

New tools to boost access to NASA Earth science data

September 26 2018, by Ellen Gray



NASA has 26 Earth-observing satellites monitoring the vital signs of our home planet. Along with airborne and ground Earth science missions, their data is stored and managed by the Earth Observing System Data and Information System. Credit: NASA

NASA has funded five new projects to develop tools and technology to make the agency's massive Earth science datasets more accessible and user-friendly.



Wake up. Turn on laptop. Start processing airborne data of the Adirondack forests in New York. Make Coffee. Eat Breakfast. Fasten the open laptop's seatbelt in the passenger seat as it continues to crunch numbers. Drive to work.

That used to be Sara Lubkin's morning routine as an early career scientist at NASA's Goddard Space Flight Center in Greenbelt, Maryland. Once at work, she would use her desktop computer, while her laptop diligently spent the next 12 hours processing airborne instrument data for the relevant information she needed to study invasive pests of hemlock trees.

Earth visualization

NASA Earth science datasets provide different perspectives and information on our planet, as seen here in this data visualization of observations of Hurricane Matthew in October 2016.Credits: NASA's Scientific Visualization Studio

"I'm not a computer scientist, I'm an Earth scientist," said Lubkin, who now works as a program officer for NASA Earth Science Data Systems' Advancing Collaborative Connections for Earth Systems Science, or ACCESS program. But her experience as a researcher is not unique.

Spending large chunks of time simply getting Earth science data into a usable form for analysis is a common situation for researchers working with the big datasets that come from NASA field, airborne and satellite missions. Downloading huge files, converting data formats, locating the same study areas in multiple datasets, writing code to distinguish different land types in a satellite image—these types of tasks eat into time scientists would rather be using to analyze the actual information in the data.



That's where the ACCESS program comes in. Part of the Earth Science Data Systems division since 2005, ACCESS finds innovative ways to streamline that cumbersome processing time. The program funds twoyear research projects to improve behind-the-scenes data management and provide ready-to-use datasets and services to scientists, Lubkin said.

In June, NASA selected five teams of NASA, university and commercial computer science researchers from the 2017 round of submissions in a range of projects that will use machine learning, cloud computing and advanced search capabilities to develop tools to improve the behind-the-scenes management for selected NASA datasets.



NASA Earth science datasets provide different perspectives and information on our planet, as seen here in this data visualization of observations of Hurricane Matthew in October 2016. Credit: NASA's Scientific Visualization Studio



"We continually invest in development and evaluation of the newest technologies to improve science data systems," said Kevin Murphy, program executive for NASA's Earth Science Data Systems at NASA Headquarters in Washington. But more than that, they want to make sure that the tools and technology help real scientists address real problems.

Each ACCESS project has Earth sciences and computer scientists involved from beginning to end, Murphy said. "With the ACCESS program, we're really trying to understand, for example, how ocean currents work, but we're trying to do that now with data that's so large that we need a team of experts who can work together to solve the big science and big data questions."

The projects will complement data management, distribution and other services provided by the Earth Observing System Data and Information System (EOSDIS), which manages and stores NASA data collected from Earth-observing satellites, aircraft and field campaigns. EOSDIS has 12 interconnected data and archive centers located across the United States, which are organized by discipline. Currently, these centers host 26 petabytes of Earth datasets—that's 26 million gigabytes, or enough data to need 52,000 computers each with 500 gigabytes of storage space. That number is expected to grow to 150 petabytes within five years with the launch of new satellites.

"Satellite data is big data," said Jeff Walter, one of the ACCESS 2017 principal investigators and lead engineer for Science Data Services at the Atmospheric Science Data Center at NASA's Langley Research Center in Hampton, Virginia. "It's very complex and sometimes difficult to use, even for expert users. In addition to the volume, which makes it difficult for users to acquire, store and manage, there's also the complexity of both the format and content. Users often have to spend a lot of time understanding how the data is organized and what the various parameters represent."



Walter's project is one of three that will use cloud computing to alleviate download and storage issues for users. Starting with two atmospheric datasets, his team will also be developing a way to convert satellite data formats into those that can be read by commercial geospatial information system (GIS) software.

"Our project aims to lower the barrier to entry for a potential new user community who might find novel ways to use this data, and who are more familiar with GIS types of tools," Walter said.

The two other cloud computing projects will be developing open source processing and analysis tools, including one designed for ocean datasets. A fourth project will use machine learning to detect changes over time in land observations, starting with the detection of landslides, floods and uplift caused by volcanic activity. The fifth project will develop an automated method for lining up datasets that observe the same location so researchers can combine more than one type of information about a place.

Upon completion, the ACCESS researchers will work closely with EOSDIS teams to incorporate their advancements into the data centers' day-to-day operations. Once those new tools are in place, that's when the real power of open and freely available Earth science datasets can flourish, according to Murphy. Easy-to-use data means it gets into the hands of decision-makers, non-governmental organizations, scientists studying related applications and researchers in different fields that may have new uses for it.

"When you make these products open and accessible, you have a lot of unintended, good scientific consequences," Murphy said, citing examples that include detecting groundwater movement from space, rapid wildfire detection and using night lights to study human energy use. "NASA has a lot of very valuable information, and the ACCESS



program really tries to help scientists to not only address primary science questions but also help us understand our environment and plan for our future."

Provided by NASA's Goddard Space Flight Center

Citation: New tools to boost access to NASA Earth science data (2018, September 26) retrieved 28 April 2024 from <u>https://phys.org/news/2018-09-tools-boost-access-nasa-earth.html</u>

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.