GEOSPATIAL DATA AS A SERVICE USE CASES

A Report of the National Geospatial Advisory Committee
December 2018

Introduction

This document was developed to complement the material in the NGAC report on Geospatial Data as a Service: A vital investment in American Enterprise and to provide real-world use cases of Data as a Service (DaaS) in action.

DaaS opens new possibilities and innovation by improving access to and use of data. Users and producers across government, industry, open source, education, and the sciences are turning to DaaS to meet their huge appetite for information, exponential storage needs, data archiving and to minimize data duplication and reduce costs. To illustrate these points, we examined four case studies that provide more in-depth use cases from across Federal and local agencies and the private sector. These case studies showcase real world scenarios with large, distributed data sources to support a broad range of end user needs. The case studies presented are not intended to be an exhaustive list of examples of DaaS in use; there are many other relevant examples and programs of DaaS across the stakeholder community.

To demonstrate use of DaaS for local governments and emergency response, we highlight the work of the Missouri Task Force One (MO-TF1) and Boone County Fire Protection District (BCFPD). MO-TF1 and BCFPD are using public cloud for basemap creation, data access, and delivery to enable a common operating picture of critical data during deployments. (Appendix 1 - Data as a Service & Cloud Computing: Disaster Response Just-in-Time Basemap Creation for Deployments).

To demonstrate the applicability of DaaS for serving a broad range of state-level geospatial data, we highlight the Montana State Library's use of Cloud based GIS DaaS web map services to complement their traditional, downloadable data and growing collection of web applications (Appendix 2 - Montana Spatial Data Infrastructure).

To demonstrate applications of DaaS for large volume data, we highlight the NOAA Big Data Project. NOAA has been experimenting with the Big Data Project to improve accessibility and usability of environmental data. This project currently exceeds 20 terabytes of data captured per day (from a variety of sources, including satellites and weather stations) and is expected to exceed 80 petabytes of archived data per year by 2021 (Appendix 3 - NOAA Big Data Project). To further highlight the extreme size of many geospatial datasets, NASA now produces 7-20 petabytes a year per project.

To demonstrate commercial use of DaaS, we highlight the private sector 'Living Atlas' platform provided by Esri (Appendix 4 - Living Atlas of the World – The Data You Need) that serves data to a variety of customers and applications – many of which are embedded into mobile devices such as smartphones and tablets.

Use Case 1. Data as a Service & Cloud Computing: Disaster Response Just-in-Time Basemap Creation for Deployments

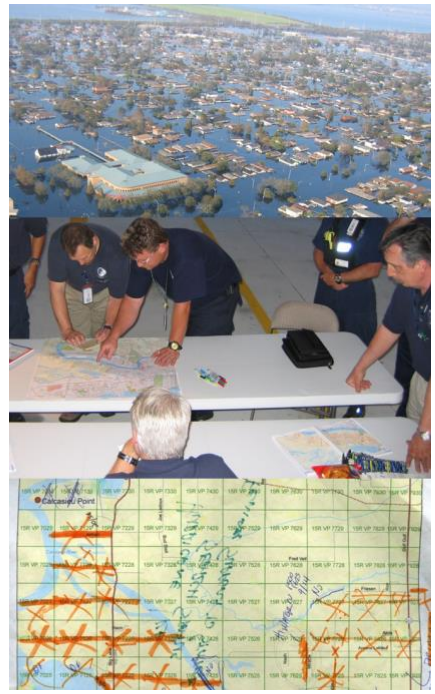
Jason Warzinik (Boone County, MO Fire District / Missouri Task Force 1)

Missouri Task Force One (MO-TF1) is managed by the Boone County Fire Protection District and is one of 28 FEMA Urban Search and Rescue teams in the United States. The Task Force is designed to assist the local emergency agencies facing a disaster response both in-state and out-of-state. Many of its missions have involved a wide area search component, including Hurricanes Katrina, Sandy, and Harvey; flooding in Colorado; the Joplin, MO tornado; and, the Oso, WA landslide. During the first few days of a deployment, the Task Force operates in areas with heavy damage to structures, infrastructure, and utilities, and must plan to operate nearly completely offline or with severely limited communications and data bandwidth. As a result, even though current geospatial information in the field is critical to effectively and efficiently plan, perform, and document searches, the Task Force continues to operate with little intelligence and maps for the response area. They often have to resort to finding local commercial paper maps for planning, performing, and documenting searches.

The Task Force can deploy anywhere in the U.S. within a few hours of the call to mobilize and needs to be ready to begin field operations with the aid of street or city block map scale imagery and base layers. Traditionally, the Task Force relied on FTP to update local GIS data caches, which quickly caused a data management nightmare of outdated basic vectors and USGS 7.5m quadrangle topographic map rasters for basemaps.

With a mission to modernize the basemap and mapping capabilities, the Task Force is building a hybrid GIS system utilizing an Esri ArcGIS based mobile GIS situational awareness vehicle and Amazon's Elastic Compute Cloud (EC2) public cloud for basemap creation, data access, and delivery to enable a common operating picture of critical data during deployments. When completed, the system will provide planning and field operations staff a near real-time geospatial wide area search and rescue situational awareness with the ability to operate in disconnected mode when needed or in connected mode using the GIS vehicle's Wi-Fi, LTE, or satellite communications when available.

One of the challenges with working in a disconnected environment is that any basemaps need to be downloaded to the devices. Esri offers download-enabled basemaps on their ArcGIS Online platform, but



these exports fail due to the large (regional) map extents required for large area deployments which is often the case for flooding, earthquakes, and hurricanes. To mitigate this issue, the group experimented with a different solution during a recent training exercise. A combined AWS Microsoft Windows EC2 instance and cloud storage client was used to directly access the 140,000 GB AWS Big/Open NAIP s3 bucket as if it was a local hard drive. As a result, within minutes, a mosaicked dataset containing the 189 GBs of imagery for the State of Missouri was created and ready to use within Esri ArcGIS Pro. Since the EC2 instance and s3 NAIP bucket are in the same AWS region, the AWS s3 sourced imagery performance for on-screen map refreshes and multi-threaded export tile package geoprocessing tasks was acceptable and came at no cost. The resulting 50-500 MB basemap tile package was downloaded from AWS to the mobile GIS situational awareness vehicle and then deployed to the field mapping devices.

The use of cloud computing, such as the AWS EC2 instances that are accessible from anywhere with an Internet connection, will also enable members of the Task Force that didn't deploy to the disaster to provide GIS and mapping support. The high network bandwidth will enable the Task Force to quickly create updated basemaps during the disaster response using other imagery and GIS services from agencies such as FEMA and NOAA and other content providers as they become available during a disaster response. This allows the teams on the ground, and those at home, to work efficiently for effective disaster response.

Use Case 2. Data as a Service: Montana Spatial Data Infrastructure

Jennie Stapp (Montana State Library)

It has long been a stated goal for the Montana State Library to make the entire Montana Spatial Data Infrastructure (MSDI) accessible through web services. Constantly changing resources – primarily in terms of funding, technical support, server software capabilities – have impacted how this has been implemented, but the overall recognition of the need to provide data as a service (DaaS) access to the MSDI has remained constant for at least 10 years.

Data from 11 of the 15 MSDI Datasets are currently served through a centralized, cloud based, ArcGIS Server Environment. The remaining datasets (Geology, Soils, Elevation, and Climate) have unique stewardship models or data structure challenges that have prevented the Library from making those available from this location. It is worth noting that the Geology data is available as a web service directly from the Montana Bureau of Mines and Geology and that some portions of the remaining three datasets are available through ArcGIS Online.

When efforts to provide DaaS access to the MSDI began in the mid-2000s, web services were viewed as an improvement over the well-established data download model. The vision was for web services to replace data downloads, simplifying access to users and reducing problems resulting from users downloading data that quickly becomes stale. Several factors weren't anticipated in realizing that vision:

- The slow pace of broadband buildout in Montana;
- The growth in the complexity of many MSDI datasets;
- The very diverse needs of the user community when it comes to data formats, access tools and formats, and data currency requirements;
- The limitations of certain web map service standards and the growth (fragmentation) in web map service types – presumably driven at least in part by those limitations;
- Additional complications and complexity that have come about as a result of the increased use and availability of online mapping platforms such as ArcGIS Online.

As the DaaS offerings have matured, and issues like those listed above have come to light, the Library no longer views web map services as a replacement for data download options. Instead, Web Map Services are viewed as complementary to downloadable data as well as a growing collection of web applications. Interest in the use of these services to support web applications (both internal and external) have pushed for the provisioning of these services using two different spatial reference systems, Web Mercator and Montana State Plan (Figure 1). And finally, while attempts to provide a single general-purpose service for each MSDI theme are made, there are many cases where themes are split across multiple services.

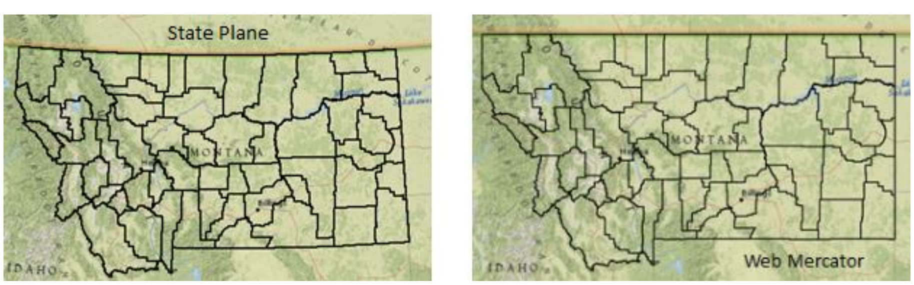


Figure 1 - The Montana State Library provides web services in two different spatial reference systems.

Moving forward there are a few challenges facing the current MSDI DaaS vision –

- Should two spatial reference systems be supported? If not, should the focus be on supporting a single spatial reference system, or select on a theme-by-theme basis based on usage patterns of one versus the other?
- Should efforts continue to be focused on making all themes available as Esri Map Services or should additional formats be explored, such as Web Feature Service, Web Coverage Services and Web Map Tile Services?
- Should web map services be supported for all themes or should efforts be focused on those which are most heavily used? Different themes currently have dramatically different usage levels with some being hit tens of thousands of times a month and others just hundreds of times a month or less.
- Should efforts be focused on offering more services by teasing out more of the specific theme components into their own dedicated services or should focus be placed on base map services that combine datasets, sometimes (often?) from across themes.
- There needs to be a review, and possibly update of the design of existing services to ensure they are mobile friendly (both in terms of look and feel and performance).

With these issues in mind, it would be beneficial for the Federal Geographic Data Committee (FGDC) to take more of a leadership role in the realm of DaaS for the National Spatial Data Infrastructure (NSDI). It has been encouraging to see Federal agencies investing more in making GIS data available as a service.

Use Case 3. Data as a Service: NOAA Big Data Project

May Yuan (University of Texas – Dallas)

Since October 2015, NOAA has been experimenting with a Big Data Project to improve accessibility and usability of environmental data. This project itself now reaches 20 terabytes (TB) of data daily. The growth of NOAA archives, in general, including atmospheric in-situ observations, geophysical data, ocean data, NEXRAD radar data, and a wide suite of satellite data, has increased to almost 35 petabytes in FY17 and is expected to exceed 80 petabytes by FY21. Along with this growth is the demand for NOAA data access, which increased to over 7 petabytes in FY16. The volume and diversity of NOAA data demand significant time and computing resources to facilitate downloads and analyses.

NOAA Big Data Project on Public Cloud Platforms
(https://ncics.org/data/noaa-big-data-project/)

The NOAA Big Data project is implemented via Cooperative Research and Development Agreements (CRADAs) at no cost to the government. Amazon Web Services, Google Cloud Platform, IBM, Microsoft Corp, and the Open Cloud Consortium have formed NOAA data alliances and provided Infrastructure-as-a-Service (IaaS) to broaden NOAA's data access and foster innovations by allowing computing directly on the data in the cloud. The Big Data Project synergizes industry's expertise in data storage and access expertise and NOAA's data archives and subject matter expertise. The alliances provide full NOAA data access openly and freely to the public and industry but charge use fees for virtual machines on their IaaS. In addition to cloud computing and free open data access, the industry provides data services to commercial entities for application developments, for example, use of weather and climate data in transportation and agriculture. An increase in public access to both data and computing facilities in the cloud enhances opportunities for innovative applications of NOAA data which may otherwise remain untouched in government archives. An effective use case is that the public can now use various apps to access real-time mosaicked NEXRAD radar images to identify the progress of rainstorms or severe weather in relation to one's current location.

Early measures show great promises for the NOAA Big Data Project. For example, requests for NOAA NEXRAD data increased from 18 TB in March 2015 to 95 TB in March 2016, with AWS facilitated two-thirds of data access. The Big Data project continues to reduce NOAA's data workload. In July 2016, NOAA processed only about 25% of the NEXRAD data requests. Since then, the vast majority of NOAA data access is now through AWS, Google Cloud Platform, IBM's NOAA Earth Systems Data Portal, and the Environmental Data Commons at Open Commons Consortium.

Use Case 4. Data as a Service: Living Atlas of the World

Pat Cummens (Esri, Inc.)

The days of satisfying data sharing requests through FTP downloads and the once revolutionary 'Clip and Ship' tools have passed. In an always connected society one has come to expect instant access, response and results in everything from connecting with friends through messaging apps to ordering groceries and having them appear within the hour. There are similar expectations of not only accessing data but quickly getting answers delivered. However, data are different than groceries: the insights extracted from big data are only as good as the sources from which they are derived. It is important to work with authoritative data. This ensures that data driven decisions are indeed driven by the right data.

The geospatial community is making progress in this direction by implementing data-as-a-service strategies, creating dynamic services of data, accessible over the web on demand, eliminating the need to download data. Esri's Living Atlas, a collection of geographic information from around the globe that includes maps, apps, and data layers, is a successful example of this modern technical approach to meet expectations and demands of the causal and professional users of geospatial data. Continually growing and evolving, this resource features items such as weather, traffic, elevation, demographics, oceans, urban systems, and imagery, that impact people's lives. This can be combined with an end user's own data or other data contained in the Living Atlas to quickly and efficiently garner results.

Data as a Service - a Proven Approach in Living Atlas

Curating and making available a diverse collection of web services improves the efficiency of applying GIS data and processes to support decision-making. Employing this build-once-use-many-times philosophy has proven to be beneficial to a vast number of users, as demonstrated in the table on the left. To streamline access and discovery, the growing collection of data and applications is organized by general themes to facilitate browsing categories but also includes several filtering options to quickly zero in on the data for which one is searching. The general organizing themes include: basemaps, people, environment, infrastructure and imagery.

Living Atlas Statistics.

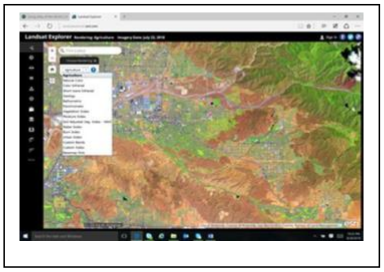
- **Content:** 7000+ Items (maps, layers, apps, etc.-including 3,500+ live services)
- **Number of Requests**: over 30 billion per month (over 1 billion per weekday)
- **Number of contributors**: several thousand
- **Most popular content**: imagery, basemaps, geocoding, live feeds (including traffic, spikes with events like fires and hurricanes), elevation
- **Top Contributors:** NOAA, European Environmental Agency, David Rumsey Collection, British Columbia, Centers for Disease Control and Prevention, US Environmental Protection Agency, USDA Farm Services Agency, UK Ordnance Survey

'Live Feeds' Exemplify Power of Data as a Service

Adopting the technical pattern of web services, which enables data as a service, allows the creation of an ever-changing collection of resources from around the world. In the Living Atlas example, some layers update more than others. Collections updated in an automated manner as soon as the source data is available are "Live Feeds." These are growing in popularity and relied upon by millions of users to provide reliable information for weather, natural disaster, and environmental applications. Live Feeds services include Multispectral Imagery feeds like Landsat & Sentinel. These satellite imagery products provide near real-time situational awareness and are useful for monitoring conditions and changes across Earth's surface. Other examples are disaster feeds for earthquake data, hurricane forecasts, flood maps, and wildfire mapping.

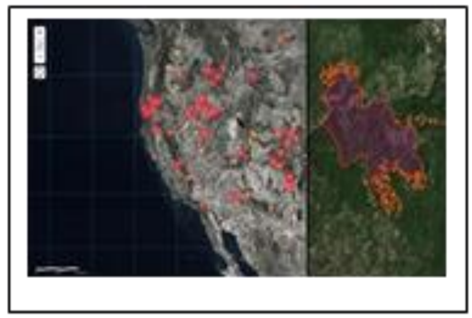
The App Revolution: Moving toward "Answers as a Service'

While making data more usable by providing data as a service, the value of the data and return on investment increases dramatically when the data services are incorporated in ready-to-use applications that instantly connect the data to useful intuitive apps. Making these accessible on mobile devices and through simple web browsers, users are more readily able to derive insights from the data. Examples include:



Landsat Explorer and the Sentinel Explorer are popular web apps that empower users, with the click of a button, to utilize different imagery bands to visualize color infrared, vegetation index, moisture index and a host of other characteristics using the imagery services.

Drought Tracker displays the current week's drought monitor information and graphs the historic information making it easy to monitor trends.



Fire Monitor leverages data feeds from Earth-observing satellite sensors capable of detecting the infrared energy released by fires, identifying hotspots and burned areas, using both thermal characteristics and visible appearance.

Of course, these ready to use apps will not answer all the questions, but since the data is available as a service, it can also be incorporated into other new apps to address specific needs. Data as a service coupled with ready to use apps delivers the best of both worlds.

Meeting the Growing Demands for Services – The Living Atlas Testing the Path

One of the great values of the shift to dynamic web services is that it taps into more than just data – as demonstrated in the Living Atlas. It includes data services, live feeds, web maps, apps, Story Maps, and other informational resources. This is a model for how to curate an immense and rapidly growing collection of ready-to-use maps, apps, imagery, and geo-referenced data for the world. This online collection of authoritative content, together with the growing application of services, and adoption of 'everything as a service' pattern, is having a profound impact on the way people interact with data and apply spatial thinking and analytical tools. The Living Atlas of the World is connecting people with information and tools they need to solve problems.

NOTE: *This set of use cases was developed to complement the NGAC paper, "Geospatial Data as a Service: A Vital Investment in American Enterprise", which was adopted by the NGAC in September 2018.*