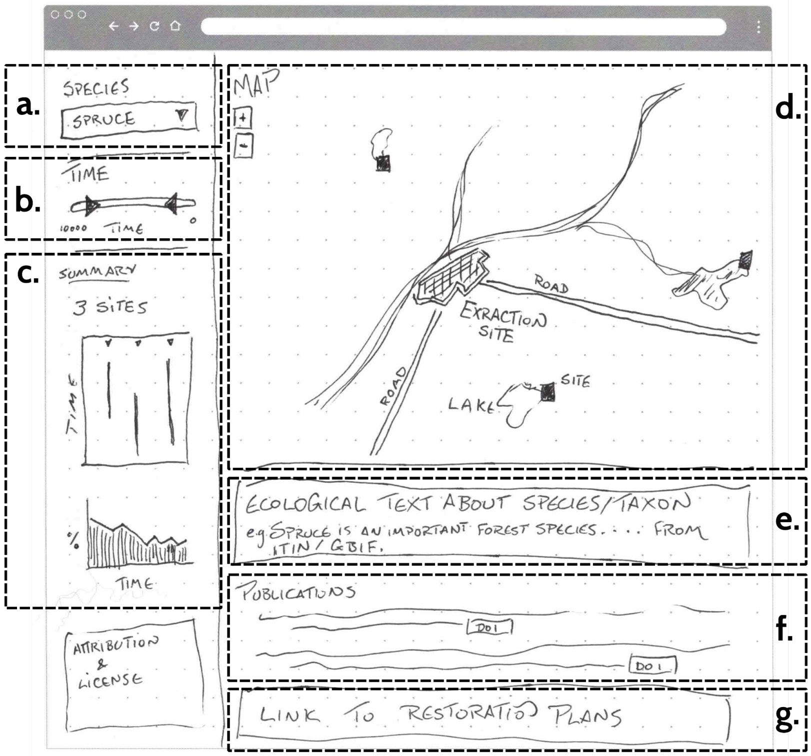


Figure 9. A hand-drawn wireframe diagram by Esther showing the layout of a hypothetical website that links paleoecological records to modern oil extraction activities and restoration efforts. The wireframe highlights the way users are expected to interact with the website: (a) selecting taxa of interest to search from Neotoma; (b) a slider bar to pick time periods of interest (over thousands of years); (c) a summary panel showing time coverage for each of the paleoecological sites, and some indication of abundance change for the taxon over time in that region; (d) an interactive web-map to show paleoecological sites of interest, extraction sites, and possibly restoration activities; (e) text related to the taxon of interest highlighting its ecological importance with external links; (f) links to publications for the paleoecological sites; (g) links to relevant restoration reports by the oil and gas industry for the region.

To undertake her work Esther plans to develop a Vue.js single page application that she has wireframed to present to project stakeholders to solicit feedback (Fig. 9). The wireframe includes a map based interface (Fig. 9d) and sliders to support search by time (Fig. 9b), along with a search bar that allows a user to enter information for particular species (Fig. 9a). To better understand what data can be obtained, and how it is represented, Esther consults the Neotoma API homepage: https://api.neotomadb.org/api-docs/.

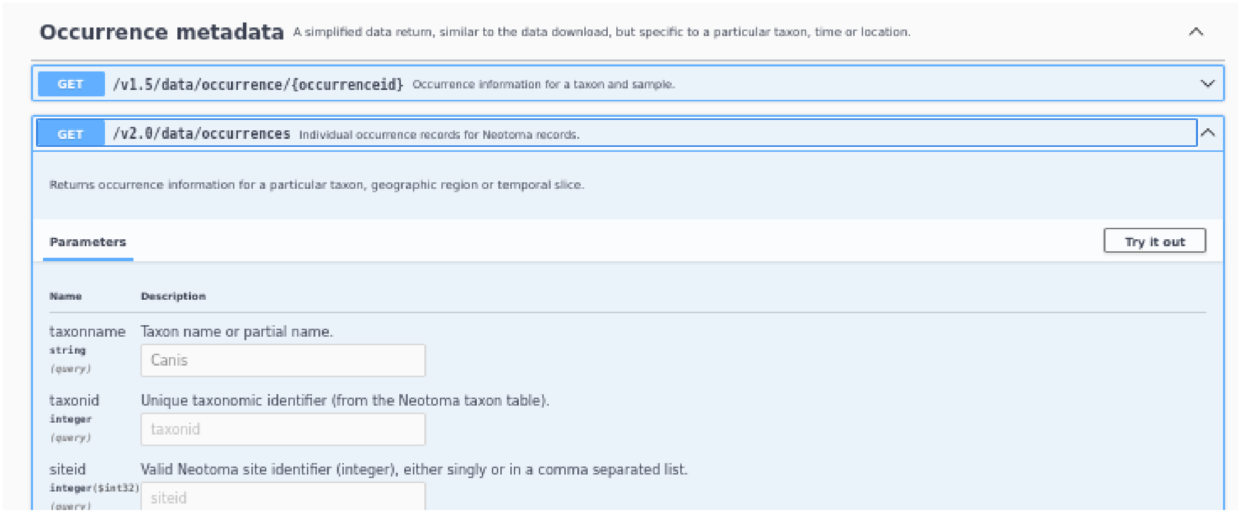


Figure 10. A subset of the Neotoma API homepage indicating some of the API routes, along with information about what the API endpoint does. Individual endpoints are identified by HTTP verb (GET, PUT, POST etc.) and can be expanded to see the parameters each endpoint takes. The “Try it out” button allows a user to execute a query and return a response from the API itself. The API homepage is rendered using the OpenAPI standard (OpenAPI v3.1.0: https://github.com/OAI/OpenAPI-Specification) and the swagger-ui-express (v5.0.0; https://github.com/scottie1984/swagger-ui-express) plugin for Node.js.

Esther looks at the API help page and notices that the heading “Occurrence metadata” seems to provide the information she is looking for (i.e., searching by taxon, time, and location). Esther expands the tab for */v2.0/data/occurrences* and is able to see the data types the endpoint accepts, some example data, and an example of the API response (Fig. 10). When she uses the “Try it out” feature in the API help she notices that the results return an error when the request is sent with the GeoJSON object to define the spatial extent of the request. However, the example returns an acceptable response when the WKT (Well Known Text) data is sent, so she chooses to use WKT spatial data for her application. From here she is able to build reactive Vue components and JavaScript functions to *fetch()* from the API, using a query structure to pass WKT data from the map bounds into the *loc* parameter, and to pass integer values from the slider into the *ageyoung* and *ageold* parameters. This produces a query similar to:

```
https://api.neotomadb.org/v2.0/data/occurrences?loc=POLYGON ((12.49
45.53, 9.847 42.879, 16.982 40.812, 20.549 44.03, 14.736 46.266, 12.49
45.53))&ageyoung=-60&ageold=11430&limit=25&offset=0
```

As Esther works on her *occurrence* queries she notices that the queries she sends using the *taxonname* parameter are not returning data she expects. For example, searching for Picea (the genera for spruce) doesn’t include the occurrences for *Picea glauca*, the specific epithet for white spruce in Alberta, Canada. The API help includes a link to the Neotoma Manual (Goring & Grimm, 2023), which contains a section explaining how taxa are managed in Neotoma. Esther chooses to provide a pre-populated list of taxa in a drop-down menu after consulting the manual (Fig. 9a). She creates this list by searching for taxa using the *v2.0/data/taxa* endpoint and grouping taxa into units, reporting them by common name, so that "Spruce" includes "*Picea*", "*Picea glauca*", "*Picea mariana,*" etc. By doing this Esther is then able to provide clearer guidance to end users, and ensure that the set of species available to users are informative and will return reliable search results.

To link to external ecological and taxonomic information Esther takes the taxonomic search from Neotoma and passes them directly into the ITIS API (https://itis.gov/web_service.html) with a separate JavaScript *fetch()* call. With her standardized set of taxa, she is able to call ITIS and obtain common names, and with the GBIF API (https://techdocs.gbif.org/en/openapi/v2/maps) she is able to overlay modern range maps of certain taxa. However, because each system uses a unique numbering system, Esther needs to take her grouped taxa and find the related identifiers for ITIS and GBIF. By selecting a small subset of species of interest she is able to do this with relatively little work.

Links to publications (Fig. 9f) are managed through the *v2.0/data/datastes/{datasetid}/publications* endpoint. As with the taxa and occurrence searches, Esther tries out the query in the API documentation. She sees that it returns a large JSON object for some datasets she tries out, but that some datasets have no responses because they are not linked to publications. Esther also sees that the response is a non-standard publication format (i.e., not BibTeX or RiS) but one that makes it relatively straightforward to re-format into an academic citation. This allows her to build a Vue component that returns article citations, and, when available, links to the article DOI, a unique web address that links to the journal article. At the same time, Esther does not realize that individual Neotoma datasets also have DOIs and that standard data citation practices can help improve outcomes for researchers and identify high value data objects within Neotoma.

In the end, Esther is satisfied that her work with Neotoma has put modern restoration efforts into a broader conservation ecology context by highlighting the past and future of the region. Her work would be further supported by direct links to ITIN and GBIF for taxonomic data in Neotoma, in addition to integration with more broadly adopted bibliographic standards, such as citation.js (Willighagen 2019) for rendering bibliographic citations.

Neotoma API Claims Analysis

Finally, we conducted one more claims analysis to synthesize positive and negative aspects of the Neotoma API (Table 3).

Table 3: SBD Claims Analysis for the Neotoma API

Component	**Feature Followed by Claims**
Neotoma API	***Spatial and temporal search (e.g., v2.0/data/occurrence/)*** + supports text-based queries for spatial polygons and temporal bounds + documented with some example data - some spatial queries may be slow depending on the complexity

	- Neotoma taxonomic structures are not straightforward and may cause confusion for web developers who are not familiar with paleoecological data
Neotoma API	***Linking to primary research/citation data for returned records* (e.g., v2.0/data/datasets/{datasetid}/publications)** + with a known dataset ID a user can return citation data and DOIs for proper data and authorship citation - not all publication data in Neotoma is properly formatted, and some records contain transcription errors or errors due to legacy data entry - proper data citation practice is not clearly identified within the API and may be unfamiliar to non-academic users
Neotoma API	***Linking between datasets/sites and taxonomic information (e.g., v2.0/data/datasets/{datasetid}/taxa)*** + with a known dataset ID a user can extract all taxa at a single site, with the taxonomic authority and ecological context of the taxon + API construction allows a user to easily return linked data using IDs returned from earlier calls - taxonomy in Neotoma is reliant on disciplinary context as documented in the Neotoma Manual, however this is not communicated in the API - taxonomy is hierarchical, however Neotoma only returns the taxon at the level identified by the investigator, thus further standardization or harmonization requires disciplinary or ecological knowledge

Part II: Usability and Utility Assessment

The second part of this user-centered design study focuses on usability and utility assessment (Ooms & Skarlatidou, 2018). Usability refers to how easy (or difficult) it is to use a tool and includes elements of learnability, memorability, efficiency of use, accuracy of results, and subjective satisfaction during use (Nielsen 1994). Utility refers to the usefulness of the tool in the context of its envisioned audience and takes into account functional scope and how many features are available to support necessary tasks (Fuhrmann 2005). The following subsections provide specifics on study design and participants, followed by a synthesis of results.

Study Design & Participants

The usability/utility assessment was administered via an anonymous online survey that was available between February 27th and April 4th, 2023. The study was reviewed by the University of Wisconsin-Madison's Minimal Risk Research Institutional Review Board and was determined to meet the criteria for exempt human subjects research based on the non-identifiable survey procedure. All participants were informed that the survey was anonymous and that their participation in the study was voluntary. The complete survey is provided in the Supplemental Material.

The survey consisted of four parts: 1) a brief background section; 2) a section focused on the Explorer application; 3) a section devoted to the R package; and 4) a section centered on the API. Each part consisted of a mix of multiple choice, short answer, and Likert-scale questions. Usability questions were based on the system usability scale (SUS; Brooke 1996). Utility questions were designed based on relevant work on user evaluation of a wide array of interactive geovisualizations (e.g., Pezanowski et al., 2018; Robinson et al., 2020; Nelson & MacEachren, 2020) but were adapted to specifically assess the extent to which Neotoma's data services improve paleodata access, enable better understanding of environmental change, facilitate interdisciplinary research, serve as relevant educational resources, and support a broad community of researchers and educators across the paleosciences.

Current Neotoma users were the target audience for the study, thus participants were recruited primarily via the Neotoma Slack space and the Neotoma Google group. Twenty members of the Neotoma community engaged in the study, completing parts of the survey that were applicable to their use cases. Over three quarters of participants had been part of the Neotoma community for >5 years while three participants were first-year members. Figure 11 provides a visual summary of participants' demographic attributes and areas of expertise.

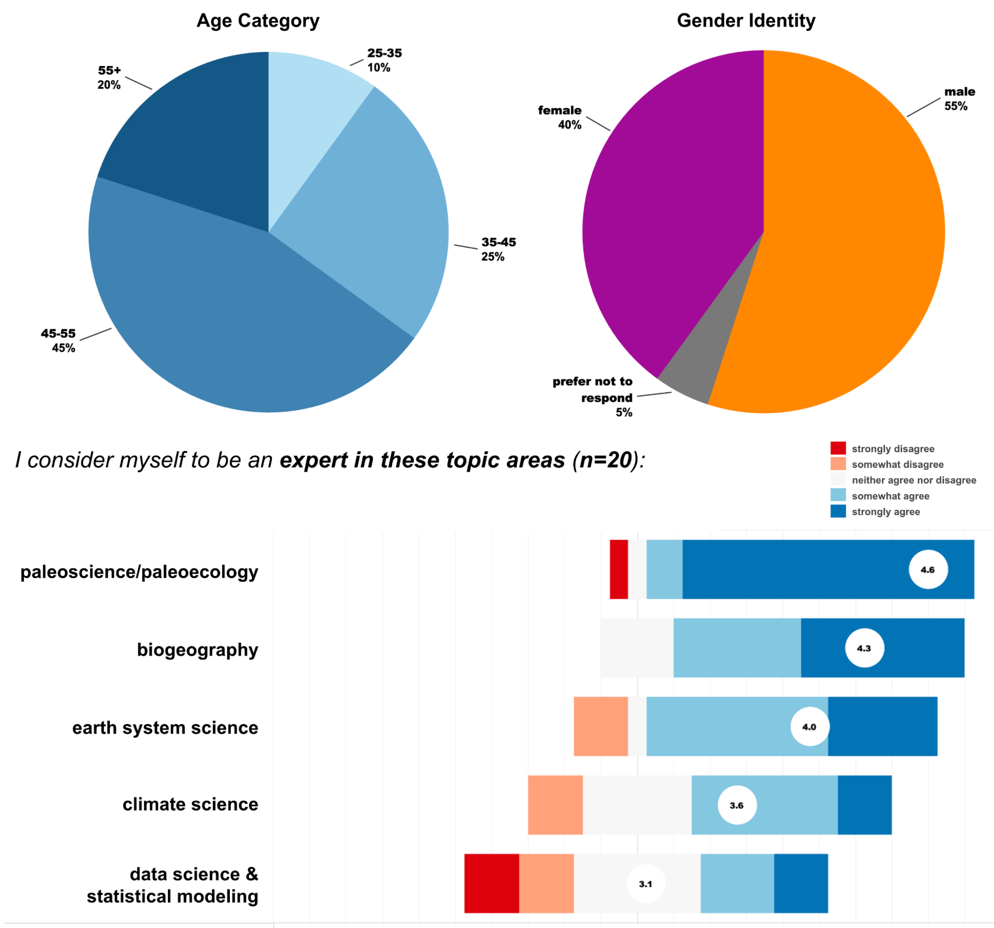


Figure 11. Visual summary of participants' demographic attributes and areas of expertise.

The majority of participants were 45 years of age or older. 35% of participants were 25-45 years of age. Just over half of participants identified as being male while 40% participants identified as being female; 5% preferred not to respond. 18 participants reported being researchers while nine of the 20 identified as being an educator. Only one participant identified as being a data scientist and no participants identified as being students. Most participants reported professional interests within the relevant domain sciences with an emphasis on paleoecology, palynology, and climate change applications. On average, most respondents strongly agreed to possessing expertise in palaeosciences and biogeography, somewhat agreed to being experts in Earth system science and climate science, and were on average less confident in their abilities in data science and statistical modeling.

Usability/Utility Results

With respect to overall service usage trends, 19 of the 20 participants reported using Explorer, nine reported using the *neotoma2* R package, and four participants indicated that they used the API. This trend aligns with the underlying goal of these services: Explorer is designed to serve a broad audience with easy data access whereas the R package and API are designed to provide a smaller group of power users with more advanced analysis capabilities and a means to develop custom paleodata applications.

Neotoma Explorer Findings

Prior to soliciting input on Explorer's usability and utility, participants were first asked about frequency of use and which "search" and "data discovery" strategies were most commonly employed. The intent of these questions was to establish a better understanding of how Neotoma users were engaging with the Explorer application. Of the 19 participants who reported using Explorer, 37% indicated weekly use, 21% indicated monthly use, and 24% reported only using the application a few times throughout the year. Search by "dataset type" was the most common search technique, with 95% of respondents indicating use of this functionality. Search by "space" and "taxa" tied for second most common search strategy with 68% of participants reporting use of these techniques. Searching by "metadata" and "time" ranked third and fourth, with 47% and 21% of participants reporting use of these techniques respectively. Once results were returned from various search parameters, participants most commonly used the "Metadata," "Chronology," and "Samples" tabs to view data attributes (Fig. 12). Over 50% of participants also reported using the "Diagram" tab and downloading data extracts for use in other applications.

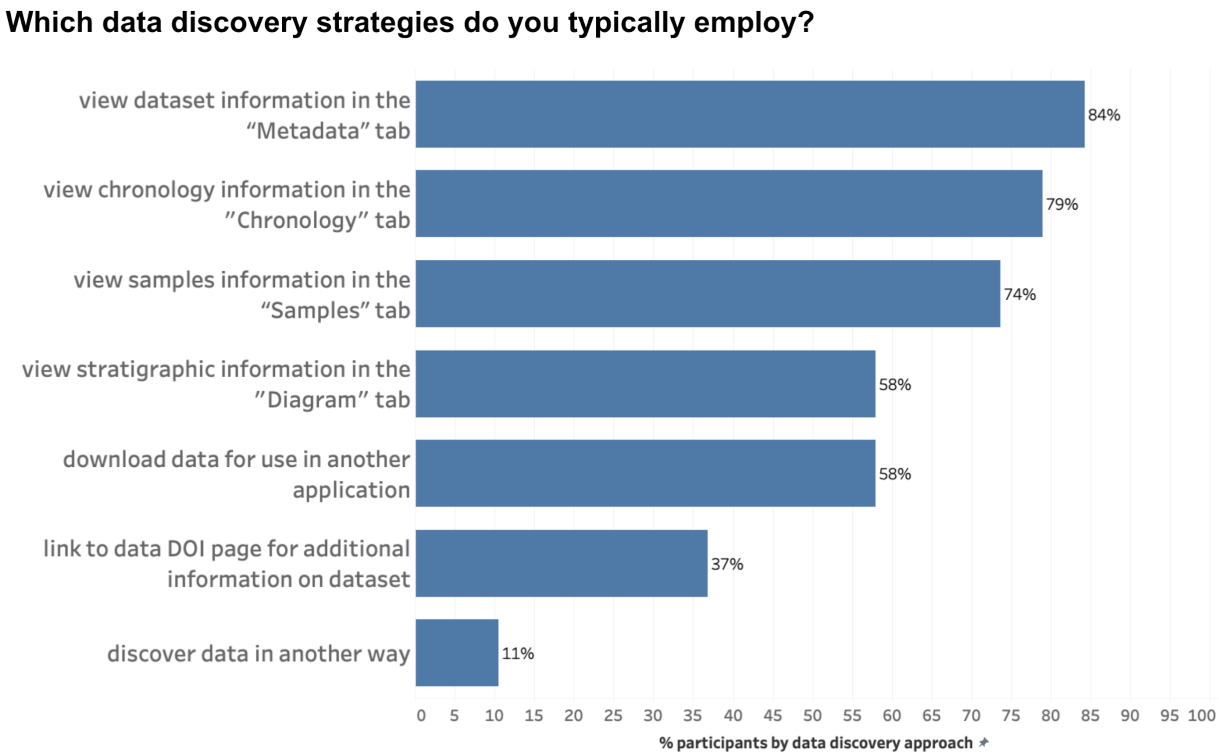


Figure 12. Visual summary of participants' data discovery strategies using Neotoma Explorer.

To systematically assess the usability of Explorer, participants were asked to rank their agreement to seven positive and three negative questions (adapted from the SUS; Fig. 13). On average, participants strongly or somewhat agreed to all positive questions and strongly or somewhat disagreed to all negative questions. For example, all participants agreed that most people could learn to use Explorer quickly and felt confident using the application, whereas participants generally disagreed that Neomtoma is unnecessarily complex and cumbersome to use. Likert scale responses were converted to numbers and used to calculate an overall usability score of 77.6 which is in the 80-84th percentile and considered a B+ or "good" from a usability perspective.

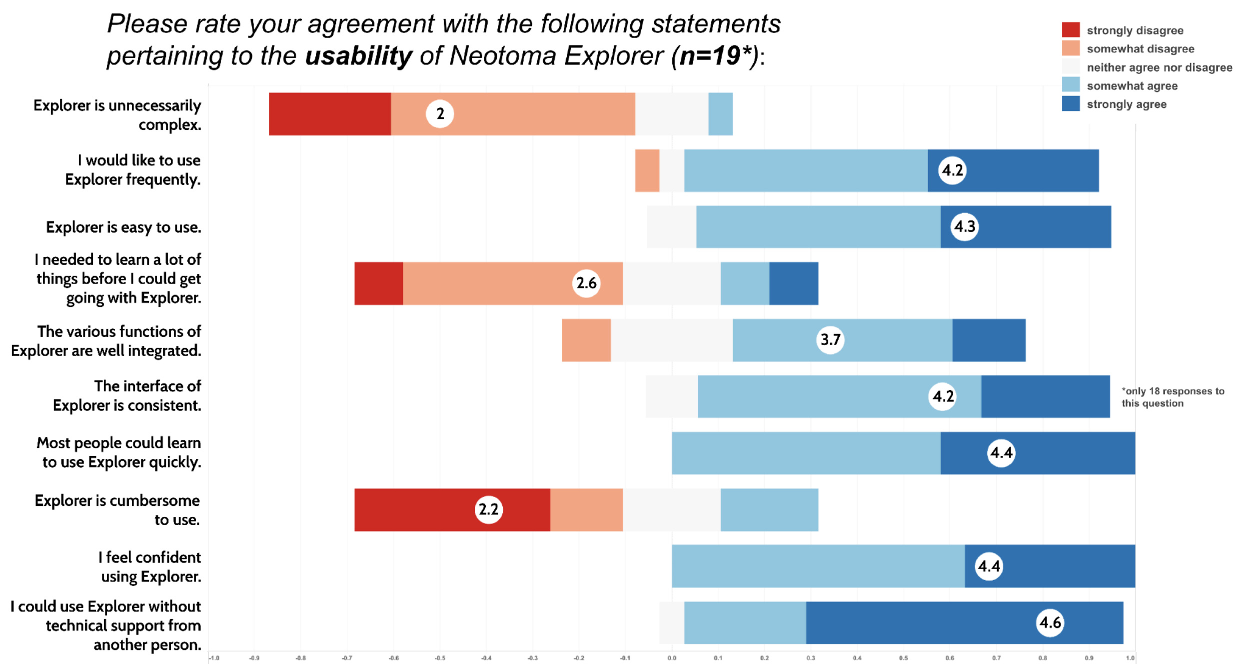


Figure 13. Visual summary of usability ratings for Neotoma Explorer.

Participants were also asked to rate their agreement to five utility-focused questions (Fig. 14). On average, all participants strongly agreed that Explorer improves data access, enables better understanding of environmental change, facilitates interdisciplinary research, serves as a relevant educational resource, and supports a broad community of researchers and educators across the paleosciences.

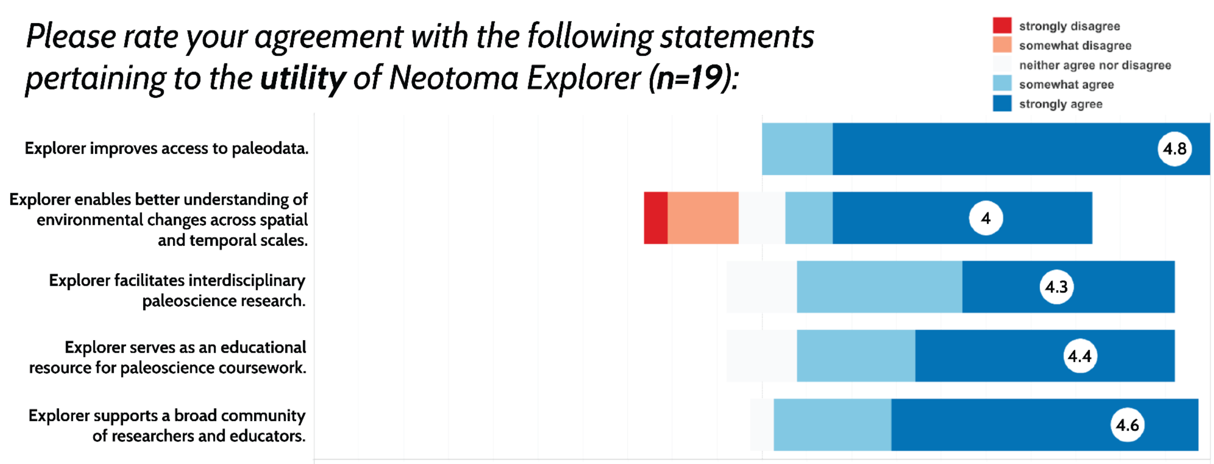


Figure 14. Visual summary of utility ratings for Neotoma Explorer.

neotoma2 R Package Findings

Of the nine participants who reported using the *neotoma2* R package, one participant reported using the package daily, five reported using it monthly, and three reported using it a few times throughout the year. When asked which query strategies were most typically employed, all respondents indicated using the *get_downloads()* function and over three-quarters of respondents reported also using the *get_datasets()* and *get_sites()* queries.

The ten usability and five utility questions were again asked in the context of the *neotoma2* R package (Figs. 15 and 16). While the number of responses does not support calculating an overall usability score, findings suggest that while most respondents strongly agreed with wanting to use the *neomtoma2* R package frequently, there was general disagreement to questions asking if people could learn how to use the package quickly and without technical support from another person. Participants' opinions were split when assessing how easy the R package is to use, how integrated its various functions are, and whether or not it is unnecessarily complex. Despite these mixed usability results (Fig. 15), there was overall strong agreement that the *neotoma2* R package has high utility (Fig. 16): it improves data access, enables better understanding of environmental change across space and time, facilitates interdisciplinary research, serves as a relevant educational resource, and supports a broad community of researchers and educators across the paleosciences.

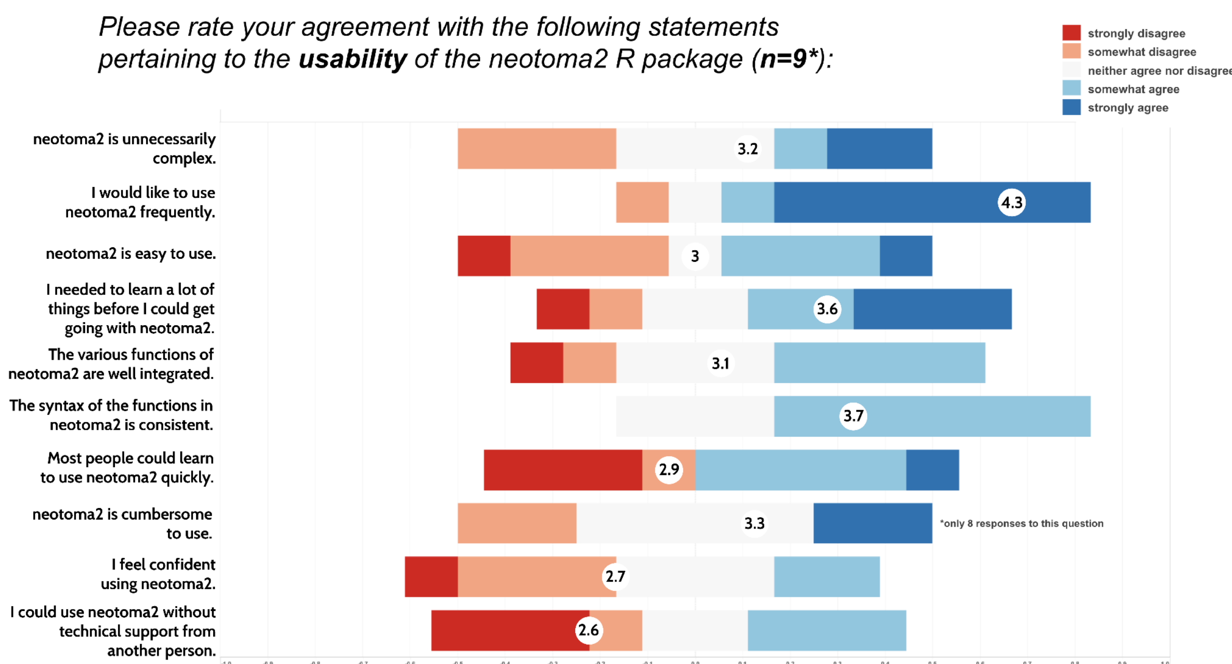


Figure 15. Visual summary of usability ratings for the *neotoma2* R package.

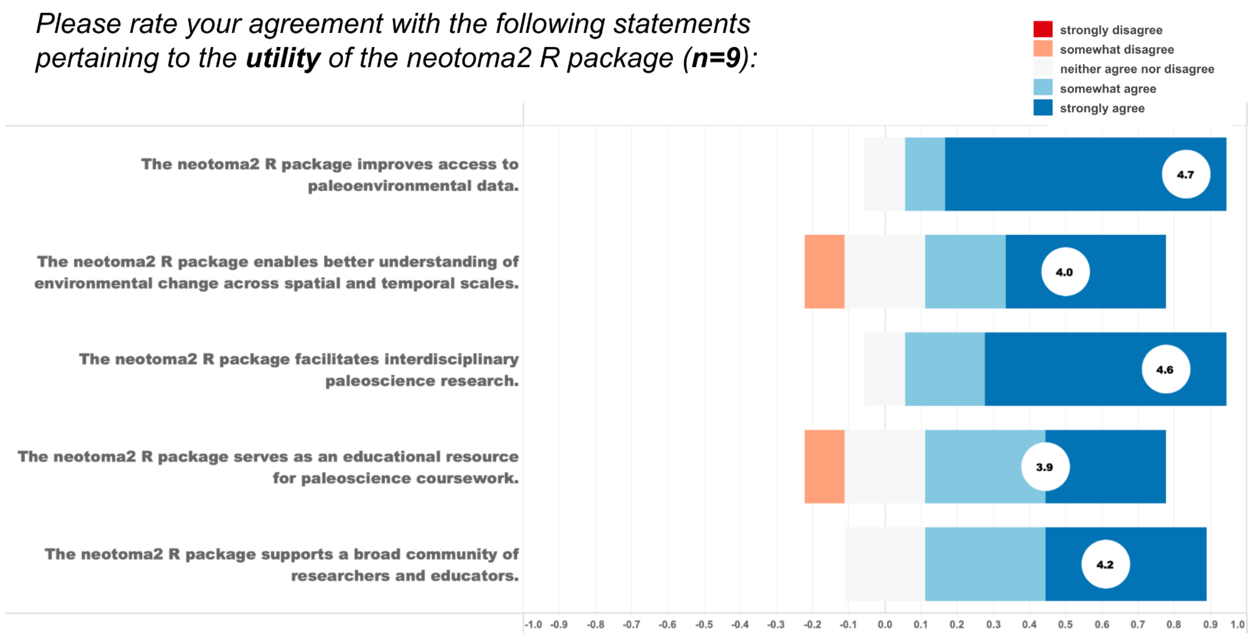


Figure 16. Visual summary of utility ratings for the *neotoma2* R package.

Neotoma API Findings

The final part of the survey sought to better assess how a select subset of the Neotoma community is leveraging the API in their scientific workflows and application development efforts. Because only four survey participants reported using the API, we only summarize usability and utility insights (Figs. 17 and 18). With respect to usability, a couple key takeaways emerged. First, most participants agreed that the various functions of the API are well integrated and the syntax of the endpoints is consistent. Second, despite participants indicating that they wanted to use the API frequently, some participants also expressed barriers in the learning process and lack of confidence using the API.

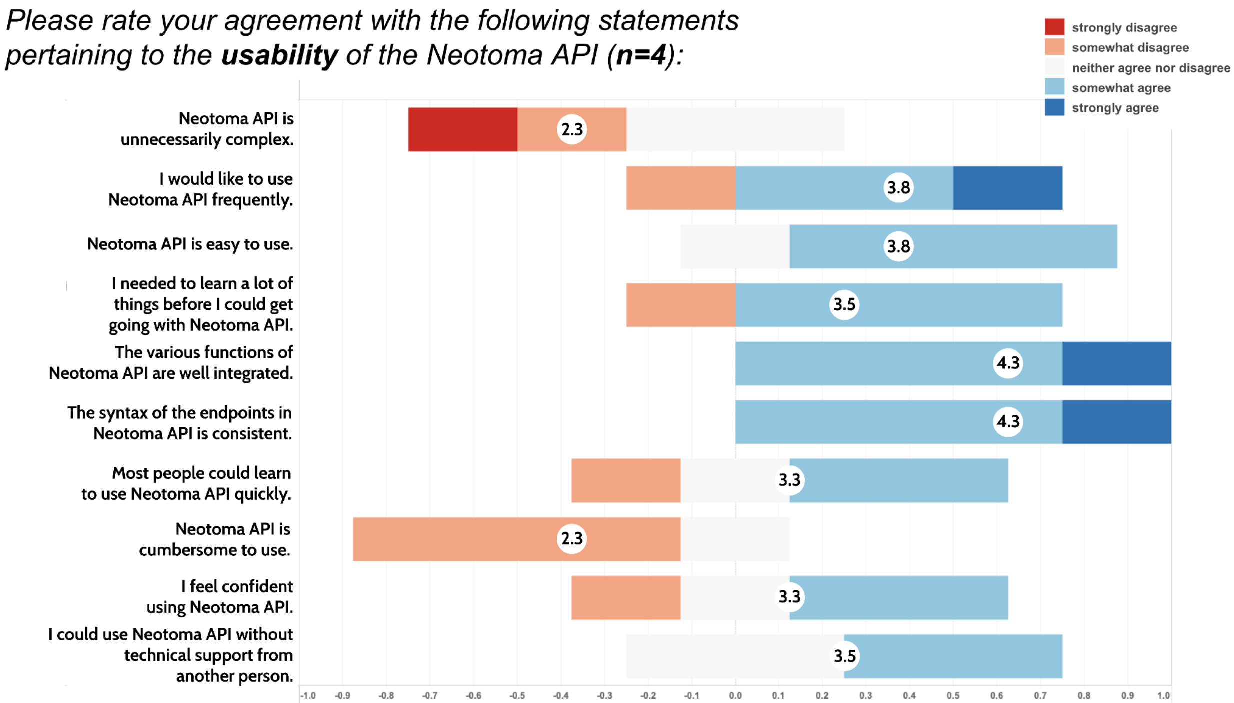


Figure 17. Visual summary of usability ratings for the Neotoma API.

With respect to utility, all participants strongly agreed that the API improves access to paleoenvironmental data. On average, participants also agreed that the API enables better understanding of environmental change, facilitates interdisciplinary research, and supports a broad community of researchers and educators across the paleosciences. The majority of participants, however, somewhat disagreed that the API serves as an educational resource, which is unsurprising considering the primary audience and goal of the API.

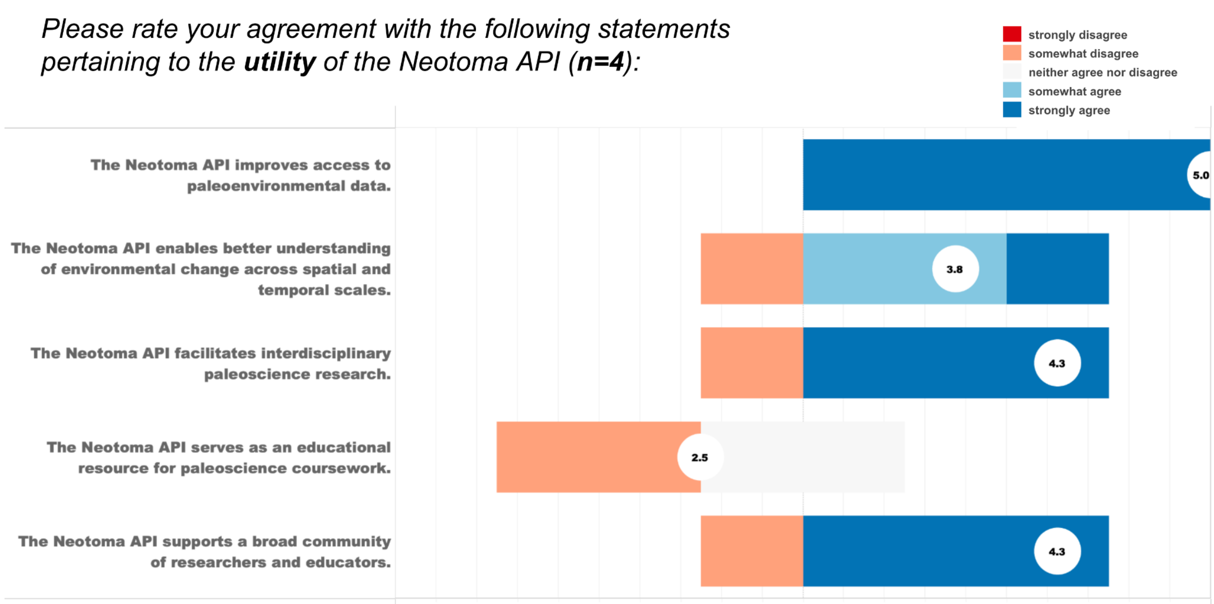


Figure 18. Visual summary of utility ratings for the Neotoma API.

Discussion

Findings from the user-centered design study offer many important insights into the extent to which Neotoma's core *data access and discovery* components support research and education activities and how they can be enhanced to improve usability and utility. The scenario-based design exercises provided a balanced view of how well specific features and functionality of each component meet the needs of the representative audiences. The personas and scenarios illustrate representative use cases and make important distinctions between the audiences served by each component and the typical pain points they encounter. Neotoma Explorer, for example, is designed to provide a broad audience of educators, researchers, and students with easy access to paleodata; whereas, the *neotoma2* R package and API are designed to enable a smaller group of power users to conduct advanced analysis and develop custom, third-party applications. Following each scenario is a complementing claims analysis that relates positive and negative characteristics of specific features and functionality to the respective use cases. The process of creating scenarios and translating them into individual claims about specific features enables the Neotoma development team to clearly identify system shortcomings and prioritize development efforts based on user-informed needs. Indeed, some of these shortcomings (e.g., inaccessibility of *neotoma2* R package vignettes) were subsequently targeted for priority updates through the process of generating the scenarios. These scenarios also helped to contextualize larger, more complex system challenges, such as the need to better support taxonomic harmonization. This needs assessment can serve as a frame of reference for collaborative efforts among Neotoma data stewards, stakeholders, and developers moving forward. Finally, Neotoma's growing user community can further reference these scenarios and claims analyses to orient themselves with the platform and learn about the features and functionality currently supported across the three core services.

A commonality across all scenarios and claims analyses is a tension between usability and utility. What a beginner may describe as a 'bug' while using a data access and discovery tool might also be described as a 'feature' to a more advanced user. The user survey identified focused opportunities to resolve this tension for each of the three core data access and discovery components. For example, Neotoma Explorer was rated "good" in the context of the SUS; however, the aggregated utility results suggest an opportunity for creating new features and functionality to improve users' understanding of environmental change across both spatial and temporal scales. One strategy for addressing this limitation is to strategically integrate Explorer with related third-party applications such as Flyover Country (Loeffler et al., 2017) and Range Mapper (George et al., 2023), which are designed to visualize how ecological patterns and trends dynamically change over both space and time. Enabling Explorer to connect with other databases, such as NOAA and Pangaea, would further enrich understanding of paleoclimatic and paleoenvironmental data. Another opportunity would be to create new services that cross-reference information about spatiotemporal distributions of taxa (from Neotoma) with independent information about past climates and human activity, thus retaining Explorer's ability to serve a broad audience through a focus on interactive data exploration and strong usability.

Aggregated results from the usability assessments also revealed opportunities to make the R package and API less complex and easier to use, through both refinement of these services and better documentation. Given that most participants indicated that the various functions of these services are well integrated, future development efforts should prioritize documentation and example workflows to improve ease of use and bolster user confidence. Striking the right balance between usability and utility will be important to both expand the audience of these services, while also ensuring their advanced capabilities are retained for developers.

This tension between usability and utility is common in the software development process, because designers are forced to grapple with the tradeoffs between a less powerful, more usable, and widely adopted solution versus a more powerful, less intuitive, and limited audience solution (Grinstein et al., 2003; Hart et al., 2022). The Neotoma platform attempts to resolve some of this tension by leveraging the FAIR data principles (Wilkinson et al., 2016) and creating innovative infrastructure to improve data **f**indability, **a**ccessibility, **i**nteroperability, and **r**eusability. Other data frameworks such as CARE (Carroll et al., 2020) and ICON principles (Goldman et al., 2021, Koren et al., 2022) can provide additional perspectives that extend beyond data infrastructure considerations. These alternate frameworks promote more ethical open science by putting both data generators and those potentially reflected in the data (e.g., indigenous communities) together with the context of their science and lived experiences at the center of the process. Future development efforts can extend the *user*-centered design concept to the *human*-centered design concept to more deeply consider how open paleodata software ecosystems can better meet the needs of not just those who use these systems but who may be impacted by those who do.

Finally, it is important to acknowledge that the Neotoma user community is much larger than the 20 members who agreed to participate in the usability/utility survey. While we purposely recruited survey participants from Neotoma community outlets (e.g., the dedicated slack space, google group, etc.), we recognize that our sample size cannot adequately represent the wealth of diversity in domain expertise, background experiences, teaching/research interests, and technical skills of the entire community. Survey results from the demographic and professional background questions suggest we reached a wide range of users across different ages, genders, and expertise areas in addition to some variation in how long participants have been part of the Neotoma user community with a bias towards participants who have been members 5+ years. With respect to occupation, participants were also relatively balanced between researchers and educators; however, participation from students and data scientists was limited. From a usability evaluation perspective, five is typically the minimum number of participants needed to uncover ~80% of usability issues with diminishing returns in insights gleaned from including more participants in a study (Nielsen 2009). Thus, future participant recruitment strategies should prioritize inclusion of at least five participants across relevant audience categories for each relevant data access and discovery component being evaluated. Given the significance of open (paleo)data software ecosystems such as Neotoma to early career scientists (Koch et al., 2018), we further recognize the need for distinguishing this audience from the broader category of “researchers” in future design studies.

Conclusion

This problem-driven research aimed to inform the question of *how to improve open paleodata software ecosystems for a global audience with a broad range of technical and scientific backgrounds?*. We evaluated the accessibility, usability, and utility of one such system, the Neotoma Paleoecology Database, which has three interrelated data access and discovery components – *an interactive web mapping application, statistical programming package, and API* – that provide different levels of functionality and target different audiences. Based on our work, we offer four key takeaways that may be useful when designing data access and discovery components for Earth science applications with similar audiences and goals:

1. ***enhance data visibility*** — Design strategies should strive to expose all relevant metadata attributes, geospatial characteristics, uncertainty measures, and data provenance information. Enhanced data visibility increases user awareness and builds system trust (Sacha et al., 2015).
2. ***extend search & filter capabilities*** — Users benefit from having many avenues to discover data (e.g., by unique IDs, proxy types, taxonomy, spatiotemporal constraints, etc.). If search and filter capabilities are employed using interactive visualization, use the visual-information seeking mantra of "overview first, zoom and filter, details on demand" (Shneiderman 2003) as a starting point to help guide design decisions.
3. ***improve data download methods*** — Some use cases require porting data from one application to another, in which case it is important to remove any object redundancies, minimize file size, and support multiple output formats.
4. ***emphasize learnability*** — System learnability is one of the most important components of usability (Nielsen 1994). Documentation should be accessible, centralized, and up-to-date, having clear variable and function definitions. Leverage vignettes and scenarios of envisioned use to provide illustrative examples of how to get started with system components.

While results from user-centered design studies are not generalizable since they tend to inform specific situations, findings may be transferable to similar use cases (Sedlmair et. al, 2012). Thus, the UCD model presented here can be adapted to inform the development of other open Earth and environmental science portals. Improving the design of open Earth science software ecosystems promotes data findability, accessibility, interoperability, and reusability (FAIR principles), and in this particular use case, fosters a better understanding of the physical environment and global change processes, which has broad implications for society, policymakers, and the future of our planet.

Acknowledgements

This user-centered design study was supported by the NSF Geoinformatics Program under Grants 1948926, 1948229, and 1948579. Additional support for this research was provided by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation. Neotoma contains data obtained on lands that have a history of colonization and displacement. The work of data contributors, data stewards, and broader community in building the Neotoma Paleoecology Database is gratefully acknowledged. We thank Mike Anderson, Brian Bills, Steve Crawford, Doug Miller, and Mike Stryker for their work building and maintaining Neotoma Explorer. We also thank all participants of the Neotoma user survey and the anonymous reviewers for their insightful feedback.

Author Contributions

Jonathan Nelson led the conceptualization, administration, and analysis of the user-centered design study; development of the Explorer scenario and supporting claims analysis; and development of the manuscript. Jessica Blois provided input on study design and led the development of the *neotoma2* R package scenario and supporting claims analysis. Simon Goring provided input on study design and led the development of the Neotoma API scenario and supporting claims analysis. John Williams provided input on study design, assisted in disseminating the survey, and guided the direction of the Explorer scenario and supporting claims analysis. Socorro Dominguez provided input on study design and contributed to the development of the *neotoma2* R package scenario and supporting claims analysis. All authors contributed to creating figures and writing the paper.

Data Availability Statement

Data in the Neotoma Paleoecology open software ecosystem are available as database snapshots (https://www.neotomadb.org/data/db-snapshots) or accessible via Explorer (https://apps.neotomadb.org/explorer/), the neotoma2 R package (https://cran.r-project.org/web/packages/neotoma2/vignettes/neotoma2-package.html), and the API (https://api.neotomadb.org/api-docs/). Individual datasets within Neotoma are referenceable by DOIs (e.g., https://doi.org/10.21233/n39w2j). The participants of this study did not give written consent for their data to be shared publicly, so individual results from the user survey are not available.

Declaration of Interest

The authors report there are no competing interests to declare.

References

Aditya, T., & Kraak, M. J. (2005, June). Reengineering the Geoportal: Applying HCI and Geovisualization Disciplines. In *Proceedings of the 11th EC-GI & GIS Workshop*, Alghero, Italy (Vol. 29).

Allen, J. M., Folk, R. A., Soltis, P. S., Soltis, D. E., & Guralnick, R. P. (2019). Biodiversity synthesis across the green branches of the tree of life. *Nature Plants, 5*(1), 11-13.

Blois, J. L., Williams, J. W., Fitzpatrick, M. C., Jackson, S. T., & Ferrier, S. (2013). Space can substitute for time in predicting climate-change effects on biodiversity. *Proceedings of the National Academy of Sciences, 110*(23), 9374-9379.

Brooke, J. (1996). Sus: a 'quick and dirty' usability scale. *Usability evaluation in industry*, *189*(3), 189-194.

Carroll, S. R., Hudson, M., Holbrook, J., Materrechera, S., & Anderson, J. (2020). Working with the CARE principles: operationalizing Indigenous data governance. *Ada Lovelace Institute*, *9*.

Dawson, A., Paciorek, C. J., McLachlan, J. S., Goring, S., Williams, J. W., & Jackson, S. T. (2016). Quantifying pollen-vegetation relationships to reconstruct ancient forests using 19th-century forest composition and pollen data. *Quaternary Science Reviews, 137*, 156-175.

Dominguez, S., & Goring, S. J. (2023). neotoma2: An R package to access data from the Neotoma Paleoecology Database. *Journal of Open Source Software*, 8(91), 5561.

Fuhrmann, S., Ahonen-Rainio, P., Edsall, R.M., Fabrikant, S.I., Koua, E.L., Tobón, C., Ware, C. and Wilson, S. (2005). Making useful and useable geovisualization: design and evaluation issues. In *Exploring geovisualization* (pp. 551-566). Elsevier.

George, A. K., Roth, R. E., Widell, S., & Williams, J. W. (2023). Range Mapper: An Adaptable Process for Making and Using Interactive, Animated Web Maps of Late-Quaternary Open Paleoecological Data.

Goldman, A. E., Emani, S., Perez-Angel, L. C., Rodriguez_Ramos, J., Stegen, J. C., & Fox, P. (2021). Special collection on open collaboration across geosciences. *Eos, Transactions American Geophysical Union (Online)*, *102*(PNNL-SA-158823).

Goring, S. J., Graham, R., Loeffler, S., Myrbo, A., Oliver, J. S., Ormond, C., & Williams, J. W. (2018). The Neotoma paleoecology database: *A research outreach nexus*. Cambridge University Press.

Goring, S. J., & Grimm, Eric C. (2023). Neotoma Paleoecology Database SQL Database Manual (v0.2.0). Zenodo. https://doi.org/10.5281/zenodo.10201999

Goring, S. J., & Stryker, M. (2023). The Neotoma Paleoecology Database API v2.0.0 (Version 2.0.0) [Computer software]. https://doi.org/10.5281/zenodo.7893836

Goring, S. J., Vidaña, S.D., Blois, J., Crawford, S., Nelson, J. K., Thomer, A., & Williams, J. W. (2022, December). Advancing Interdisciplinary Global Change Science Through Linked

Research Services in the Neotoma Paleoecology Database. In *AGU Fall Meeting Abstracts* (Vol. 2022, pp. IN35B-0408).

Grinstein, G., Kobsa, A., Plaisant, C., Shneiderman, B., & Stasko, J. T. (2003, October). Which comes first, usability or utility?. In *Visualization Conference, IEEE* (pp. 112-112). IEEE Computer Society.

Hart, D. A., Prestby, T., & Roth, R. E. (2022). Design and evaluation of coastal web atlases: Best practices and future opportunities for map representation, interaction, and usability. *Coastal Management, 50*(6), 514-548.

Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government information quarterly, 30*(4), 508-513.

Kaufman, D., McKay, N., Routson, C., Erb, M., Davis, B., Heiri, O., Jaccard, S., Tierney, J., Dätwyler, C., Axford, Y. & Brussel, T. (2020). A global database of Holocene paleotemperature records. *Scientific data, 7*(1), 115.

Koch, A., Glover, K. C., Zambri, B., Thomas, E. K., Benito, X., & Yang, J. Z. (2018). Open-data practices and challenges among early-career paleo-researchers. *Past Global Change Magazine, 26*(2), 54-55.

Koren, Gerbrand, Vincenza Ferrara, Madeleine Timmins, and Monica Ainhorn Morrison. "Global Environmental Change Perspectives on Integrated, Coordinated, Open, and Networked (ICON) Science." *Earth and Space Science* 9, no. 5 (2022): e2022EA002231.

Lacourse, T., & Adeleye, M. A. (2022). Climate and species traits drive changes in Holocene forest composition along an elevation gradient in Pacific Canada. *Frontiers in Ecology and Evolution, 10*, 838545.

Lézine, A. M., Ivory, S. J., Gosling, W. D., & Scott, L. (2021). The African Pollen Database (APD) and tracing environmental change: State of the Art. *Quaternary Vegetation Dynamics*, 5-12.

Loeffler, S., Thorn, R., Myrbo, A., Roth, R., Goring, S. J., & Williams, J. (2017, December). Multivariate spatiotemporal visualizations for mobile devices in Flyover Country. In *AGU Fall Meeting Abstracts* (Vol. 2017, pp. IN31D-05).

MacEachren, A.M., Jaiswal, A., Robinson, A.C., Pezanowski, S., Savelyev, A., Mitra, P., Zhang, X. & Blanford, J. (2011, October). Senseplace2: Geotwitter analytics support for situational awareness. In *2011 IEEE conference on visual analytics science and technology (VAST)* (pp. 181-190). IEEE.

Mottl, O., Flantua, S.G., Bhatta, K.P., Felde, V.A., Giesecke, T., Goring, S., Grimm, E.C., Haberle, S., Hooghiemstra, H., Ivory, S. and Kuneš, P. (2021). Global acceleration in rates of vegetation change over the past 18,000 years. *Science, 372*(6544), 860-864.

Nelson, J. K., & MacEachren, A. M. (2020). User-centered design and evaluation of a geovisualization application leveraging aggregated quantified-self data. *Cartographic Perspectives*, (96), 7-31.

Nielsen, J. (1994). Usability engineering. Boston, MA: Academic Press, Inc

Nielsen, J. (2009). Discount usability: 20 years. Jakob Nielsen's Alertbox Available at http://www.useit.com/alertbox/discount-usability.html

Nieto-Lugilde, D., Blois, J. L., Bonet-García, F. J., Giesecke, T., Gil-Romera, G., & Seddon, A. (2021). Time to better integrate paleoecological research infrastructures with neoecology to improve understanding of biodiversity long-term dynamics and to inform future conservation. *Environmental Research Letters, 16*(9), 095005.

Nogués-Bravo, D., Veloz, S., Holt, B.G., Singarayer, J., Valdes, P., Davis, B., Brewer, S.C., Williams, J.W. & Rahbek, C. (2016). Amplified plant turnover in response to climate change forecast by Late Quaternary records. *Nature Climate Change, 6*(12), 1115-1119.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.

Oliver, J. S., III & Graham, R. (2015). Teaching With Neotoma, SERC, https://serc.carleton.edu/neotoma/teaching_with_Neotoma.html.

Ooms, K. & Skarlatidou, A. (2018). Usability Engineering and Evaluation. *The Geographic Information Science & Technology Body of Knowledge* (1st Quarter 2018 Edition), John P. Wilson (ed).

Pezanowski, S., MacEachren, A. M., Savelyev, A., & Robinson, A. C. (2018). SensePlace3: a geovisual framework to analyze place–time–attribute information in social media. *Cartography and Geographic Information Science*, *45*(5), 420-437.

Quinn, S. D., & MacEachren, A. M. (2018). A geovisual analytics exploration of the OpenStreetMap crowd. *Cartography and Geographic Information Science, 45*(2), 140-155.

Robinson, A. C., Anderson, C. L., & Quinn, S. D. (2020). Evaluating geovisualization for spatial learning analytics. *International Journal of Cartography*, *6*(3), 331-349.

Rollinson, C.R., Dawson, A., Raiho, A.M., Williams, J.W., Dietze, M.C., Hickler, T., Jackson, S.T., McLachlan, J., JP Moore, D., Poulter, B. & Quaife, T. (2021). Forest responses to lastmillennium hydroclimate variability are governed by spatial variations in ecosystem sensitivity. *Ecology Letters, 24*(3), 498-508.

Rosson, M.B., & Carroll, J.M. (2002). Scenario-based Design. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, edited by J. Jacko and A. Sears, 1032–1050. New York: Lawrence Erlbaum Associates.

Sacha, D., Senaratne, H., Kwon, B. C., Ellis, G., & Keim, D. A. (2015). The role of uncertainty, awareness, and trust in visual analytics. *IEEE transactions on visualization and computer graphics, 22*(1), 240-249.

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics, 18*(12), 2431-2440.

Shneiderman, B. (2003). The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization* (pp. 364-371). Morgan Kaufmann. Chicago

Stegner, M. A., & Spanbauer, T. L. (2023). North American pollen records provide evidence for macroscale ecological changes in the Anthropocene. *Proceedings of the National Academy of Sciences, 120*(43), e2306815120.

Veloz, S. D., Williams, J. W., Blois, J. L., He, F., Otto-Bliesner, B., & Liu, Z. (2012). No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models. *Global Change Biology, 18*(5), 1698-1713.

Wang, Y., Pineda-Munoz, S., & McGuire, J. L. (2023). Plants maintain climate fidelity in the face of dynamic climate change. *Proceedings of the National Academy of Sciences, 120*(7), e2201946119.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data, 3*(1), 1-9.

Williams, J. W., Grimm, E. C., Blois, J. L., Charles, D. F., Davis, E. B., Goring, S. J., Graham, R.W., Smith, A.J., Anerson, M., Arroy-Cabrales, J., & Ashworth, A.C. (2018). The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. *Quaternary Research*, 89(1), 156-177.

Willighagen, L. G. (2019). Citation. js: a format-independent, modular bibliography tool for the browser and command line. *PeerJ Computer Science, 5*, e214.

Zhu, X., & Freeman, M. A. (2019). An evaluation of US municipal open data portals: A user interaction framework. *Journal of the Association for Information Science and Technology, 70*(1), 27-37.

Supplemental Material

NeotomaDB: Usability & Utility Assessment [survey]

This anonymous online survey aims to evaluate a community-curated, open source paleodata platform, NeotomaDB. The Neotoma Paleoecology Database serves as an online hub for data, research, education, and discussion about paleoenvironments. This research is being conducted by Dr. Jonathan Nelson and Dr. Jack Williams from UW-Madison's Geography Department to assess the usability and utility of various applications and services that comprise the NeotomaDB platform.

This survey should take approximately 15 minutes to complete.

Participation in this study is voluntary and you may end your participation at any time without penalty. You may ask any questions about the research at any time. If you have questions, concerns, or complaints, or think that participating in the research has hurt you, talk to the research team or contact the Principal Investigator, Jonathan Nelson at jknelson3@wisc.edu.

If you have any questions about your rights as a research participant or have complaints about the research study or study team, call the confidential research compliance line at 1-833-652-2506. Staff will work with you to address concerns about research participation and assist in resolving problems.

If you agree to participate, please select the "I agree" option below.
☐ I agree to participate in this anonymous survey.
☐ I do not agree to participate in this anonymous survey.

You have elected to participate in a research study focused on assessing NeotomaDB, a community-curated open paleoecological data platform.

The purpose of this research is to evaluate the usability and utility of Neotoma's core *data access & delivery* components. Specific components that will be evaluated include: the Neotoma Explorer user interface for exploring, finding, downloading, and visualizing data (https://apps.neotomadb.org/explorer/); a dedicated R statistical computing package (https://github.com/NeotomaDB/neotoma2); and an application programming interface (https://api.neotomadb.org/api-docs/) that supports Explorer, the R package, and other custom web applications. The primary intent of these applications and services is to enable research, education, and discussion about paleoenvironments through improving data access.

To sustain this effort and continue to improve its utility, we want to better understand the extent to which you leverage these tools and how they support your research and teaching efforts. This anonymous survey consists of four parts: 1) a brief background section; 2) a section focused on the Explorer application; 3) a section devoted to the R package; and 4) a section centered on the application programming interface (API). This survey will take approximately 15 minutes to complete. We appreciate your time!

Part 1: Background Information

Please answer the following background questions to the best of your ability. *Your answers will not affect your eligibility for this study.*

1. What is your age category?

   ☐ 18-25
   ☐ 25-35
   ☐ 35-45
   ☐ 45-55
   ☐ older than 55
   ☐ prefer not to answer

2. What is your gender identity?

   ☐ woman
   ☐ man
   ☐ transgender
   ☐ non-binary/non-conforming
   ☐ prefer not to answer

3. Which of the following best describe(s) your occupation (select all that apply)?

   ☐ student
   ☐ researcher
   ☐ educator
   ☐ data scientist
   ☐ application developer

☐ other (please specify): ______________________

4. Please rank your agreement with the following statements.

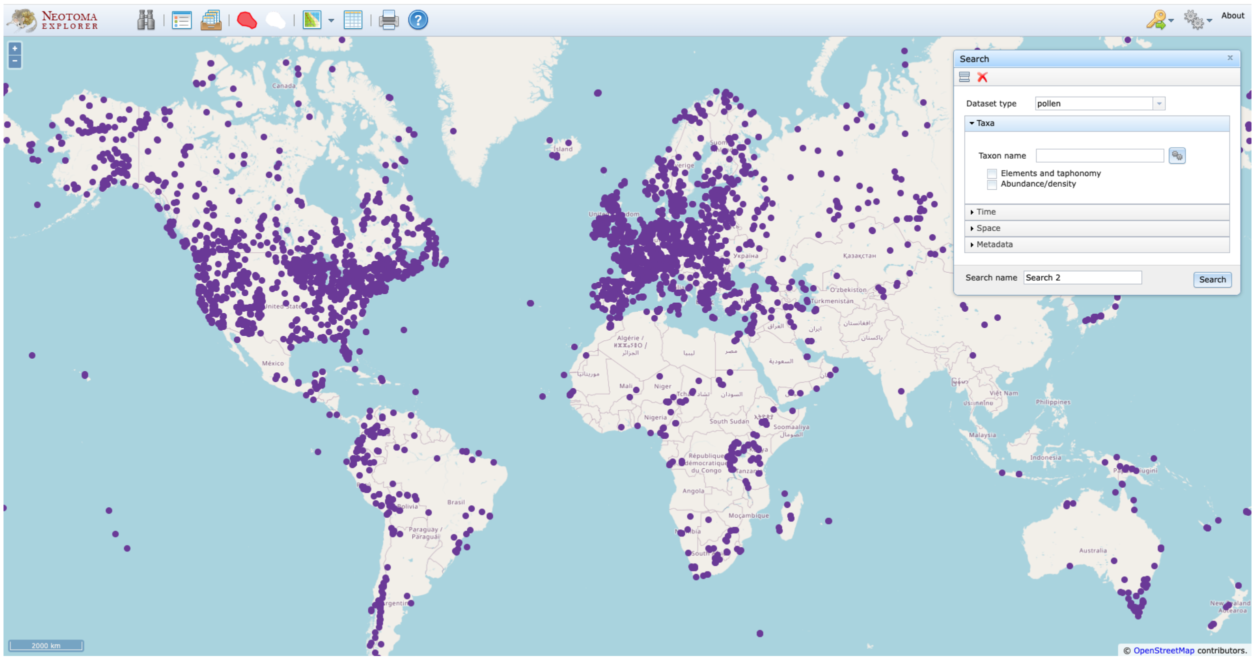
I consider myself to be an expert in these topic areas:

	strongly disagree	somewhat disagree	neither agree nor disagree	somewhat agree	strongly agree
paleoscience/ paleoecology	☐	☐	☐	☐	☐
biogeography	☐	☐	☐	☐	☐
earth system science	☐	☐	☐	☐	☐
climate science	☐	☐	☐	☐	☐
data science & statistical modeling	☐	☐	☐	☐	☐
interactive maps & data visualization	☐	☐	☐	☐	☐

5. Please provide three to five keywords that you think best describe your professional interests and/or research:
____
____

6. How many years have you been part of the NeotomaDB user community?
☐ this was my first year
☐ 2-3 years
☐ 3-5 years
☐ more than 5 years

Part 2: Neotoma Explorer



[Neotoma Explorer](#) provides an interactive starting place for interested users to quickly discover, visualize, and download data stored in NeotomaDB. The following questions are designed to help us understand how you interact with the web-based data explorer and to assess the extent to which the tool is both useful and usable.

1. Do you use Neotoma Explorer?
   ☐ yes
   ☐ no *(if no, participant is sent to part 3 of the survey)*

2. How often do you use Neotoma Explorer?
   ☐ daily
   ☐ weekly
   ☐ monthly
   ☐ a few times throughout the year

3. Neotoma Explorer allows users to search for data in a variety of different ways. Which of the following search strategies do you typically employ? (*select all that apply*)
   ☐ search by dataset type
   ☐ search by taxa
   ☐ search by time

☐ search by space
☐ search by metadata

4. Neotoma Explorer allows users to view and interact with data in a variety of different ways once a search result is returned. Which data discovery strategies do you typically employ? (*select all that apply*)
☐ view dataset information in the “Metadata” tab
☐ view samples information in the “Samples” tab
☐ view chronology information in the “Chronology” tab
☐ view stratigraphic information in the “Diagram” tab
☐ download data for use in another application
☐ link to data DOI page for additional information on dataset

5. Please rate your agreement with the following statements pertaining to the usability of Neotoma Explorer:

	strongly disagree	**somewhat disagree**	**neither agree nor disagree**	**somewhat agree**	**strongly agree**
Neotoma Explorer is unnecessarily complex.	☐	☐	☐	☐	☐
I would like to use Neotoma Explorer frequently.	☐	☐	☐	☐	☐
Neotoma Explorer is easy to use.	☐	☐	☐	☐	☐
I needed to learn a lot of things before I could get going with Neotoma Explorer.	☐	☐	☐	☐	☐

The various functions of Neotoma Explorer are well integrated.	☐	☐	☐	☐	☐
The interface of Neotoma Explorer is consistent.	☐	☐	☐	☐	☐
Most people could learn to use Neotoma Explorer quickly.	☐	☐	☐	☐	☐
Neotoma Explorer is cumbersome to use.	☐	☐	☐	☐	☐
I feel confident using Neotoma Explorer.	☐	☐	☐	☐	☐
I could use Neotoma Explorer without technical support from another person.	☐	☐	☐	☐	☐

6. If you could make one improvement to the usability of Neotoma Explorer (*in terms of how easy or difficult it is to navigate and understand the interface*), what would that improvement be?

   ____
   ____
   ____
   ____
   ____
   ____

7. Please rate your agreement with the following statements pertaining to the utility of Neotoma Explorer:

	strongly disagree	**somewhat disagree**	**neither agree nor disagree**	**somewhat agree**	**strongly agree**

Neotoma Explorer improves access to paleoenvironmental data.	☐	☐	☐	☐	☐
Neotoma Explorer enables better understanding of environmental change across spatial and temporal scales.	☐	☐	☐	☐	☐
Neotoma Explorer facilitates interdisciplinary paleoscience research.	☐	☐	☐	☐	☐
Neotoma Explorer serves as an educational resource for paleoscience coursework.	☐	☐	☐	☐	☐
Neotoma Explorer supports a broad community of researchers and educators.	☐	☐	☐	☐	☐

8. If you could make one improvement to the utility of Neotoma Explorer (*in terms of its ability to support paleoenvironmental research and education initiatives*), what would that improvement be?
____
____
____
____

Part 3: Neotoma2 R Package

````
---
title: "Working with diatom data in the `neotoma2` R package"
output:
  rmarkdown::html_vignette: default
  html_vignette:
    toc: yes
    toc_depth: 4
vignette: |
%\VignetteIndexEntry{Working with diatom data in the `neotoma2` R package} %\VignetteEngine{knitr::rmarkdown} %\VignetteEncoding{UTF-8}
---

```{r, include = FALSE}
knitr::opts_chunk$set(
  collapse = TRUE,
  comment = "#>",
  eval = FALSE
)
```

```{r setup}
library(neotoma2)
library(dplyr)
library(purrr)
library(sf)
```
````

The Neotoma2 R package serves as a starting point for a fully interactive experience with NeotomaDB through R. The following questions are designed to help us understand how you interact with the Neotoma2 R package.

1. Do you use the Neotoma2 R package?
   ☐ yes
   ☐ no *(if no, participant is sent to part 4 of the survey)*

2. How often do you use the Neotoma2 R package?
   ☐ daily
   ☐ weekly
   ☐ monthly
   ☐ a few times throughout the year

3. Do you reference the Neotoma2 R vignettes?
   ☐ no
   ☐ yes; if yes, which one(s): ____________________

4. The Neotoma2 R package allows users to query for data in a variety of different ways. Which of the following query strategies do you typically employ? (*select all that apply*)
☐ get_sites query
☐ get_datasets query
☐ get_downloads query
☐ other: please specify ___________________

5. The Neotoma2 R package allows users to plot and analyze data in a variety of different ways once data has been retrieved. In a short paragraph, please describe a typical workflow you employ to analyze Neotoma data using the R package.
____
____
____
____
____
____

6. Please rate your agreement with the following statements pertaining to the usability of the Neotoma2 R package:

	strongly disagree	**somewhat disagree**	**neither agree nor disagree**	**somewhat agree**	**strongly agree**
The Neotoma2 R package is unnecessarily complex.	☐	☐	☐	☐	☐
I would like to use the Neotoma2 R package frequently.	☐	☐	☐	☐	☐
The Neotoma2 R package is easy to use.	☐	☐	☐	☐	☐

I needed to learn a lot of things before I could get going with the Neotoma2 R package.	☐	☐	☐	☐	☐
The various functions of the Neotoma2 R package are well integrated.	☐	☐	☐	☐	☐
The syntax of the functions in the Neotoma2 R package is consistent.	☐	☐	☐	☐	☐
Most people could learn to use the Neotoma2 R package quickly.	☐	☐	☐	☐	☐
The Neotoma2 R package is cumbersome to use.	☐	☐	☐	☐	☐
I feel confident using the Neotoma2 R package.	☐	☐	☐	☐	☐
I could use the Neotoma2 R package without technical support from another person.	☐	☐	☐	☐	☐

7. If you could make one improvement to the usability of the Neotoma2 R package (*in terms of how easy or difficult it is to use the R functions*), what would that improvement be?

   ____
   ____
   ____
   ____

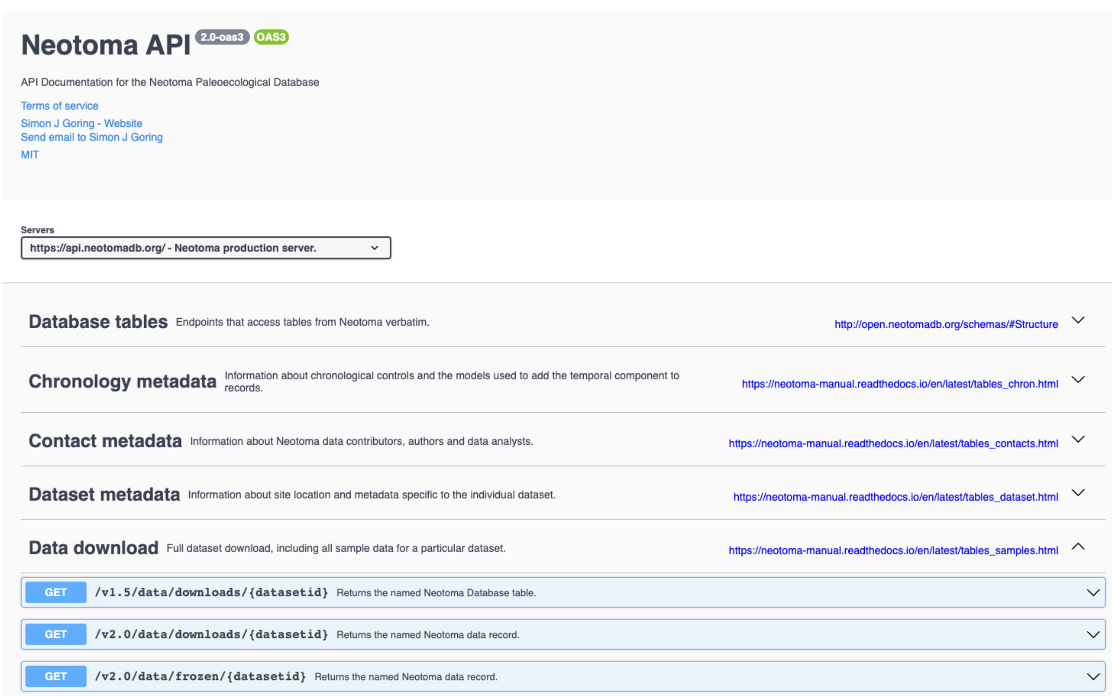
8. Please rate your agreement with the following statements pertaining to the utility of the Neotoma2 R package:

	strongly disagree	**somewhat disagree**	**neither agree nor disagree**	**somewhat agree**	**strongly agree**
The Neotoma2 R package improves access to paleoenvironmental data.	☐	☐	☐	☐	☐
The Neotoma2 R package enables better understanding of environmental change across spatial and temporal scales.	☐	☐	☐	☐	☐
The Neotoma2 R package facilitates interdisciplinary paleoscience research.	☐	☐	☐	☐	☐
The Neotoma2 R package serves as an educational resource for paleoscience coursework.	☐	☐	☐	☐	☐

The Neotoma2 R package supports a broad community of researchers and educators.	☐	☐	☐	☐	☐

9. If you could make one improvement to the utility of the Neotoma2 R package (*in terms of its ability to support paleoenvironmental research and education initiatives*), what would that improvement be?

Part 4: Neotoma API



Data in NeotomaDB are also made openly available via the Neotoma API. The Neotoma API supports internal tools, such as Explorer and the R package, as well as third party applications, such as Flyover Country, Range Mapper, and Where the Wild Things Were. The following questions are designed to help us understand how you leverage the Neotoma API for custom analysis and application development.

1. Do you use the Neotoma API?
   ☐ yes
   ☐ no *(if no, survey ends and participant is thanked for their time)*

2. How often do you use the Neotoma API?
   ☐ daily
   ☐ weekly
   ☐ monthly
   ☐ a few times throughout the year

3. The Neotoma API allows users to GET data in a variety of different formats. Which API endpoints do you most frequently use? (*please specify URLs*)
   ____
   ____
   ____
   ____
   ____
   ____

4. The Neotoma API supports custom analysis and application development. In a short paragraph, please describe an analysis and/or application you have developed (or are actively developing) using the Neotoma API.
   ____
   ____
   ____
   ____
   ____
   ____

5. Please rate your agreement with the following statements pertaining to the usability of the Neotoma API:

	strongly disagree	somewhat disagree	neither agree nor disagree	somewhat agree	strongly agree
The Neotoma API is unnecessarily complex.	☐	☐	☐	☐	☐
I would like to use the Neotoma API frequently.	☐	☐	☐	☐	☐
The Neotoma API is easy to use.	☐	☐	☐	☐	☐
I needed to learn a lot of things before I could get going with the Neotoma API.	☐	☐	☐	☐	☐
The various functions of the Neotoma API are well integrated.	☐	☐	☐	☐	☐
The syntax of the endpoints in the Neotoma API is consistent.	☐	☐	☐	☐	☐
Most people could learn to use the Neotoma API quickly.	☐	☐	☐	☐	☐
The Neotoma API is cumbersome to use.	☐	☐	☐	☐	☐
I feel confident using the Neotoma API.	☐	☐	☐	☐	☐
I could use the Neotoma API without technical support from another person.	☐	☐	☐	☐	☐

6. If you could make one improvement to the usability of the Neotoma API (*in terms of how easy or difficult it is to use the endpoints*), what would that improvement be?
____
____
____

7. Please rate your agreement with the following statements pertaining to the utility of the Neotoma API:

	strongly disagree	**somewhat disagree**	**neither agree nor disagree**	**somewhat agree**	**strongly agree**
The Neotoma API improves access to paleoenvironmental data.	☐	☐	☐	☐	☐
The Neotoma API enables better understanding of environmental change across spatial and temporal scales.	☐	☐	☐	☐	☐
The Neotoma API facilitates interdisciplinary paleoscience research.	☐	☐	☐	☐	☐

The Neotoma API serves as an educational resource for paleoscience coursework.	☐	☐	☐	☐	☐
The Neotoma API supports a broad community of researchers and educators.	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐

8. If you could make one improvement to the utility of the Neotoma API (*in terms of its ability to support paleoenvironmental research and education initiatives*), what would that improvement be?
   ____
   ____
   ____
   ____