

July 2020 Monthly Meeting – Science Data Management and New Applications for Field Data Collection

For July's CDI Monthly Meeting, we heard two presentations: one on science data management within USGS, and another on the NGA's new mobile and web applications for field data collection!

For more information, questions and answers from the presentation, and a recording of the meeting, please visit the [CDI wiki](#).

It's July 2020: Do you know what your data are doing? SAS Science Data Management: Contributing to USGS progress in management of scientific data - Viv Hutchison, USGS



CSS SAS Science Data Management Team

Mission: Optimize and share USGS data management practices and workflows to ensure that our science data are fully described, preserved, and publicly accessible
Building Communities



Community for Data Integration
usgs.gov/cdi



USGS Powell Center
powellcenter.usgs.gov

ScienceBase Data Release Network

Promoting Best Practices



USGS Data Management Website
usgs.gov/datamanagement

Leadership in Policy



Developing Enterprise Tools

- ScienceBase
sciencebase.gov
- DOI Creation Tool
www1.usgs.gov/csas/doi
- Online Metadata Editor
www1.usgs.gov/csas/ome
- Metadata Wizard
<https://code.usgs.gov/usgs/fort-pymdwizard>
- Science Data Catalog
data.usgs.gov

Overview of CSS-SAS-SDM

Most of us have probably heard of data management, but why should we take care to manage our scientific data well? Good data management increases reproducibility and integrity for Earth science data. As such, it's important that data is FAIR (findable, accessible, interoperable, and reusable) and well-maintained.

The Science Data Management (SDM) branch within Science Analytics and Synthesis (SAS) leverages tools and expertise for good data management, and encourages community engagement around this topic. SDM has made strides towards better data management and measuring impact of data.

ScienceBase (SB), an online digital repository for science data, became a trusted digital repository (TDR) in 2017, meeting rigorous standards to attain this status. Many journals require that data accompanying an article is made publicly available, and ScienceBase is an easy way to accomplish this requirement. The ScienceBase Data Release (SBDR) Tool, which allows scientists to easily complete a data release, connects seamlessly to other USGS tools such as the DOI Tool, IPDS, and the [Science Data Catalog \(SDC\)](#). The SBDR Tool can be customized to reflect a science center's specific workflow as well. There are currently 92 USGS science centers use SB for data release. The TDR has seen a steady increase in usage over time, and is now accommodating approximately 1,000 data releases per year. The upcoming SBDR Summary Dashboard will share data release metrics by science center, program, region, mission area. For help with ScienceBase data release, find the instructions page [here](#), or contact sciencebase_datarelease@usgs.gov with other questions.

SDM has strengthened the connection between USGS publications and supporting data. The team worked with the [USGS Pubs Warehouse](#) to collect information on known primary publications related to a data release, then added those related publications to the ScienceBase landing page and the DOI Tool. This link has proven useful for letting data authors know how others are using their data, and for understanding some impacts of the data. In a similar vein, SDM uses xdd (previously [GeoDeepDive](#)) to track references to USGS data DOIs, with plans to display these data citations on ScienceBase landing pages in the future.

Citation of data is an emerging practice, but many data releases in ScienceBase have seen multiple instances of reuse in subsequent scientific research. For example, data for a [Geospatial Fabric data release](#) has been cited or reused by seventeen publications. Another data release on the [global distribution of minerals](#) was cited in U.S. public policy on critical minerals.

Other projects in the works are aimed at analyzing USGS data for reuse. A recent "state of the data" project, consisting of analyzing 165 random data release samples against several established data maturity matrices, aims to determine how mature and FAIR USGS data contained in ScienceBase is, and to document an assessment methodology that is scalable to other bureau data repositories and platforms.

SDM has undertaken initiatives in the past few years to make data easier to work with, access, and publish. USGS Data Tool, a python wrapper around a set of system APIs, is one such tool. USGS Data Tools creates a bridge between various systems (DOI Tool, Pubs Warehouse, BASIS plus, Metadata Parser, SBDR), making data management easier and more intuitive. Other systems have also recently gained connections, such as the SBDR Tool which now contains an option to auto fill information from IPDS.

The USGS Model Catalog is another recent project spearheaded by SDM. The goals for the Catalog are to increase discovery and awareness of scientific models and link models to their related literature, code, data and other resources. The Model Catalog effort is informing practices that will allow the latest information for models to be dynamically updated. CDI is currently assisting by gathering input from modelers across the Bureau - contact Leslie Hsu at lhsu@usgs.gov for more information.

So, in sum, what are your data doing?:

1. Your data are contributing to USGS successes in open and FAIR data
2. Citations to your data are being tracked
3. Data are connected between ScienceBase and your publication in Pubs Warehouse
4. Your data are accessible to the scientific community!

For more information, see the [USGS Data Management Website](#).

Field Data Collection using NGA's free and open source Mobile Awareness GEOINT Environment (MAGE) and MapCache Mobile Apps 338px- Ben Foster and Justin Connelly, National Geospatial-Intelligence Agency

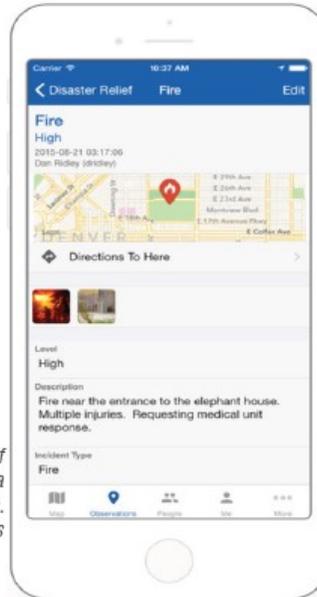
What is MAGE?



MAGE is a mobile app designed to provide situational awareness while operating in a low-bandwidth or disconnected environment. Ideal for disaster response personnel, urban search and rescue teams, and others that need interactive content to assist with their tasks, MAGE users can take geotagged pictures, audio, or video, and use these media files to annotate entries and instantly share that information with others on the team. Users can access MAGE from a variety of mobile devices. Disclaimer: MAGE is an open-source product that NGA has enabled for use throughout the geospatial enterprise.



(Left) Example content of the MAGE app displayed via Android. Disaster relief data. Source: GEOINT Services



(Right) Example content of the MAGE app displayed via iPhone. Disaster relief data. Source: GEOINT Services

What can users do with MAGE?

Using a mobile device's GPS, MAGE can record a user's position in real time. These locations can be automatically shared with other members of the user's team. When disconnected from the network, MAGE will use local data layers to continue providing relevant geospatial information to teammates.

Overview of the MAGE mobile application.

The MAGE and MapCache mobile applications are open source field data collection apps designed to reach a wide audience, even those without GIS experience. The GitHub repository for these applications can be found here: <http://ngageoint.github.io/MAGE/>

Please see the meeting [recording](#) for a live demonstration. Some highlights from the demonstration:

1. The mobile application allows users to upload geo-located video and photo observations.
2. The mobile app also allows creation of lines and polygons.
3. Information added will be visible to other team members on the app.
4. The web application has a similar interface, but with more robust features.

To join the CDI Event on MAGE for govt employees: (1) Request an account from NGA's Protected Internet Exchange (PIX): <https://www.pixtoday.net> (use government email address for account registration) Once you have an account on PiX, send an email to help@pixtoday.net and request to be added to the "MAGE USGS CDI" event.

--
[All CDI Blog Posts](#)