<RRH: Goring, et al.: CCDR Teaching Nexus>

THE NEOTOMA PALEOECOLOGY DATABASE: A RESEARCH-OUTREACH NEXUS

Goring, Simon James[1]

Graham, Russell[2]

Loeffler, Shane[3]

Myrbo, Amy[3]

Oliver, James S, III[2]

Ormand, Carol J[4]

Williams, John W[1,5]d

[1] Department of Geography, University of Wisconsin-Madison, Madison WI

goring@wisc.edu, jww@geography.wisc.edu

[2] Department of Geosciences, Pennsylvania State University, State College, PA

rwg12@psu.edu, j.s.oliver0156@gmail.com

[3] Department of Earth Sciences, University of Minnesota, Minneapolis, MN

loeff081@d.umn.edu, amyrbo@umn.edu

[4] Science Education Resource Center, Carleton College, Northfield, MN

cormand@carleton.edu

[5] Center for Climatic Research, University of Wisconsin-Madison WI

ABSTRACT

Paleoecological data from the Quaternary Period (2.58 – 0.012 Mya) provides a fruitful opportunity for educational outreach for the Earth and biological Sciences. Data repositories that steward paleoecological data are uniquely situated to help produce teaching modules and resources that support broader engagement, since they serve as technical hubs and focal points within their disciplinary research communities.

By supporting outreach and educational activities the Neotoma Paleoecology Database can provide support to educators at a number of educational levels, from primary schools to graduate students. In collaboration with pedagogical experts, the Neotoma Paleoecology Database team has developed several teaching modules

and model workflows. Early education is centered on discovery, while higher-level educational tools, such as programmatic workbooks, focus on illustrating best practices for technical tasks performed by researchers.

More broadly, we make the case that collaborations between pedagogic experts, technical experts and data stewards, centered around community curated data resources such as Neotoma, can provide an important role within research communities, and an important service to society, helping to support best practices, translate current research advances to interested audiences, and communicate the importance of individual research disciplines.

INTRODUCTION

The Quaternary Period (2.6 million years ago to present) is a highly active area of integrative ecological, geological, and climatic research. Quaternary timescales range from the decadal and centennial scales of our lives and historical memory to deep-time geological timescales of millions of years and longer. In addition, the Quaternary has been a time of planetary-scale climate changes, operating at many timescales, recorded in many different types of data sets, and with strong and clear effects on cryogenic, biological, and physical systems. This combination of factors – the range of timescales including human history plus the data-rich record of planetary climate change – makes the Quaternary a natural entry point for teaching and learning about paleoecology, geologic time, and climate change via authentic data exploration (e.g., Manduca and Mogk, 2002; Hunter et al., 2007; Laursen et al., 2010; Resnick et al., 2011, 2012).

The Quaternary is of particular interest to scientists studying climate change (Masson-Delmotte et al., 2013), the ecological effects of climate change (Araújo et al., 2008), and the causes and effects of species range shifts and extinctions (Ordonez and Svenning, 2017), providing 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 such as Heinrich Events and Dansgaard-Oeschger events (Masson-Delmotte et al., 2013). During the Holocene interglacial (11,700 years ago to present), temperature variations have been muted but hydrological variability has intensified (Mayewski et al., 2004), linked to both internal processes such as shifts in ENSO variability and external forcings such as individual volcanic eruptions and variations in solar luminosity (Wanner et al., 2008).

Ecologically, species distributions repeatedly retracted and expanded during the glacial-interglacial cycles (Davis and Shaw, 2001), with high local community turnover during climatic events (Blois et al., 2010). 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 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 alkanes from leaf waxes. These paleoenvironmental records are collected from a diverse array of depositional environments, including excavations, lake sediment core sampling, or other fossil localities. Understanding the large-scale spatial phenomena that define the Quaternary requires gathering these many diverse datasets into larger databases that store information about age, taxonomy, spatial coordinates, depth or stratigraphy, and other related information (Williams et al., 2018).

In response to these needs, the Neotoma Paleoecology Database (Neotoma) has emerged as a community-curated data resource (CCDR) with the mission of gathering, curating, and sharing paleoecological and paleoenvironmental data, to enable open, global-scale science (Williams et al., 2018). This resource also provides students and educators with a public portal for accessing data that can be used to understand ecological change through the last several million years. Authentic inquiry into real, complex data sets engages students in the learning process, thus deepening their understanding of species dynamics in changing environments, and simultaneously increasing their retention of key concepts (e.g. Laursen et al., 2010; Lopatto, 2010; Freeman et al., 2014). In recent years, new systems have been developed for finding, visualizing, exploring, and obtaining Neotoma data (Williams et al., 2018). 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 (https://serc.carleton.edu/neotoma/index.html). Phone-based apps have been developed for travelers to understand the geological and ecological history of the world around them (FlyOver Country: fc.umn.edu).

This chapter reviews these resources and introduces Neotoma as a new platform for supporting authentic teaching and outreach, able to support a wide range of users and levels of expertise. We begin with an overview of the Neotoma database and its software ecosystem, then move to an introductory exploration of Neotoma's data holdings, using the map-based Neotoma Explorer and a variant of exercises developed with the Science Education Resource Center (SERC) at

Carleton College (https://serc.carleton.edu/index.html). We then summarize teaching exercises for upper level undergraduate and graduate students using the software package R (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 comprises both a data repository (Neotoma DB) and an ecosystem of interlinked software for accessing and visualizing Neotoma data. Neotoma DB stores more than 25700 unique datasets from 12500 distinct sites, each containing observations of fossil organisms and associated geochronological and paleoenvironmental information spanning the last several million years. Neotoma DB is built around a flexible data model designed to describe paleoecological data records for many types of sites, depositional environments, taxonomic groups and geochemical and sedimentological measurements (Williams et al., 2018; Grimm, 2008). Neotoma DB 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 portal and discovery interface for finding and exploring paleoecological records within Neotoma (Fig. 2). Neotoma Explorer relies on a map-based user interface, in which users can search for data by taxonomic name, site name, spatial domain, time range, or many other dimensions. Discovered data are shown as sites on maps, which users can click to discover more information and download data. Neotoma Explorer is powered by an application program interface (API; http://api.neotomadb.org) that can be separately and directly accessed through any standard browser or programmatically. The use of an API provides a mechanism to build new tools on top of Neotoma's data resources. The API also supports the neotoma R package (Goring et al., 2015) and the inclusion of Neotoma data within the Flyover Country® mobile app (http://flyovercountry.io), and search tools that can jointly search Neotoma and other paleobiological resources (http://earthlifeconsortium.org/).

Neotoma Explorer: Data discovery and exploration

The following introductory exercise is based on Neotoma Explorer and is designed to help individuals understand how paleoecologists interpret and process proxy data, ranging from a single site to continental scales and from the recent past to

deep time records. The exercise, based on "Exploring the Neotoma Paleoecology Database" (https://serc.carleton.edu/neotoma/activities/121251.html), is directed toward college undergraduates and explores changes in mammal and vegetation distributions 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 paleoecologists to work with science education researchers to develop high-quality teaching materials for a variety of educational levels. Additional educational modules are available on the SERC-Neotoma webpage (https://serc.carleton.edu/neotoma/activities.html) and are described further below.

"Exploring Neotoma" teaches users how to search for paleoecological data through a variety of starting points (by site name, dataset type, and taxon name), find publications, and create simple mapped visualizations that show a) how species ranges changed as climates changed and ice sheets retreated and b) changing associations among plant and animal taxa. The exercise introduces students to research questions that paleoecologists might ask; for example, questions about species co-occurrence and shifting distributions under changing climate scenarios. "Exploring Neotoma" engages students through the authentic exploration of paleoecological datasets, and provides students with a foundation of skills they can use for further exploration and inquiry in subsequent exercises.

Searching for data with Neotoma Explorer: A student-centered walk-through

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 (Fig. 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 (**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 A table showing individual pollen counts, depth information and summary chronology.

Diagram	A tool to view the change in the changes in taxon presences or abundances over depth or time.
Site	Site level information for the data.
Chronology	The age-depth model used at that site and dataset.
Publications	Publications related to the dataset.

These window tabs provide you with a pretty good overview of the dataset, long term changes at the site, and a list of the publications available to learn more about the sample record.

Searching By dataset type

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*, perform the spatial search and then *click on a site* to view the data, as you did for Marion Lake. How do the variables in the water chemistry diagram differ from those in the pollen diagram? What similarities in the data can you see? Are rates of change different? What might be driving differences or similarities between these records?

Searching by taxon name and time

One very quick, informative search in Neotoma Explorer is to look for a particular species or taxon, within a set time range. (You can also restrict searches spatially; Explorer's default is to look within the bounds set by the mapped view on your computer.) Let's look for fossil pollen sites with *Picea* (spruce) present between 15,000 and 12,000 years ago. This time period is of interest because it is right at the end of the Pleistocene, as ice sheets are retreating, humans are arriving and dispersing across the Americas, and plant species ranges are shifting north. *Go to the 'Advanced' tab*, make sure all prior searches are cleared, then *enter 'pollen' in the 'Dataset' field and* Picea *in the 'Taxon Name' field*. Then, *check 'Abundance/density' and set abundance to >20%* (to only show sites with a lot of *Picea* pollen at them). Last but not least, *click on the 'Time' bar* and *enter a time range of 15,000 – 12,000 ybp*.

Mapping ice sheets

After the last search you should have a map with dots showing all the fossil pollen sites with at least 20% spruce pollen between 15,000 and 12,000 years ago. Feel free to explore them further. But where were the ice sheets? *Click on*

the white polygon on the top bar this supports the addition of overlays to map glacier extents through time. *Enter a date* (say 12,800 years ago) *hit enter,* and then *zoom out the map* so it shows most or all of North America. You should now see a white polygon, representing the continental ice sheets, overlain on the map of sites. You may also see blue polygons, that show the large lakes ('proglacial lakes') that formed around the glacial margins, fed by meltwater. If you still have the spruce sites displayed, you should see that spruce was growing just south of the ice sheets.

Research questions

We all know that a whole bestiary of large animals roamed the earth during the last age, many of which are now extinct. In North America, representative species of the 'Pleistocene megafauna' include mastodons (*Mammut*), mammoths (*Mammuthus*), saber-toothed tigers (*Smilodon*), and many others. Where did these animals live, did their distributions change over time, and were they associated with particular habitats? We've already got the spruce trees and ice sheets mapped, so now let's see where were the big animals. Let's use mastodons in this example. Do a taxon search for *Mammut* (mastodons) and look at the site distribution. What happened? Describe the site distribution. Mastodons are known to be browsers, eating leaves and woody plant tissue. What can we say about the kinds of habitats or environments that mastodons may have occupied?

Now, let's add mammoths, using Neotoma Explorer's multi-taxon search. **Figure 4** shows the button needed to select multiple related taxa through the "Basic Search" window. *Click the gear box* to search for multiple taxa. First, *make sure that 'Mammals' is entered under Taxa Group.* Then, *search and select all the* Mammut *entries in the taxon list* – this will return both mastodons (*Mammut*) and mammoths (*Mammuthus*). How has the apparent distribution changed?

More exploration with Explorer: Additional SERC teaching exercises

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,"** is a unit for undergraduate students, focused on examining environmental gradients over short spatial scales and the ways these gradients can 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 to answer the question (e.g., Hmelo-Silver et al., 2007).

Educational modules for climate change, paleoecology, and biogeography

An additional series of eleven modules was developed by vertebrate paleontologists at Penn State University for participants to learn about climate change, paleoecology, and biogeography using the Neotoma Paleoecology Database. These modules have also been posted to the SERC website for distribution ((https://serc.carleton.edu/neotoma/activities.html)[]). These modules are primarily focused on mammals but can easily be adapted to other organisms like beetles, plants, ostracodes, and other groups with rich data holdings in Neotoma. There are two basic module types: (1) Modules that provide background information about 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 to introduce learners to fundamental concepts in climate change, paleoecology, and biogeography. For instance, in the Paleoecology modules, relationships among climate variables (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 variables. From this, students can see that the eastern and western limits of the species' distributions in the eastern United States are limited by water availability, whereas the northern and southern range limits are temperature-dependent. The paleoecology module also briefly discusses underlying assumptions linking climate variables and species' distributions and how additional factors can come into play. For instance, the distribution of the prairie dog (*Cynomys ludovicianus*) may be controlled by climate but also soil properties such as soil depth, grain 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 provide further explanation of the principles.

In the applied modules, learners build on the knowledge they have gained from the background modules and apply it to analyses of Neotoma data. Each module begins with a statement about the process to be learned and an example of similar analysis is provided. Learners develop a hypothesis (e.g., "If the climate warms, then a species' distribution will expand westward"). The student can

then test the hypothesis, by examining changes in the distribution of the species during different times of climate change using climate data from the Greenland Ice Core Project (GRIP: Andersen et al., 2004), covering the transition from full glacial to late glacial (~20,000 to 13,000 years ago) or late glacial to late Holocene (~13,000 years ago to present). Mammal species are assigned to students as they continue the exercise. If the hypothesis is false, then the student is asked to think about why it failed. The learner is then taken to a file that explains what potentially happened with their analysis. Finally, based upon these analyses, the student is asked to draw lessons from the past, thinking about the potential response of these species to current climate trends. 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 persists. With the addition of FAUNMAP II and MIOMAP data that is currently being uploaded, participants will be able to work with not only the glacial to interglacial warming of the late Quaternary, but will be able to go back through the Miocene, 25 million years ago, a time period that spans a long-term cooling in the Earth System, the onset of ice sheets in the Northern Hemisphere, and regionally increasing aridity. These exercises will involve rates of immigration and extinction for various mammal species.

We are developing additional modules with an emphasis on the emerging field of conservation paleobiology (Barnosky et al., 2017; Dietl and Flessa, 2011; Dietl et al., 2015) and more sophisticated analyses of groups of species and past community dynamics.

Programmatic access for macro-scale paleoecological research (APIs, neotoma R, and GitHub)

For advanced undergraduates and graduate students, we have created educational resources to introduce them to the programmatic tools for accessing Neotoma data: Neotoma Explorer is excellent for quick-look visualizations, data exploration, 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 (https://api.neotomadb.org/) 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; Goring et al., 2015) offers a convenient system for exporting Neotoma data into the R environment, which is widely used by ecologists and paleontologists for statistical data analysis. The R package is documented in Goring et al. (2015) and includes worked examples drawn from the literature to showcase the ways that the R package can be used in research workflows.

Over the past several years, we and others have led several short courses and workshops to show early career scientists how to use the neotoma R package and APIs to obtain data from Neotoma and pass it to their analytical workflows, with

workshops at the Society for Vertebrate Paleontology, International Biogeography Society, American Quaternary Association, and other venues. As part of these workshops, we have developed training materials that are hosted on GitHub (http://neotomadb.github.io/workbooks.html). GitHub (http://github.com) is a particularly useful system for open-source software development because of 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/), providing the opportunity to host educational material, along with the source code for 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). All materials are posted with an MIT license, so any interested student or educator can access workshop resources, download copies, and modify them for their own research or educational purposes.

Teaching materials on GitHub use the software principle of version control to borrow elements across workshops, while allowing workshop leaders to adopt best practices from 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. For example, materials from the 2016 SVP meeting are posted there (https://neotomadb.github.io/workbooks/Workshop_SVP2016_v0.1.html#the-neotoma-package), and new workshop materials are being actively developed.

Model workflows: Chronology construction & climate reconstruction

Two of the workflows posted to the Neotoma workbooks page (http://neotomadb.github.io/workbooks/AgeModels were not produced for specific workshops but instead were designed as model workflows for common tasks in paleoecology and paleoclimatology: chronology construction and paleoclimatic reconstruction via transfer functions. Both are complex tasks. Chronology construction is central to making age inferences and placing geological events in a temporal framework (Harrison et al., 2016; Grimm et al., 2014), while transfer functions and paleoeceological data are often used to reconstruct past climates (e.g. Marsicek et al., 2018; Bartlein et al., 2011; Viau et al., 2012), but fraught with methodological issues (Juggins, 2013). Within the Quaternary geosciences there are many techniques that can be used to construct chronologies, and packages that can be used to reconstruct climate from proxy data. Age-depth models such as Bacon (Blaauw and Christen, 2011) and BChron (Parnell, 2016) for chronologies, or rioja (Juggins, 2017) and analogue (Simpson, 2007) for climate reconstruction, have a number of settings and requirements. However, the paleoecological community is highly distributed among institutions, and few departments or institutions offer advanced courses in quantitative paleoecological methods, so many students

might not encounter a course on chronology construction or paleoclimate reconstruction. Hence, the educational materials hosted on Neotoma's GitHub Pages site offer example code and narrative discussion of a best practices workflow process in an approachable format, for both chronology construction (http://neotomadb.github.io/workbooks/AgeModels.html) and paleoclimatic reconstruction with fossil pollen records (https://neotomadb.github.io/workbooks/clim_notebook.html). These can be used in a lab setting for Earth Science educators, or for students (graduate or senior undergraduates) who wish to undertake similar analysis.

Flyover Country®: Beyond the classroom

Creating open-access, high-quality community-curated scientific data resources such as Neotoma opens up novel opportunities to engage with entirely new audiences. One such is the use of Neotoma data in the Flyover Country® mobile app for geoscience (http://flyovercountry.io, supported by the National Science Foundation), available for iOS and Android devices. The app is designed to be used by travelers (or armchair explorers) to learn more about the geological world around them as they travel over it. Flyover Country® draws upon multiple resources, including Macrostrat, Paleobiology Database, and Neotoma. For Neotoma, Flyover Country® uses the Neotoma API to find fossil occurrences of extinct "charismatic megafauna" (mastodons, saber-tooth tigers, giant beavers, etc.) along a traveler's path. Users click on the map to establish their travel waypoints, and then Flyover Country® 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. Mobile and web app development is within the reach of undergraduate courses, and Neotoma's well-documented API and the ready availability of Neotoma data through the API means that those apps can be easily populated with research-quality scientific data. Neotoma's spatial, multivariate, and temporal data provide additional interesting opportunities for data visualization in such courses.

DISCUSSION

CCDRs at the research-outreach nexus

The development of community curated data resources (CCDRs) is often originally motivated by research objectives, by a community of researchers attempting to answer macro-scale research questions that can only be answered by analyzing many individual site-level datasets with some standard data representation. Traditionally, CCDRs have been a resource for experts only, serving systems of record and as vehicles for collating, aggregating, cleaning, and analyzing data (Kapoor et al., 2015). As mechanisms to access, share, and visualize data improve

across the geosciences, these community-curated research data resources generate new opportunities for individuals and groups outside of the traditional research community to access and interact with scientific data. This in turn moves these databases from systems of record and research to systems of engagement, opening up new social and collaborative elements to these platforms. As such, building educational activities that closely interlink with CCDRs represents a major opportunity for broader engagement, both within dispersed networks of scientific researchers and out to new and previously unreached communities.

The role for CCDRs in outreach

By acting as a portal for outreach, CCDRs can serve multiple, overlapping communities, and can bring them together in collaborative endeavors (Fig. 6). For example, the project "Neotoma: Community-led Cyberinfrastructure for Global Change Research" (NSF-Geoinformatics), in partnership with SERC, was able to bring together a broad group of researchers Earth science educators, connecting and leveraging their diverse expertises. The strength of these curricular materials is multi-faceted: the exercises make use of the 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 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 researchers to educators, and educators to the broad range of cutting-edge research activities, building on the strengths of each group.

The on-going informatics revolution, the rise of data sciences as a distinct field of inquiry (Blei and Smyth, 2017), and associated emergence of paleoecoinformatics (Brewer et al., 2012) has enhanced the paleoecology community's suite of tools, bringing geospatial tools, and the technical capacity to combine web or mobile mapping tools with data resources. There is no one-size-fits-all solution to finding and accessing data and using these data to understand core concepts; different pathways serve different student and user opulations. 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 (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 further opportunities to introduce technology as a core component of curriculum design within the geosciences. CCDRs act as important translational actors in this ecosystem, drawing on best practices in both the data and educational sciences, and applying it to scientific outreach programs with the assistance of education experts (Williams et al., 2018).

In summary, there are a number of educational opportunities available online for the Neotoma Paleoecology Database, with more being actively developed. All are open source and can be modified to explore different questions or data types (e.g., plants, mammals, diatoms or limnological datasets). They serve a broad range of audiences, from secondary to undergraduate to graduate-level students. CCDRs thus can serve dual missions, both as critical research infrastructure pedagogical infrastructure, supporting both distributed research communities and opportunities for education and outreach.

ACKNOWLEDGEMENTS AND AUTHORSHIP

REFERENCES

Andersen, K.K., Azuma, N., Barnola, J.-M., Bigler, M., Biscaye, P., Caillon, N., Chappellaz, J., Clausen, H.B., Dahl-Jensen, D., Fischer, H., and others, 2004, High-resolution record of Northern Hemisphere climate extending into the last interglacial period: Nature, v. 431, p. 147.

Araújo, M.B., Nogués-Bravo, D., Diniz-Filho, J.A.F., Haywood, A.M., Valdes, P.J., and Rahbek, C., 2008, Quaternary climate changes explain diversity among reptiles and amphibians: Ecography, v. 31, p. 8–15.

Barnosky, A.D., Hadly, E.A., Gonzalez, P., Head, J., Polly, P.D., Lawing, A.M., Eronen, J.T., Ackerly, D.D., Alex, K., Biber, E., and others, 2017, Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems: Science, v. 355, p. eaah4787.

Bartlein, P., Harrison, S., Brewer, S., Connor, S., Davis, B., Gajewski, K., Guiot, J., Harrison-Prentice, T., Henderson, A., Peyron, O., and others, 2011, Pollen-based continental climate reconstructions at 6 and 21 ka: A global synthesis: Climate Dynamics, v. 37, p. 775–802.

Blaauw, M., and Christen, J.A., 2011, Flexible paleoclimate age-depth models using an autoregressive gamma process: Bayesian Analysis, v. 6, p. 457–474.

Blei, D.M., and Smyth, P., 2017, Science and data science: Proceedings of the

National Academy of Sciences, v. 114, p. 8689–8692.

Blois, J.L., McGuire, J.L., and Hadly, E.A., 2010, Small mammal diversity loss in response to late-Pleistocene climatic change: Nature, v. 465, p. 771.

Brewer, S., Jackson, S.T., and Williams, J.W., 2012, Paleoecoinformatics: Applying geohistorical data to ecological questions: Trends in Ecology & Evolution, v. 27, p. 104–112.

Davis, M.B., and Shaw, R.G., 2001, Range shifts and adaptive responses to Quaternary climate change: Science, v. 292, p. 673–679.

Dietl, G.P., and Flessa, K.W., 2011, Conservation paleobiology: Putting the dead to work: Trends in Ecology & Evolution, v. 26, p. 30–37.

Dietl, G.P., Kidwell, S.M., Brenner, M., Burney, D.A., Flessa, K.W., Jackson, S.T., and Koch, P.L., 2015, Conservation paleobiology: Leveraging knowledge of the past to inform conservation and restoration: Annual Review of Earth and Planetary Sciences, v. 43.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., and Wenderoth, M.P., 2014, Active learning increases student performance in science, engineering, and mathematics: Proceedings of the National Academy of Sciences, v. 111, p. 8410–8415.

Goring, S., Dawson, A., Simpson, G., Ram, K., Graham, R., Grimm, E., and Williams, J., 2015, neotoma: A programmatic interface to the Neotoma Paleoecological Database: Open Quaternary, v. 1.

Grimm, E., 2008, Neotoma: An ecosystem database for the Pliocene, Pleistocene, and Holocene: Illinois State Museum Scientific Papers E Series, v. 1.

Grimm, E.C., Blaauw, M., Buck, C., and JW, W., 2014, Age models, chronologies, and databases workshop: Complete report and recommendations: PAGES Magazine, v. 22, p. 104.

Harrison, M., Baldwin, S., Caffee, M., Gehrels, G., Schoene, B., Shuster, D., and Singer, B., 2016, Geochronology: It's about time: Eos, v. 97, p. 12–13.

Hmelo-Silver, C.E., Duncan, R.G., and Chinn, C.A., 2007, Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006): Educational Psychologist, v. 42, p. 99–107.

Hunter, A.-B., Laursen, S.L., and Seymour, E., 2007, Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development: Science education, v. 91, p. 36–74.

Juggins, S., 2013, Quantitative reconstructions in palaeolimnology: New paradigm or sick science? Quaternary Science Reviews, v. 64, p. 20–32.

Juggins, S., 2017, Rioja: Analysis of quaternary science data: http://www.staff.

ncl.ac.uk/stephen.juggins/.

Kapoor, S., Mojsilovic, A., Strattner, J.N., and Varshney, K.R., 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.

Laursen, S., Hunter, A.-B., Seymour, E., Thiry, H., and Melton, G., 2010, Undergraduate research in the sciences: Engaging students in real science: John Wiley & Sons.

Lopatto, D., 2010, Undergraduate research as a high-impact student experience: Peer Review, v. 12, p. 27.

Manduca, C., and Mogk, D., 2002, Using data in undergraduate science classrooms, *in* Final Report of the National Science Digital Library Workshop,

Marsicek, J., Shuman, B.N., Bartlein, P.J., Shafer, S.L., and Brewer, S., 2018, Reconciling divergent trends and millennial variations in Holocene temperatures: Nature, v. 554, p. 92.

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., Rouco, J.G., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., et al., 2013, Information from paleoclimate archives: Climate change, v. 383464, p. 2013.

Mathewes, R.W., and Heusser, L.E., 1981, A 12 000 year palynological record of temperature and precipitation trends in southwestern British Columbia: Canadian Journal of Botany, v. 59, p. 707–710.

Mayewski, P.A., Rohling, E.E., Stager, J.C., Karlén, W., Maasch, K.A., Meeker, L.D., Meyerson, E.A., Gasse, F., Kreveld, S. van, Holmgren, K., and others, 2004, Holocene climate variability: Quaternary research, v. 62, p. 243–255.

Ordonez, A., and Svenning, J.-C., 2017, Consistent role of Quaternary climate change in shaping current plant functional diversity patterns across European plant orders: Scientific Reports, v. 7, p. 42988.

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

Resnick, I., Atit, K., Shipley, T., and others, 2012, Teaching geologic events to understand geologic time: Earth and mind II: a synthesis of research on thinking and learning in the geosciences, p. 41–43.

Resnick, I., Shipley, T.F., Newcombe, N., Massey, C., and Wills, T., 2011, Progressive alignment of geologic time, *in* Geological Society of America Annual Meeting, Minneapolis, MN, http://w3w.spatiallearning.org/archives/showcase_archive/showcase_pdfs/showcase_resnick_sep2011.pdf.

Simpson, G.L., 2007, Analogue methods in palaeoecology: Using the analogue

package: Journal of Statistical Software, v. 22, p. 1–29.

Viau, A., Ladd, M., and Gajewski, K., 2012, The climate of North America during the past 2000 years reconstructed from pollen data: Global and Planetary Change, v. 84, p. 75–83.

Wanner, H., Beer, J., Bütikofer, J., Crowley, T.J., Cubasch, U., Flückiger, J., Goosse, H., Grosjean, M., Joos, F., Kaplan, J.O., and others, 2008, Mid-to late Holocene climate change: An overview: Quaternary Science Reviews, v. 27, p. 1791–1828.

Williams, J.W., Grimm, E.C., Blois, J., Charles, D.F., Davis, E., Goring, S.J., Graham, R.W., Smith, A.J., Anderson, M., Arroyo-Cabrales, J., Ashworth, A.C., Betancourt, J.L., Bills, B.W., Booth, R.K., et al., 2018, The Neotoma Paleoecology Database: A multi-proxy, international community-curated data resource: Quaternary Research, v. 89, p. 156–177.