Deep-time Digital Earth aims to liberate data from collections such as the British Geological Survey's.

DATA SHARING

Earth scientists plan a 'geological Google'

China is backing an international effort to meld geoscience databases into a one-stop data shop

By **Dennis Normile**

The British Geological Survey (BGS) has amassed one of the world's premier collections of geologic samples. Housed in three enormous warehouses in Nottingham, U.K., it contains about 3 million fossils gathered over more than 150 years at thousands of sites across the country. But this data trove "was not really very useful to anybody," says Michael Stephenson, a BGS paleontologist. Notes about the samples and their associated rocks "were sitting in boxes on bits of paper." Now, that could change, thanks to a nascent international effort to meld earth science databases into what Stephenson and other backers are describing as a "geological Google."

This network of earth science databases, called Deep-time Digital Earth (DDE), would be a one-stop link allowing earth scientists to access all the data they need to tackle big questions, such as patterns of biodiversity over geologic time, the distribution of metal deposits, and the workings of Africa's complex groundwater networks. It's not the first such effort, but it has a key advantage, says Isabel Montañez, a geochemist at University of California, Davis, who is not involved in the project: funding and infrastructure support from the Chinese government. That backing "will be critical to [DDE's] success given the scope

of the proposed work," she says.

In December 2018, DDE won the backing of the executive committee of the International Union of Geological Sciences, which said ready access to the collected geodata could offer "insights into the distribution and value of earth's resources and materials, as well as hazards—while also providing a glimpse of the Earth's geological future." At a meeting this week in Beijing, 80 scientists from 40 geoscience organizations including BGS and the Russian Geological Research Institute are discussing how to get DDE up and running by the time of the International Geological Congress in New Delhi in March 2020.

DDE grew out of a Chinese data digitization scheme called the Geobiodiversity Database (GBDB), initiated in 2006 by Chinese paleontologist Fan Junxuan of Nanjing University. China had long-running efforts in earth sciences, but the data were scattered among numerous collections and institutions. Fan, who was then at the Chinese Academy of Sciences's Nanjing Institute of Geology and Paleontology, organized GBDB around the stacks of geologic strata called sections and the rocks and fossils in each stratum.

Norman MacLeod, a paleobiologist at the Natural History Museum in London who is advising DDE, says GBDB has succeeded where similar efforts have stumbled. In the past, he says, volunteer earth scientists tried to do nearly everything themselves, including

informatics and data management. GBDB instead pays nonspecialists to input reams of data gleaned from earth science journals covering Chinese findings. Then, paleontologists and stratigraphers review the data for accuracy and consistency, and information technology specialists curate the database and create software to search and analyze the data. Consistent funding also contributed to GBDB's success, MacLeod says. Although it started small, Fan says GBDB now runs on "several million" yuan per year.

Earth scientists outside China began to use GBDB, and it became the official database of the International Commission on Stratigraphy in 2012. BGS decided to partner with GBDB to lift its data "from the page and into cyberspace," as Stephenson puts it. He and other European and Chinese scientists then began to wonder whether the informatics tools developed for GBDB could help create a broader union of databases. "Our idea is to take these big databases and make them use the same standards and references so a researcher could quickly link them to do big science that hasn't been done before," he says.

The Beijing meeting aims to finalize an organizational structure for DDE. Chinese funding agencies are putting up $75 million over 10 years to get the effort off the ground, Fan says. That level of support sets DDE apart from other cyberinfrastructure efforts "that are smaller in scope and less well funded," Montañez says. Fan hopes DDE will also attract international support. He envisions nationally supported DDE Centers of Excellence that would develop databases and analytical tools for particular interests. Suzhou, China, has already agreed to host the first of them, which will also house the DDE secretariat.

DDE backers say they want to cooperate with other geodatabase programs, such as BGS's OneGeology project, which seeks to make geologic maps of the world available online. But Mohan Ramamurthy, project director of the U.S. National Science Foundation–funded EarthCube project, sees little scope for collaboration with his effort, which focuses on current issues such as climate change and biosphere-geosphere interactions. "The two programs have very different objectives with little overlap," he says.

Fan also hopes individual institutions will contribute, by sharing data, developing analytical tools, and encouraging their scientists to participate. Once earth scientists are freed of the drudgery of combing scattered collections, he says, they will have time for more important challenges, such as answering "questions about the evolution of life, materials, geography, and climate in deep time." ∎

Science

Earth scientists plan a 'geological Google'

Dennis Normile

View the article online
Permissions

Use of this article is subject to the Terms of service