The Neotoma Paleoecology Database: A research-outreach nexus

Goring, Simon James[1]

Graham, Russell[2]

Loeffler, Shane[3]

Myrbo, Amy[3]

Ormand, Carol J[4]

Oliver, James S, III[2]

Williams, John W[1,5]

[1] Department of Geography, University of Wisconsin-Madison, Madison WI; [2] Department of Geosciences, Pennsylvania State University, State College, PA; [3] Department of Earth Sciences, University of Minnesota, Minneapolis, MN; [4] Science Education Resource Center, Carleton College, Northfield, MN; [5] Center for Climatic Research, University of Wisconsin-Madison WI

Corresponding Author: goring@wisc.edu

Neotoma: A teaching nexus.

Introduction

The Quaternary period (2.6 million years ago to present) is a highly active area of integrative ecological, geological, and climatic research and is well suited to a number of educational outreach opportunities. The Quaternary is a natural entry point for students to think about processes operating across a wide range of timescales, because 1) the Quaternary overlaps with the evolution, migration, and development of humans and societies and 2) the key processes operate at timescales of decades to hundreds of thousands of years. Thus, the Quaternary includes our own human history, while its timescales are a bridge between those of our lives and historical memory (decades-centuries) and deep-time geological timescales (millions of years and longer).

The Quaternary is of particular interest to scientists studying climate change, the ecological effects of climate change, and the causes and effects of species range shifts and extinctions, all of which can also be engaging questions for students. Climatically, the defining features of the Quaternary include: 1) the oscillation between long glacial periods and short interglacials, accompanied by the growth and collapse of continental ice sheets, sea-level falls and rises, and variations in global temperature and atmospheric greenhouse gas concentrations; 2) a pacing of these glacial-interglacial changes by variations in the Earth's orbit, known as Milankovitch cycles; and 3) large and abrupt temperature changes during glacial terminations, Heinrich Events, Dansgaard-Oeschger events, and other millennial-scale climate events (these events are called millennial-scale because they are spaced thousands of years apart). During the Holocene interglacial (11,700 years ago to present), temperature variations have been muted but hydrological variability has intensified, linked to both internal processes such as shifts in ENSO variability and external forcings such as individual volcanic eruptions and variations in solar luminosity.

Ecologically, species distributions repeatedly retracted and expanded during the glacial-interglacial cycles, with high local community turnover during climatic events. The worldwide expansion of humans before and during the last deglaciation was accompanied by a global wave of extinctions of large animals (the Pleistocene megafauna) and other hominid species (e.g. Neanderthals, Denisovians). During the Holocene, post-glacial range expansions of tree species continued, with population dynamics and range expansions often paced by hydrological variability and drought. Understanding these past ecological dynamics helps global change ecologists and biogeographers understand how species are likely to respond to current environmental changes and test the predictive ability of Earth system models.

Our understanding of climatic and ecological dynamics during the Quaternary is founded upon many individual site-level analyses of fossil organisms and

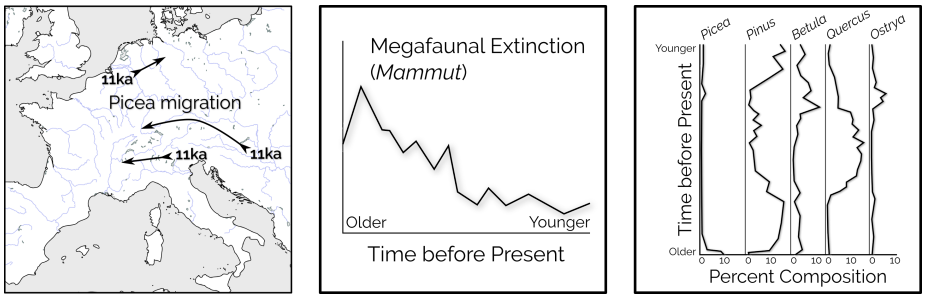

Figure 1: Examples of Quaternary datasets: (1) Species migration through time beginning 11,000 years before present (11ka), here, Picea (spruce) in Europe after the end of the last ice age (from Latalowa and van der Knaap, 2006); (2) Megafaunal extinction interpreted from the decline in the fungal spore Sporormiella, present in the dung of large herbivores (after Gill et al. 2009) and (3) change in mammal body size in the ground squirrel *Spermophilus beecheyi* (after J. L. Blois, Feranec, and Hadly 2008).

paleoenvironmental proxies (Figure 1). Fossil data are many and varied and include micropaleontological data (e.g. pollen, diatoms, ostracodes, foraminifera), macropaleontological data (e.g. vertebrate fossils, plant macrofossils), and, increasingly, molecular and organic geochemical tracers such as ancient DNA or organic compounds such as the alkanes from leaf waxes. These records are collected from a diverse array of depositional environments, including excavations, lake sediment core sampling, or other fossil localities. Understanding large-scale spatial phenomena such as the climate-driven range shifts of species and Pleistocene megafaunal extinction and the role of humans, climate change, and requires gathering these many diverse datasets into larger databases that store information about age, taxonomy, spatial coordinates, depth or stratigraphy, and other related information.

In response to these needs, the Neotoma Paleoecological Database (Neotoma) has emerged as a Community-Curated Data Resource (CCDR) with the mission of gathering, curating, and sharing paleoecological data, to enable open, global-scale science (J. W. Williams et al. in press). This resource also provides students and educators with a public portal for accessing data that can be used to understand ecological change through the Quaternary. Authentic inquiry into real, complex data sets engages students in the learning process, thus deepening their understanding of paleoecology and simultaneously increasing their retention of key concepts (e.g. Laursen et al. 2010; Lopatto 2010; Freeman et al. 2014). Multiple avenues have been developed for finding, viewing, exploring, and obtaining Neotoma data. Several recent educational and outreach activities have resulted in the production of a number of new resources for a wide range of audiences, including high school and college students, graduate students, and the general public.

This chapter introduces Neotoma as a tool for teaching and outreach and its multiple systems for data access and visualization, which together support a wide range of users and levels of expertise. We begin with an introductory exploration of Neotoma's data holdings, using the map-based Neotoma Explorer and a variant of exercises that were developed with the Science Education Resource Center at Carleton College (https://serc.carleton.edu/index.html). We then provide a summary of teaching exercises for upper level undergraduate and graduate students using the software package R (S. Goring et al. 2015) that have been developed through a series of workshops. Next, we show how the Flyover Country® mobile app (http://fc.mnu.edu) can be used to discover and learn about Earth's history, drawing on and integrating resources at Neotoma and elsewhere. We conclude this chapter by examining the role of community curated data resources as bridges between the research community and the public, with regard to educational outreach in particular.

An Introduction to Neotoma

Neotoma is both a repository and a data access portal. Neotoma contains more than 17500 unique datasets, containing observations of fossil organisms and associated geophysical information spanning the last 2.6 million years. These records and their associated metadata are contained within a web-accessible database that has been designed for research and outreach activities (J. W. Williams et al. in press).

Neotoma relies upon a software ecosystem with several components. At its heart is a flexible data model designed to to describe paleoecological data records for many types of sites and measurements [J. W. Williams et al. (in press);grimm2008neotoma]. It includes information about researchers, locations, publications, time and the raw paleoecological and paleoenvironmental records that underpin our understanding of Quaternary environments.

Neotoma Explorer (https://apps.neotomadb.org/explorer/) is the primary data discovery interface for finding and exploring Neotoma data (Figure 2), powered by an application program interface (API; http://api.neotomadb.org) that can be directly accessed through the world-wide web. The use of an API provides a mechanism to build new tools on top of Neotoma's data resources. The API has also allowed the development of the neotoma R package (S. Goring et al. 2015) and the inclusion of Neotoma data within the Flyover Country® mobile app (http://flyovercountry.io).

The following introductory exercise is based on Neotoma Explorer and is designed to help individuals understand how paleoecologists interpret and process proxy data, from a single site, to continental scales, from the near-modern to deep time records. It teaches users how to search for individual paleoecological sites, find publications, and visualize change through time. It also introduces the kinds of research questions, for example, species co-occurrence, that paleoecologists are

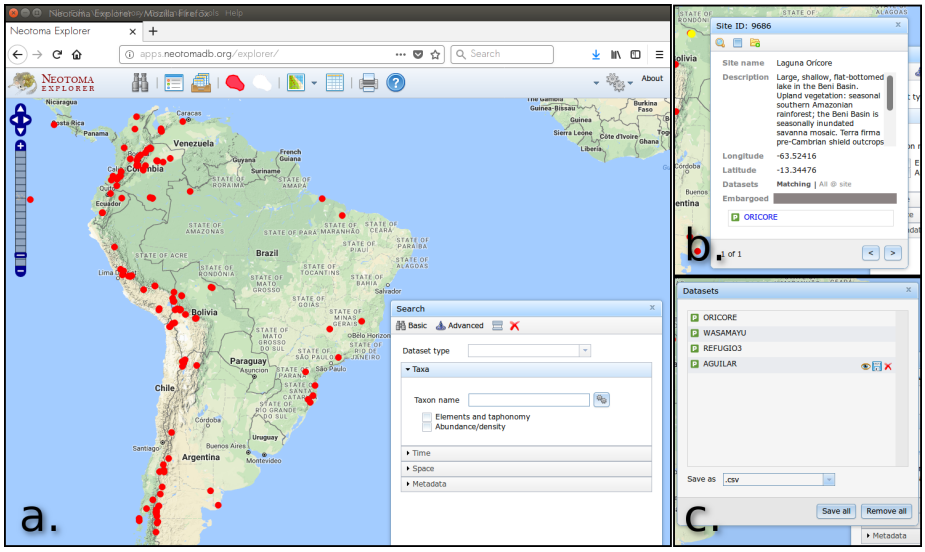

Figure 2: *The Neotoma Explorer is a web based data discovery tool, that allows a map based search for paleoecological sites using species names, site locations, site names, or other associated metadata (a). The tool provides the opportunity to examine individual sites (b) and to select a number of sites for download and later processing or study (c).*

interested in. This exercise is based on one originally developed at a workshop of researchers and teachers hosted by SERC. It is included here as an example of how the Neotoma database can engage students in learning the principles of paleoecology through authentic exploration of scientific data sets. While it is primarily intended to be an introduction to using the database, it provides students with a foundation of skills they can use for further exploration and inquiry in subsequent exercises.

Your First Search

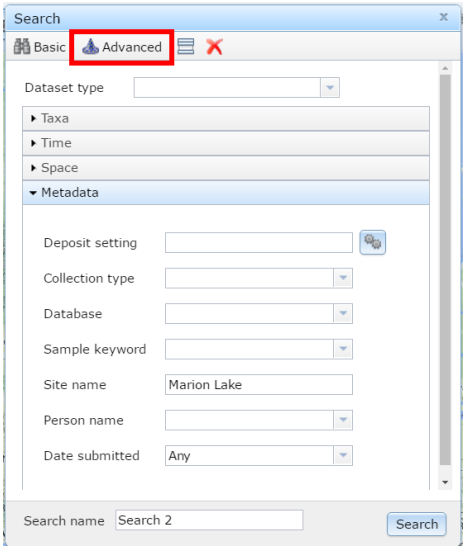

Figure 3: *Searching for Marion Lake using the Neotoma Explorer's search tools (https://apps.neotomadb.org/explorer/).*

The Neotoma Explorer provides the spatial overview and search capabilities for Neotoma. One of the simplest things you can do with the database is to search for a single research site. To do this, look for the "*Metadata*" search option under the "*Advanced*" tab of the search panel (Figure 3). "Marion Lake" is an important paleoecological record from western North America, the record provided one of the first quantitative reconstructions of Holocene climate from western North America (Mathewes and Heusser 1981). *Search for Marion Lake.* You may find more than one site. Which one is the pollen dataset? How do you know?

The northern-most site is the Marion Lake pollen sample site. When you click on the point you should see some site information and a description, along with a green **P**, as well as a small clock icon. Both have the word `MARION` beside them.

The icons represent pollen data and geochronological data respectively. *Click on the pollen data* first and scroll through the tabs available to you. You'll see:

- **Samples** – Individual pollen counts, depth information and summary chronology.
- **Diagram** – A tool to draw the change in the data values over time for datasets where diagrams are possible (pollen, diatoms, *etc.*)
- **Site** – Site level information for the data.
- **Chronology** – The age model used for the record.
- **Publications** – Publications related to the dataset.

These tabs provide you with a pretty good overview of the dataset, long term changes at the site, and any papers you might want to read to get more information about the record.

More Searching

Neotoma can also be used to look for a large number of datasets with a particular dataset type. *Create a new search for a* ***Dataset Type*** *(make sure you clear the site name).* What do you see? Select "water chemistry" datasets and choose a site. How does that dataset look different from the pollen dataset we looked at earlier? Why?

More Details

Now *click on the white polygon on the top bar.* If you enter a date (say `21100`), hit enter, and then zoom out. At its full extent, you'll see a white polygon, representing the continental ice sheets, overlain on the map of sites. It is possible to use the ice sheet distributions to help understand paleo-distributions of taxa, their presence or absence.

Research Questions

Mastodon is a progressive heavy metal band, thrice nominated for a Grammy award, but it was also a species of extinct mammutid proboscideans.

We tend to picture late-glacial environments as being cold, dry, and empty of trees, similar to modern tundra environments. *Do a taxon search for* Mammut *(mastodons) and look at the site locations that appear.* What happened? Where did you see site locations?

Figure 4 shows the button needed to select multiple related taxa through the "Basic Search" window. More records can be found by doing a multi-taxon search. *Click the gear box to search for multiple taxa.* Search and select all the *Mammut*s we can find in the list. Make sure to choose the right *Taxa Group* and then search. Where are the mastodons?

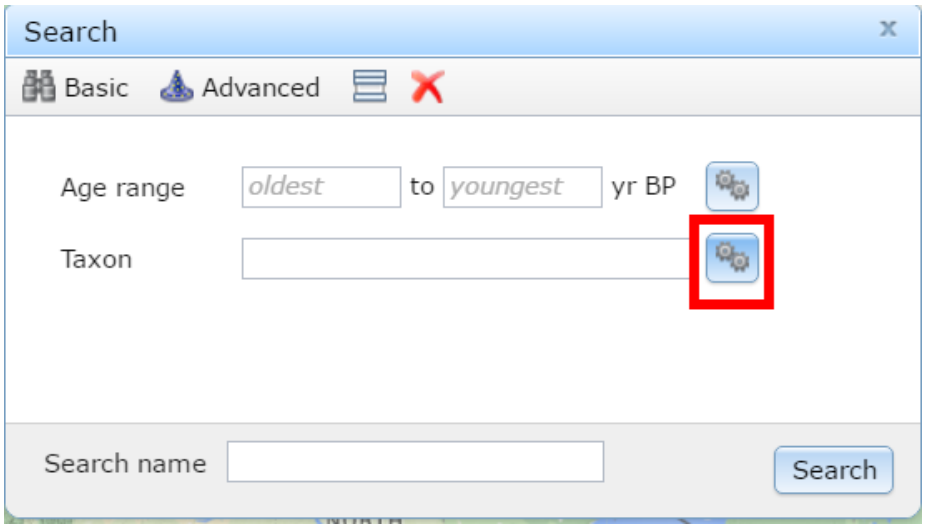

Figure 4: The search tools provide the opportunity to search for a single taxon or for multiple taxa simultaneously.

We can also look for sites with common forest taxa. Let's look for sites with *Picea* in them. Since *Mammut* went extinct somewhere around 12,000 ybp (Widga et al. 2017), lets limit our search to all sites with *Picea* taxon, abundance over 20% and a time range of 15,000 – 12,000 ybp. When we look at the distribution of sites with *Picea* and sites with records of mastadons, what can we say about our friends the mastodons? Does it appear that they may have co-occurred? Were their distributions completely separate? What does this tell us about the kinds of environments mastadons may have occupied.

More Exploration with Explorer: Additional SERC Teaching Exercises

The above exercise is based on "Exploring the Neotoma Paleoecology Database" (https://serc.carleton.edu/neotoma/activities/121251.html), an exercise directed toward college undergraduates that covers mammal and vegetation distribution changes since the Last Glacial Maximum. This exercise and others were developed in 2015 at a workshop hosted by SERC. The workshop provided an opportunity for paleoecology domain experts to work with science education researchers to develop high-quality teaching materials at a variety of educational levels.

Additional exercises from this workshop, focused on undergraduate, or senior high school students, are available on the SERC website (https://serc.carleton.edu/neotoma/activities.html). Other examples include: (1) "Species distributions in response to environmental gradients in the

Upper Midwest of the United States - an example using the Neotoma database", a unit focused on environmental gradients over short spatial scales and how these change over time; and (2) "Climate Change and Mammal Dispersal" which is directed toward senior high school and early undergraduate students, helping them understand dispersal in small mammals through time in response to climate change.

Each of these exercises guides students through the process of accessing and exploring data from the Neotoma database to answer a scientific question or questions. This kind of scaffolding, generally referred to as "guided inquiry," allows students to discover their own answers while preventing the frustration they might experience if simply turned loose the answer the question without guidance (Hmelo-Silver, Duncan, and Chinn 2007).

Educational Modules for Climate Change, Paleoecology, and Biogeography

A series of eleven modules have been developed for participants to learn about climate change, paleoecology, and biogeography using the Neotoma Paleoecology Database. These modules are primarily focused on mammals but can easily be adapted to other organisms like coleoptera (beetles), plants, ostracodes, and other dataset types supported by Neotoma. There are two basic module types: (1) Modules that provide background information about each of climate change, paleoecology, and biogeography (background modules), and (2) modules in which learners apply knowledge from the first set of modules to examine hypotheses using data taken from the Neotoma database (applied modules).

Background modules can be used by learners who may not be familiar with the basic concepts of topics such as climate change, paleoecology, and biogeography. For instance, in the Paleoecology modules, relationships between climate (moisture and temperature) are discussed and illustrated (*e.g.* the latitudinal correspondence between temperature and the longitudinal correlation with temperature in the eastern United States). Relationships between the modern distributions of mammals within the Neotoma database are compared to climate parameters. From this, it becomes apparent that the eastern and western limits of the species' distributions in the eastern United States are functions of moisture; whereas, the northern and southern ranges are temperature-dependent. The paleoecology module also provides a brief discussion about underlying assumptions linking climate variables and species' distributions and how these assumptions may sometimes be violated. For instance, the distribution of the prairie dog (*Cynomys ludovicianus*) may be controlled by climate as well as parameters such as soil depth, particle size, and soil moisture. A series of questions reinforce the learning objectives at the end of each module. The modules also provide links to other web pages and literature that will provide further explanation of the principles.

In the second modules, learners take knowledge they have gained from the background modules and apply it to analyses of data taken from Neotoma. Each module begins with a statement about the process to be learned and an example of similar analysis is provided. Learners develop a hypotheses (*e.g.*, "If the climate warms, then the species' distribution will expand westward"). The participant may then test the hypothesis, by examining changes in the distribution of the species during different times of climate change using the Greenland Ice Core Project (GRIP: Andersen et al. 2004) climate data, covering the transition from full glacial to late glacial or late glacial to late Holocene. A number of mammal species are given to the student as they continue the exercise. If the hypothesis is false, then the learner is asked why it failed. The learner is then taken to a file that then explains what potentially happened with their analysis. In some cases, multiple species may be involved in the exercise. Finally, based upon these analyses, the student is asked about the potential response of these species to future global warming.

For biogeographic analyses, students are asked to examine changes in species' distributions through time. Various exercises then focus on when a species arrived in an area and how long it is maintained. With the addition of FAUNMAP II data that is currently, being uploaded, participants will be able to work with not only the late Quaternary, the last 20,000 years but will be able to go back through the Miocene to 30 million years ago. These exercises will involve rates of immigration and extinction for various mammal species.

Future work on all these modules will involve more exercises on conservation paleobiology and more sophisticated analyses of groups of species (paleocommunities).

Programmatic Access for Macro-Scale Paleoecological Research (APIs, neotoma R, and GitHub)

Advanced undergraduates and graduate students will likely be interested in learning about the programmatic tools for accessing Neotoma data: Neotoma Explorer is excellent for quick-look visualizations and introductory teaching, but Neotoma's programmatic interfaces are better suited for large-scale data downloads and analyses. Neotoma supports programmatic data access through an API that provides users with the ability to access data from Neotoma using programming languages such as JavaScript, R or Python. The neotoma R package (https://github.com/ropensci/neotoma; S. Goring et al. 2015) offers a convenient system for exporting Neotoma into the R environment, which is widely used by ecologists and paleontologists for statistical data analysis.

GitHub (http://github.com) is a particularly useful system for open-source software development, with its ability to host and version programmatic code that can be accessed and modified by multiple developers. GitHub can also host public webpages (using GitHub Pages; https://pages.github.com/), which provides the

opportunity to host educational material, along with the source code that powers any programmatic workflows. Our educational and training workshops for graduate students and early career scientists make frequent use of GitHub pages and much material can be found there (http://neotomadb.github.io/workbooks.html). These materials use the software principles of "version control" to borrow elements across workshops, while allowing workshop leaders to take the strengths of core workshop materials. This provides the ability to both reuse well established teaching materials and to tailor for new audiences. Workshop material for five separate workshops is currently available in the Neotoma Workshops GitHub repository, and new workshop materials are being actively developed. In the remainder of this section, we walk through a common graduate-level introductory exercises for accessing and analyzing Neotoma data.

Accessing Data with `neotoma`'s R package

More complete resources for accessing Neotoma data are provided in online materials, including online workshop materials e.g., https://neotomadb.github.io/workbooks/Workshop_SVP2016_v0.1 neotoma-package. The R package is also documented in Goring et al. (2015) using worked examples drawn from the literature to showcase the ways that the R package can be used to undertake high impact research.

Programmatic Access and Paleoecology

Neotoma supports data access through an Application Program Interface (API) that provides users with the ability to access data from Neotoma using programming languages such as JavaScript, R or Python. Since the publication of Goring *et al.* (S. Goring et al. 2015), a number of resources using the API have been published online in a publically accessible format. In particular, GitHub has the ability to host programmatic code, which acts as a resource, but also publically hosted pages (using GitHub Pages), which provides the opportunity to host educational material, along with the source code that powers any programmatic workflows (Figure 5).

Chronology Construction & CLimate

Chronology construction and pollen-based climate reconstruction are both complex tasks. Chronology construction is central to the field of the paleogeosciences (Harrison et al. 2016), while climate reconstruction is widely used, but fraught with methodological issues (Juggins 2013). Within the Quaternary geosciences there are a number of techniques that can be used to construct chronologies, and packages that can be used to reconstruct climate from proxy data. Tools such as Bacon (Blaauw and Christen 2011) and BChron (Parnell 2016) for chronologies, or `rioja` (Juggins 2017) and `analogue` (G. L. Simpson 2007) for climate reconstruction, have a number of settings and requirements, however, the

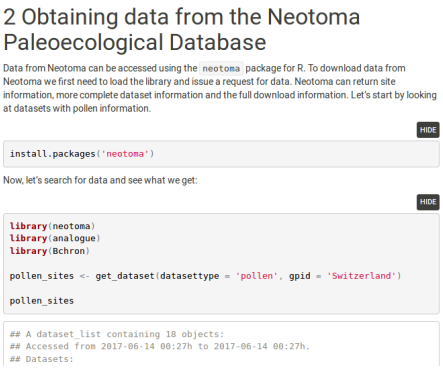

Figure 5: Educational material presented on Neotoma's GitHub Pages show the initial stages of constructing chronologies for paleoecological records.

paleoecological community is highly distributed, and many students might not encounter a course on chronology construction or paleoclimate reconstruction.

Flyover Country®: Beyond the Classroom

Technological developments within the CCDR community can lead to unexpected opportunities. One such is the use of Neotoma in the National Science Foundation (NSF) funded Flyover Country® mobile app for geoscience (http://flyovercountry.io), available for iOS and Android devices. The app is used by travelers (or armchair explorers) to see what geoscientific discoveries have been made in a particular region of interest worldwide. Flyover Country® uses the Neotoma API to find fossil occurrences of extinct "charismatic megafauna" (mastodons, saber-tooth tigers, giant beavers, etc.) within an area. Because of its user-friendly API and interesting content, Neotoma was one of the original set of databases to be included in Flyover Country®. App users click on map, which submits a spatial query directly to Neotoma. The API structure allows Flyover Country® to link the taxon name (e.g., *Mammut*) with the Wikipedia article about that animal, which is then also displayed in the app, enriching the learning experience of the user. Neotoma's well-documented API facilitates the use of Neotoma data in other mobile and web apps beyond Flyover Country®. Mobile and web app development is within the reach of undergraduate courses, and the ready availability of Neotoma data through the API means that those apps can be easily populated with interesting information. Neotoma's spatial, multivariate, and temporal data provide additional interesting challenges for data visualization in such courses.

Discussion: CCDRs at the Research-Outreach Nexus

The development of community supported data resources (CCDRs) is often a by-product of a community of researchers attempting to answer research questions that can only be answered through the use of a number of datasets with some standard data representation. Traditionally, CCDRs have been a resource for data storage: systems of record; collating, aggregating and cleaning data, largely supervised by domain experts (Kapoor et al. 2015). As mechanisms to access and visualize data have improved across the geosciences, it has become possible to generate opportunities for individuals and groups outside of the traditional research community to access and interact with this data, moving these databases from systems of record to systems of engagement, adding a social and collaborative element to the platform. As such, engagement with educational activities represents a major opportunity to help transform CCDRs, providing an opportunity for broader engagement, and for the development of a robust user community, adding value beyond just data content.

The Role for CCDRs in Outreach

The development of community supported data resources is often a by-product of a community of researchers attempting to answer research questions that can only be answered through the use of a number of datasets with some standard data representation. Traditionally, CCDRs have been a resource for data storage: systems of record; collating, aggregating and cleaning data, largely supervised by domain experts (Kapoor et al. 2015). As mechanisms to access and visualize data have improved across the geosciences, it has become possible to generate opportunities for individuals and groups outside of the traditional research community to access and interact with this data, moving these databases from systems of record to systems of *engagement*, adding a social and collaborative element to the platform. As such, engagement with educational activities represents a major opportunity to help transform CCDRs, providing an opportunity for broader engagement, and for the development of a robust user community, adding value beyond just data content.

By acting as a portal for outreach with the research community, CCDRs can serve multiple, overlapping communities, and can bring them together in collaborative endeavors (Figure 6). For example, the NSF Funded project "Neotoma: Community-led Cyberinfrastructure for Global Change Research" was able to leverage a broad group of researchers, and to connect with Earth Science Educators, through SERC. Through this collaboration, the project generated a collection of high-quality educational resources. The strength of these curricular materials is multi-faceted: the exercises make use of the extraordinarily rich data available through the Neotoma database, as well as its data visualization tools; they engage students in authentic data inquiry to answer questions of interest to scientists; and they illustrate the use of paleoecological data to understand

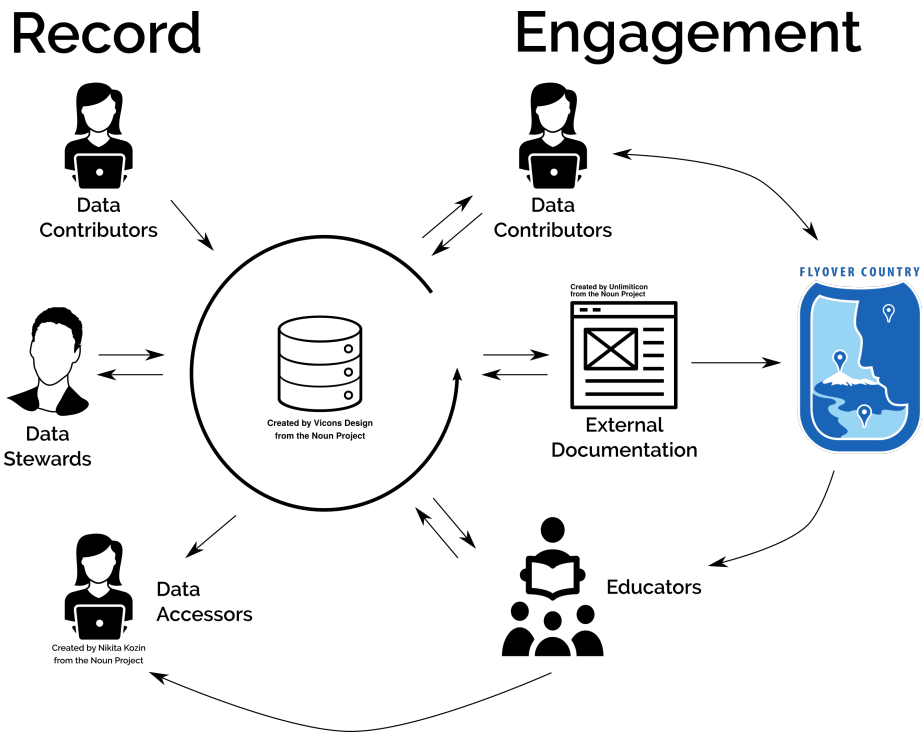

Figure 6: CCDRs involved in engagement can access new communities outside of the traditional sphere of activity in the research community.

the impacts of climate change processes happening today. Tapping the expertise of both communities - paleoecologists and educational researchers - produced more robust educational materials than either community would have produced alone. Thus, CCDRs provide an opportunity to link research communities to educators, and educators to the broad range of cutting-edge research activities, building on the strengths of each group.

The emerging field of paleoecoinformatics (Brewer, Jackson, and Williams 2012) has enhanced the paleoecology communities suite of tools, bringing geospatial tools, and the technical capacity to combine web or mobile mapping tools with data resources. The development of the Neotoma Explorer (http://apps.neotomadb.org/explorer) provides an easy interface for elementary and secondary school students, while access through the Neotoma API and the neotoma R package (S. Goring et al. 2015) allows the integration of paleoecological data into post-secondary curriculum. With the emergence of larger cyberinfrastructure programs, such as NSF's EarthCube program, there will be increased pressure to introduce technology as a core component of curriculum design within the geosciences. CCDRs act as important translational actors in this ecosystem, using core technological expertise, and applying it to scientific outreach programs with the assistance of education experts.

Here we have shown a range of educational opportunities available online for the Neotoma Paleoecological Database. Each of these can be modified for a range of data types, including plants, mammals and limnological datasets. Given this, a range of undergraduate students can access and integrate these modules into their learning programs. Sustained support for CCDRs thus provides long term opportunities for the maintenance of scientific and pedagogical infrastructure, and offers a key avenue for supporting distributed research communities, and the continued development of new opportunities for public outreach.

Acknowledgements & Authorship

Authorship is ordered by lead author (SJG) and subsequently by alphabetical order. SJG produced the main text and figures. Co-authors provided text and editing, and assisted in the workshop development and module creation. The authors would like to acknowledge both contributors to the Neotoma Paleoecological Database and data stewards. Work on this paper was supported by NSF Awards NSF-1541002, NSF-1550707 and NSF-1550707.

References

Andersen, Katrine K, N Azuma, J-M Barnola, Matthias Bigler, P Biscaye, N Caillon, J Chappellaz, et al. 2004. "High-Resolution Record of Northern

Hemisphere Climate Extending into the Last Interglacial Period." *Nature* 431 (7005). Nature Publishing Group: 147.

Blaauw, Maarten, and J Andrés Christen. 2011. "Flexible Paleoclimate Age-Depth Models Using an Autoregressive Gamma Process." *Bayesian Analysis* 6 (3). International Society for Bayesian Analysis: 457–74.

Blois, Jessica L, Robert S Feranec, and Elizabeth A Hadly. 2008. "Environmental Influences on Spatial and Temporal Patterns of Body-Size Variation in California Ground Squirrels (Spermophilus Beecheyi)." *Journal of Biogeography* 35 (4). Wiley Online Library: 602–13.

Brewer, Simon, Stephen T Jackson, and John W Williams. 2012. "Paleoecoinformatics: Applying Geohistorical Data to Ecological Questions." *Trends in Ecology & Evolution* 27 (2). Elsevier: 104–12.

Freeman, Scott, Sarah L Eddy, Miles McDonough, Michelle K Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. 2014. "Active Learning Increases Student Performance in Science, Engineering, and Mathematics." *Proceedings of the National Academy of Sciences* 111 (23). National Acad Sciences: 8410–5.

Gill, Jacquelyn L, John W Williams, Stephen T Jackson, Katherine B Lininger, and Guy S Robinson. 2009. "Pleistocene Megafaunal Collapse, Novel Plant Communities, and Enhanced Fire Regimes in North America." *Science* 326 (5956). American Association for the Advancement of Science: 1100–1103.

Goring, Simon, Andria Dawson, Gavin Simpson, Karthik Ram, Russ Graham, Eric Grimm, and John Williams. 2015. "Neotoma: A Programmatic Interface to the Neotoma Paleoecological Database." *Open Quaternary* 1 (1). Ubiquity Press.

Harrison, Mark, Suzanne Baldwin, Marc Caffee, George Gehrels, Blair Schoene, David Shuster, and Brad Singer. 2016. "Geochronology: It's About Time." *Eos* 97 (5). American Geophysical Union: 12–13.

Hmelo-Silver, Cindy E, Ravit Golan Duncan, and Clark A Chinn. 2007. "Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006)." *Educational Psychologist* 42 (2). Taylor & Francis: 99–107.

Juggins, Steve. 2013. "Quantitative Reconstructions in Palaeolimnology: New Paradigm or Sick Science?" *Quaternary Science Reviews* 64. Elsevier: 20–32.

———. 2017. *Rioja: Analysis of Quaternary Science Data*. http://www.staff.ncl.ac.uk/stephen.juggins/.

Kapoor, Shubir, Aleksandra Mojsilovic, Jade Nguyen Strattner, and Kush R Varshney. 2015. "From Open Data Ecosystems to Systems of Innovation: A Journey to Realize the Promise of Open Data." In *Bloomberg*

Data for Good Exchange Conference. https://pdfs.semanticscholar.org/ebfa/cdfc9da14c5b54791e6fba89bd7a6c7809d0.pdf.

Latałowa, Małgorzata, and Willem Oscar van der Knaap. 2006. "Late Quaternary Expansion of Norway Spruce Picea Abies (L.) Karst. in Europe According to Pollen Data." *Quaternary Science Reviews* 25 (21). Elsevier: 2780–2805.

Laursen, Sandra, Anne-Barrie Hunter, Elaine Seymour, Heather Thiry, and Ginger Melton. 2010. *Undergraduate Research in the Sciences: Engaging Students in Real Science*. John Wiley & Sons.

Lopatto, David. 2010. "Undergraduate Research as a High-Impact Student Experience." *Peer Review* 12 (2). Association of American Colleges; Universities: 27.

Mathewes, Rolf W, and Linda E Heusser. 1981. "A 12 000 Year Palynological Record of Temperature and Precipitation Trends in Southwestern British Columbia." *Canadian Journal of Botany* 59 (5). NRC Research Press: 707–10.

Parnell, Andrew. 2016. "Package 'Bchron'." https://cran.r-project.org/web/packages/Bchron/index.html.

Simpson, Gavin L. 2007. "Analogue Methods in Palaeoecology: Using the Analogue Package." *Journal of Statistical Software* 22: 1–29.

Widga, Chris, Stacey N Lengyel, Jeffrey Saunders, Gregory Hodgins, J Douglas Walker, and Alan D Wanamaker. 2017. "Late Pleistocene Proboscidean Population Dynamics in the North American Midcontinent." *Boreas*. Wiley Online Library.

Williams, John W., Eric C. Grimm, Jessica Blois, Donald F. Charles, Edward Davis, Simon J. Goring, Russell W. Graham, et al. in press. "The Neotoma Paleoecology Database: A Multi-Proxy, International Community-Curated Data Resource." *Quaternary Research*.