

Storing vertebrates in the cloud

August 24 2011, By Robert Sanders



A trumpet fish and tiger shark, two vertebrate species represented in the world's museum collections. Collection information on these and all other vertebrates will soon to be stored in the cloud for use by researchers and citizen scientists.

What Google is attempting for books, the University of California, Berkeley, plans to do for the world's vertebrate specimens: store them in "the cloud."

Online [storage](#) of information from vertebrate collections at the Smithsonian Institution, [American Museum of Natural History](#), [National Museum of Natural History](#) in Paris, UC Berkeley's Museum of Vertebrate Zoology and from hundreds of other animal collections around the world — or at least, all collections that include animals with backbones — will make them readily available to academic researchers and citizen scientists alike.

Computing clouds are shared pools of servers that can be accessed from anywhere, anytime, and are far more reliable than computer servers at individual institutions.

The project to create VertNet, a cloud-based collection of vertebrate specimens, got off the ground this summer thanks to a three-year, \$2.4 million grant from the National Science Foundation. The effort is led by UC Berkeley museum curators and involves colleagues from the University of Colorado, Boulder; University of Kansas in Lawrence; and Tulane University in New Orleans.

VertNet coordinator David Bloom said VertNet would have been invaluable after the 2010 oil spill in the Gulf of Mexico, telling scientists and others about the animals potentially affected.

"But no one knew where the fauna were and how they might be impacted, even though the information that could tell them resided in bird and fish collections around the country," Bloom said. "There was just no single place to go to find out where the collections were, and no easy way to put these data together, even now. We have to get started."

VertNet will pave the way for similar cloud-based resources consolidating animal and plant information from state, national and international collections. UC Berkeley itself is involved in NSF-funded projects to digitize audio, plant, insect and paleontology collections.

VertNet merges four earlier databases

VertNet will merge four successful but financially unsupported database networks — MaNIS for mammals, ORNIS for birds, HerpNET for reptiles and amphibians; and FishNet. UC Berkeley museum scientists have played a leadership role in creating and maintaining all but FishNet.

The current version of VertNet allows scientists to search these four specialized networks, but only by sending simultaneous queries to online databases at 74 institutions housing 174 separate collections.

"The architecture of these networks was not able to keep up with the demand, and queries were getting dropped or encountering servers that were offline or down," said Carla Cicero, staff curator for birds at the MVZ and principal investigator for the VertNet project. "VertNet will create a completely new cloud-based platform that eliminates the need for any individual collection to have servers or hardware to maintain or manage."

Collection use skyrockets

The MVZ has seen firsthand the impact of digitizing collections and making them searchable through the Internet, according to MVZ information architect John Wieczorek. Researchers used to call the museum and ask for printouts of information about specimens, but they now can do that online, leaving museum staff more time to double check the online data and add GPS locations to each, a process called georeferencing. The information on a museum specimen can range from basic — the species name and where and when it was collected — to extensive field notes, photographs, audio recordings and information about tissue samples.

"After we put information on our specimen and tissue collections online, usage skyrocketed," Wieczorek said. "The number of specimen records delivered in response to queries went from the hundreds of thousands to tens of millions per year."

Once the collections data are in the cloud, an increasing number of online applications will allow manipulation and display of the data beyond the basic maps now possible.

"The power of VertNet lies in new ways of discovering and integrating data for biodiversity science," Bloom said.

Seaweed, lichens and bryophytes

VertNet's success should lead other museums, at UC Berkeley and elsewhere, to upload their data to the cloud.

"The cloud idea — getting data out where it's accessible in a variety of ways — is very attractive," said Richard L. Moe, an information technologist and seaweed expert at the University and Jepson Herbaria. "Its capacity for indexing is unparalleled."

For now, though, Moe is working on an NSF-funded project to merge the California plant collections from 17 state institutions into one online database, which has many of the capabilities of a cloud-based collection. UC Berkeley's herbaria have already digitized their 360,000 specimens of California vascular plants (ferns, flowering plants and gymnosperms). Moe is collaborating with institutions as diverse as Cal State Chico and the Rancho Santa Ana Botanic Garden, all members of the Consortium of California Herbaria, to help bring their specimens online.

These online collections were used a few years ago by UC Berkeley researchers to predict the impact of climate change on California's endemic land plants.

"The value of online specimens is not only to document existing and new species, but also to investigate the spread of invasive species and future changes to distributions of native species and communities," said Brent Mishler, UC Berkeley professor of integrative biology and director of the campus herbaria. "A similar dataset is needed to document changes in the marine flora."

Mishler garnered a grant from NSF to bring all of UC Berkeley's California seaweed collections, including images, online. The herbaria also is part of a nationwide digitization effort, funded by NSF's

Advancing Digitization of Biological Collections program, aimed at all museum lichen and bryophyte (mosses and their kin) collections.

The Essig Museum of Entomology is also digitizing its collections as part of the Calbug consortium of eight state institutions that have insect collections.

VertNet in cloud by 2012

By summer of next year, Bloom, Cicero and Wieczorek should have the current vertebrate collections "mobilized" to the cloud and have started to bring onboard some 75 institutions now on the waiting list, including museums in Ecuador and Africa.

VertNet also can draw more amateur or citizen scientists into research. Already, ORNIS, by far the largest of the vertebrate databases at 75 million of VertNet's 85 million records, incorporates field observations from birders via eBird.

But the public is clearly interested in other resources in VertNet, Cicero said. "Digitization of original collectors' notes for the MVZ egg collection resulted in a jump in online searches from about 900 per month to more than 9,000 per month, and a lot of our hits are from [Google](#) searches," said Cicero, who was a principal investigator for ORNIS. "Putting these data online increased the exposure of our collection tenfold."

She hopes soon to add links to VertNet of UC Berkeley's bird call recordings, as well as links to photos of bird eggs.

VertNet also will incorporate paleontological data, which expands the scope of the original four vertebrate database networks. This will enable researchers to study how species have changed through both space and

time, Cicero said.

"Initially, there was a lot of hesitation on the part of institutions who didn't want to let go of their data," Wieczorek said of earlier efforts, such as MaNIS, to combine collection information online. "But we showed them that there are tangible benefits and that more people use their collections. Institutions are now eager to be part of it."

Provided by University of California - Berkeley

Citation: Storing vertebrates in the cloud (2011, August 24) retrieved 24 April 2024 from <https://phys.org/news/2011-08-vertebrates-cloud.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.