MINNESOTA IT SERVICES

Minnesota Geospatial Commons

State of Minnesota — Minnesota IT Services

CATEGORY:
Information Communications Technology (ICT) Innovations

CONTACT:
Cambray Crozier
Director of Communications
cambray.crozier@state.mn.us
(O) 651-201-1063
(C) 651-367-4616

INITIATION DATE:
January 2012

END DATE: March 2015

Executive Summary

From transportation to environmental conservation, all manner of public, private, and non-profit organizations benefit from readily accessible geographic information. This information lets us analyze and interpret maps to help us understand relationships, patterns and trends. Minnesota citizens benefit from the use of this information in myriad ways, and from a history of publishers making it available on open data distribution sites. However, those sites were supported by several different organizations, leading to disjointed data systems and processes for distributing and using the information.

Enter the Minnesota Geospatial Commons, a collaborative public website where geospatial technologists can access hundreds of geospatial resources composed of data, metadata, maps, services, and applications. This site provides users a single place to search for resources contributed by a variety of publishers, including state, regional, county, academic, and nonprofit organizations. Operated by the Minnesota Geospatial Information Office (MnGeo), a program of Minnesota IT Services (MNIT), it averages approximately 40,000 page views a month.

The primary audience for the Commons is geographic information systems (GIS) professionals, who otherwise spend a high percentage of their time gathering data for any particular project. The Commons leverages the open source data management system CKAN for search and discovery, and an FTP site for data distribution. A custom-developed "broker" interface allows publishers to stage new data sets within a standardized data system, validate the contents, and approve them for sharing with state agencies or the public. After initial approval, most data updates are processed automatically, streamlining operations for all publishers.

The data from this site recently played a role in the successful management of an avian influenza outbreak across Minnesota. MnGeo has also documented significant usage by the private sector. When an engineering firm or real estate company needs access to spatial data, they've discovered they can obtain the latest and greatest from one website. They can quickly find information, assess the suitability of a resource, and download data in a variety of formats. Standards and best practices on documentation and metadata have ensured that the site contains useful, understandable data.

By working across boundaries, the Commons supports MNIT's strategic goal of improving government operations through collaboration and shared technology. When an authoritative data source provides spatial data to be published on the Commons, it is providing that data to all state agencies and the public via one process. The success of the Commons is evident in its growth, starting with just nine publishers and less than 200 data resources when it launched in March of 2015, and growing to over 600 resources from 25 publishers within two years. It can be found at https://gisdata.mn.gov/.

Exemplar

Driven by the example of data.gov and open data policies at the federal level, MNIT made several decisions early on to ensure the success of this innovative open data solution. For discovery and cataloging, we chose the open source CKAN software, based on its success at data.gov and its large developer community. We also built a bridge between the existing "Geospatial Data Resource System," or GDRS, developed by the Minnesota Department of Natural Resources (DNR). This data organization and replication scheme was already in use by several Minnesota state agencies, so those agencies did not need to reorganize their internal data systems. MNIT also established minimum requirements, best practices and business rules for geospatial metadata, which defined the quality of shared data.

These decisions laid the framework to develop the "GeoBroker" for validating resources from a GDRS before publication. While this middleware system leverages advanced techniques, such as the CKAN API, perhaps the most innovative aspect of the Commons is the governance decisions that were made early on about the quality of data that would be published. By only allowing well-documented data from authoritative sources, we ensured that current and future users would not be burdened by sifting through an overabundance of irrelevant or unusable data. Visitors to the Commons are confident that the data they seek will be simple to find, easy to evaluate for their purposes, and quickly downloadable.

Easy access to open data supports several policy goals set by Minnesota state leaders, such as Governor Dayton's focus on water quality. For example, Minnesota's buffer law establishes new perennial vegetation buffers of up to 50 feet along rivers, streams and ditches that will help filter out phosphorus, nitrogen and sediment. The DNR produces and maintains a map of public waters and public ditch systems that require permanent vegetation buffers, and used the Commons to release this buffer protection map and data in July 2016. Consistent access to this data is now guiding the implementation of the law by landowners with the aid of state and local government agencies.

Since the Commons launch in 2015, nine Minnesota Counties have either passed a "free and open data policy" or implemented open data rules by practice. Currently, eight counties are leveraging the Commons for their data distribution mechanism, meaning they do not have to invest in their own technological solution to meet their goals of distributing data to the public. This capability has been a critical selling point to county commissioners when considering the adoption of an open data policy.

Concept

Before this project, Minnesota had rich sources of geographic information, but they were scattered throughout many organizations and levels of government. A person looking for this information needed to know what entity produced it, make a request, then wait for that request to be processed – usually by manual means. The concept for creating the Geospatial Commons was simple: create one place for publishers and users to easily share and access geospatial data. When the original vision was articulated, input was gathered from more than 500 survey responses from the public sector (39%), education (24%), and private sector (22%). This helped project sponsors ensure that the final solution would represent a wide variety of stakeholders, not just state agencies or government staff.

The project to create the Commons was initiated with a dedicated project manager and technical staff that supported state agency data publishers. Team members were also frequent users of data, making them ideally suited to business analysis and development roles. The team compared three open source platforms using 12 evaluation criteria that included the health of their developer community. The final choice was CKAN, currently with more than 17,000 contributions from 158 developers on GitHub.

In the fall of 2013, a pre-release demonstration for stakeholders identified work needed, but also validated the choice of CKAN. From there, the team built the custom GeoBroker to manage transactions between the GDRS and CKAN, completing the first phase in June 2014, and a second final phase in March 2015. By the end of 2015, data was consolidated from three major data publishers: The DNR, MnGeo, and MetroGIS (funded and administered by the Metropolitan Council).

The project team used agile development methodologies, with input from statewide stakeholders after each major release. The team identified and prioritized requirements through user stories that were mapped to functionality and achieved during two-week sprints. A shared code repository, and centrally provided development, test, and production environments helped control the scope of the project and drive success by focusing on the critical requirements.

The GeoBroker and CKAN each underwent an application security assessment, and the systems are patched on a bimonthly basis. While most users of the site browse anonymously, all traffic is forced into HTTPS. Accessibility improvements to the metadata are being made, and we are working toward a more responsive and accessible design.

MnGeo publishes regular news articles, actively seeks and cultivates publishing partnerships, and onboards new publishers on a regular basis. We also engage in continuous operational improvements, such as providing GDRS node administrators a way to “unsubscribe” from datasets that might be useful for the public, but redundant for state agencies. Such improvements mitigate costs while continuing to improve the appeal of the site and its support systems.

The Commons is currently operating on an annual budget of approximately $60,000. It was built with an estimated $600,000 of staff time over two years. Expensive custom development was limited by using existing GDRS nodes and leveraging the work of CKAN developers and data.gov. In this way, the project team demonstrated how to create a solution to a long-standing problem, despite having mostly part-time staff and no official budget for hardware or software.

Significance

Today, Minnesota citizens in need of geospatial resources about their state can go directly to one place for search and discovery — the Minnesota Geospatial Commons. The robust search function and ability to download resources in a multitude of formats greatly reduces the time spent acquiring information.

Governance is established and maintained by the Minnesota Geospatial Advisory Council. Our governing principles have ensured that it is unique from other open data platforms that may suffer from too little

oversight or an overwhelming number of resources. The Commons emphasizes quality over quantity, holding publishers to basic metadata standards and expectations:

- Data must be free and open, covering at least part of Minnesota.
- Data must be published in either geospatial or tabular formats that can be spatially enabled.
- Each resource must have sufficient documentation to determine if they are fit for a particular use.
- Publishers must be organizations, not individuals.

This level of governance not only makes the site attractive for users, but inevitably leads to time savings for publishers as well, who have to spend less time answering questions about their data. Growth has not suffered as a result; since its creation, 25 organizations (including 12 state agencies) have published their information via the Commons. Three major original data discovery sites were decommissioned, focusing the state's geospatial efforts into one space instead of many redundant sites. Merging these publishers onto a single platform sent a clear signal to Minnesota geospatial professionals, and established the Geospatial Commons as a pivotal data source for their work.

Long-term success of the Commons depends on growth in authoritative publishers, particularly from local government agencies who are stewards of key data layers such as land ownership, addresses, and street centerlines. As more of those governments adopt open data policies and share their data with the public via the Commons, it becomes easier to aggregate that data in a standardized way. Continuous improvement includes creating additional methods of "harvesting" data from local government sites to the Commons, and in turn, publishing our metadata records in a manner that can be harvested by data.gov.

The Commons reflects the values of Minnesota IT Services by partnering across organizational boundaries in order to transform the way state agencies work. This enterprise strategy uses solid governance and prioritization to develop enterprise technologies that reduce duplication—a key point in the Minnesota IT Services Master Plan.

Data distributed on the Commons supports Governor Dayton's Cleaner, Safer, Healthier Communities agenda by identifying opportunities to clean up past pollution. For example, the Minnesota Pollution Control Agency identifies closed landfill facilities and hazardous waste, solid waste, and other remediation opportunities from a large agency database and publishes that data to the Commons.

The Commons meets NASCIO priorities of consolidating services and data resource, using Enterprise IT Governance to empower partnerships and inter-jurisdictional collaboration, and provides a platform for data management and analytics.

Finally, the Commons helps meet MnGeo's legislative mandate for coordinated geospatial data in state agencies and the broader geospatial community.

Impact

By incorporating local government publishing partners as well as state agencies, the Commons has transformed the operations of both users and publishers of geospatial information in Minnesota. As concluded by Minnesota's MetroGIS in their research on free and open data, the true value of data is derived from its use. Making authoritative data easily discoverable in a single place for Minnesota ensures that its usage is maximized, providing an exponential return on investment for the Commons and the publishers who use it to distribute their data.

The Geospatial Commons project produced a modern, next-generation site. With increased transparency and access to data in a single location, it reduces duplication of effort and redundant delivery mechanisms. Traffic averages approximately 40,000 page views per month over the past two years. In 2016, the site received over 488,000 page views (326,000 unique) with a bounce rate of 35%, and 71,634 successful downloads. The amount of resources available on the site is a powerful measure of success, with steady growth over time, from the very first seven publishers and less than 50 resources to 25 publishers and 647 total resources published today.

Tangible benefits are being experienced:

- With consistent access to statewide foundational data layers, adjacent public safety answering points can more effectively execute joint powers agreements, and environmental agencies can work with landowners on best management practices. Before the Commons was implemented, planners at the Minnesota Department of Transportation spent an estimated 80-120 hours per project searching for and obtaining data from others. Using the Commons likely cuts that time in half, resulting in a $600,000 annual cost savings. (200 projects per year, 80 hours per project, $75 per hour.)

- If seven different state agencies independently seek out one data set from each of the 87 counties in Minnesota, it would require an estimated 2,436 hours of search and retrieval time per county to obtain a license, have an attorney review and then obtain the data. Therefore, harvesting local government data into the Commons should save the equivalent of more than one full-time employee.

- In 2016, the Census Bureau was looking for specific geospatial information from 22 counties. Fifteen of those counties have requested their information be made accessible through the Geospatial Commons, saving time for the Census Bureau, and saving time for individual county staff who will not have to fulfill the data requests.

The true impact of the Commons is perhaps best conveyed in the words of those that benefit from it:

Education and Academic Research:

"As large consumers of data, student and faculty at the University of Minnesota regularly use the Commons for their learning, research and outreach projects. The Commons provides efficient access to a

huge amount of data, allowing more time to be spent on turning data into valuable discoveries. Having free and open geospatial data from the Geospatial Commons is critical to meeting the mission of the University of Minnesota." —Len Kne, GISP, U-Spatial Associate Director, Research Computing, Office of the Vice President for Research, University of Minnesota

Transforming government:

"This website is a shining example of interagency cooperation and transparent, responsive government. We at the Council are proud to be part of this collaborative resource. We expect having an efficient way to search and access vast amounts of data will be a catalyst for innovation…" —Mark Kotz, GIS Manager, Metropolitan Council

"One Minnesota":

"We are excited to be one of the first counties publishing data on the Commons... Now, we will simply direct people to the Commons, and avoid having to maintain our own website and data repository." —Randy Knippel, GIS Manager, Dakota County

"MN Geospatial Commons has proven an invaluable resource for many of Pro-West's rural county clients. These GIS applications are vital tools to enable rural counties to become more connected, more engaged with citizens, and more efficient organizations." —Jennifer Ward, GISP | GIS Consultant Pro-West & Associates Inc., Geographic Information System Specialists

Citizens and Businesses:

"NorthstarMLS leverages publicly available data in the Commons … to improve the quality and accuracy of Real Estate listings in our service area. [The] Commons allows NorthstarMLS to provide some very advanced and intricate search capabilities, [including] things like identifying the service area for an individual school, or finding properties within a calculated commute distance to work, within 1000' of a bus line, or on a good bass-fishing lake. The Commons has been an indispensable tool for us and allows us to continue to serve the residents of Minnesota with highly accurate Real Estate services." —Curtis Carlson, GIS Coordinator, NorthstarMLS

Environment:

"As a non-profit organization, having access to current and comprehensive public data via the Geospatial Commons is imperative to our mission of advocating for the sustainability of our common resources. Our current and future work relies largely on the continued growth of this irreplaceable spatial data repository resource." —Andra Bontrager, Minnesota Center for Environmental Advocacy