ASSOCIATION OF CANADIAN MAP LIBRARIES AND ARCHIVES
BULLETIN

Geospatial Data and Software Reviews

University of Manitoba GISHub: ESRI Site License Integration using ArcGIS Online, Hub and Enterprise

Meg Miller & Mullai Manickavalli
University of Manitoba

Abstract

This review provides an outline of the solution the University of Manitoba Libraries has implemented to integrate their ESRI Educational Site License. In looking at the tools available the project came to encompass the following:

1. Semi-automated management and integration of UM ESRI site license using campus authentication methods
2. Discovery and access point for proprietary and open researcher data
3. Secure local environment for active-use geospatial datasets using ArcGIS Enterprise

The following discusses the software specifics, use cases, and lessons learned in a Canadian academic library context.

Tags

ESRI, ArcGIS Online, ArcGIS Hub, ArcGIS Enterprise, authentication, data, repository, academic library, proprietary license, OCAP, data management

Background

The following examines the integration of the University of Manitoba's ESRI Educational Site License into what is referred to as UM GISHub. The major impetus for this project was the current push for research to contain a formal data management plan and there was a gap in campus infrastructure needed to support it. Campus Libraries put a plan forward to the University IT Advisory Committee with a request for funding to support a local geospatial data repository. While the project began as a solution to address the lack of storage and management options for active-use geospatial data on campus, upon stepping back, it evolved into a solution for the semi-automated management and integration of the university's site license, a discovery and access point for proprietary and open researcher data as well as that secure local environment piece.

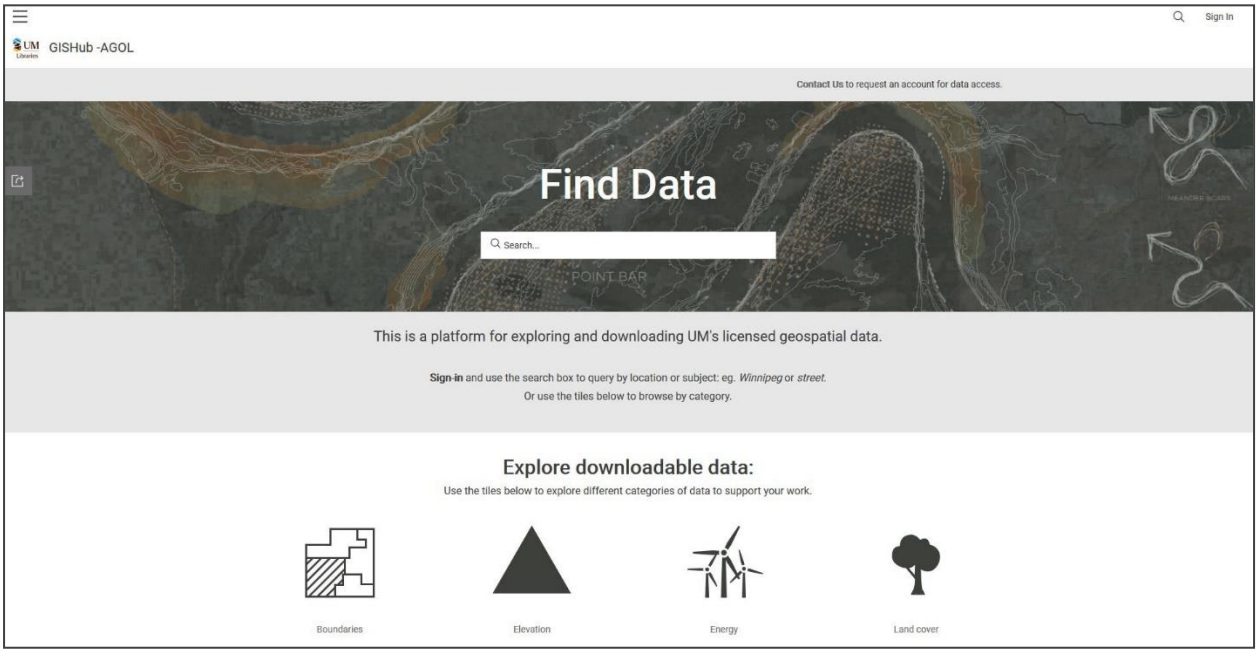

Figure 1: Landing page for University of Manitoba GIS Hub

Project Goals

Account Management

Previously access and accounts were being managed manually, which as campus usage grew became unsustainable. Tying into campus authentication allows users to sign-in with their university credentials, it also ensures that when they depart from the institution software access is removed. Some legacy accounts had been created using partial university credentials. Before the campus authentication gateway could be opened these user accounts and content needed to be manually migrated with proper system mapping, which was a very labour intensive process.

Secure Local Data Environment

The locally hosted Enterprise portion of the Hub allows researchers to honor their data management plans and keep data on university servers while making use of ESRI analysis tools. Direct connections can be established with ArcGIS Pro or users can also make use of the analysis options available in the virtual environment itself (think a locally hosted version of ArcGIS Online with some additional features and analytics capabilities). Users require institutional credentials to access the system and have the ability to restrict and share their data at many different levels. Additionally, Hub administrators can facilitate collaboration between other organizations with ESRI licenses outside of the university to create secure shared workspaces for cross-institutional work.

Discovery and Access

The final goal of the system was to create a portal for the discovery and access of proprietary data Libraries purchases and, in the future, open researcher data.

By making use of the ability to set up distributed collaboration between ArcGIS Online and Enterprise instances we are able to host proprietary datasets in Enterprise and share them with our ArcGIS Online organizations for our users to link to no matter the platform they are logged into. This gives Libraries a hands-off way of distributing data while staying within the terms of the data sharing agreements we have with different vendors. The system also has the ability to track usage of these datasets which is useful for institutional reporting. This sharing can be done at an organizational level (all members of the UM community) or at a group level (members of a certain lab group).

An ArcGIS Hub page was set up to act as a landing page to tie all these components together. Researchers can search for data by location or keyword or use category tiles to browse (municipal and DMTI datasets have been tagged) for data in which they are interested. Links are provided to popular open government portals, as well as to the organizational tools and documentation available.

Technology

The University of Manitoba is part of a provincial consortium for purchasing their institutional site license (MEGIC). As such, Libraries access their licenses from the institutional representative who is part of a separate department (Faculty of Environment, Earth, and Resources). The diagram below illustrates the software components as well as faculty owners.

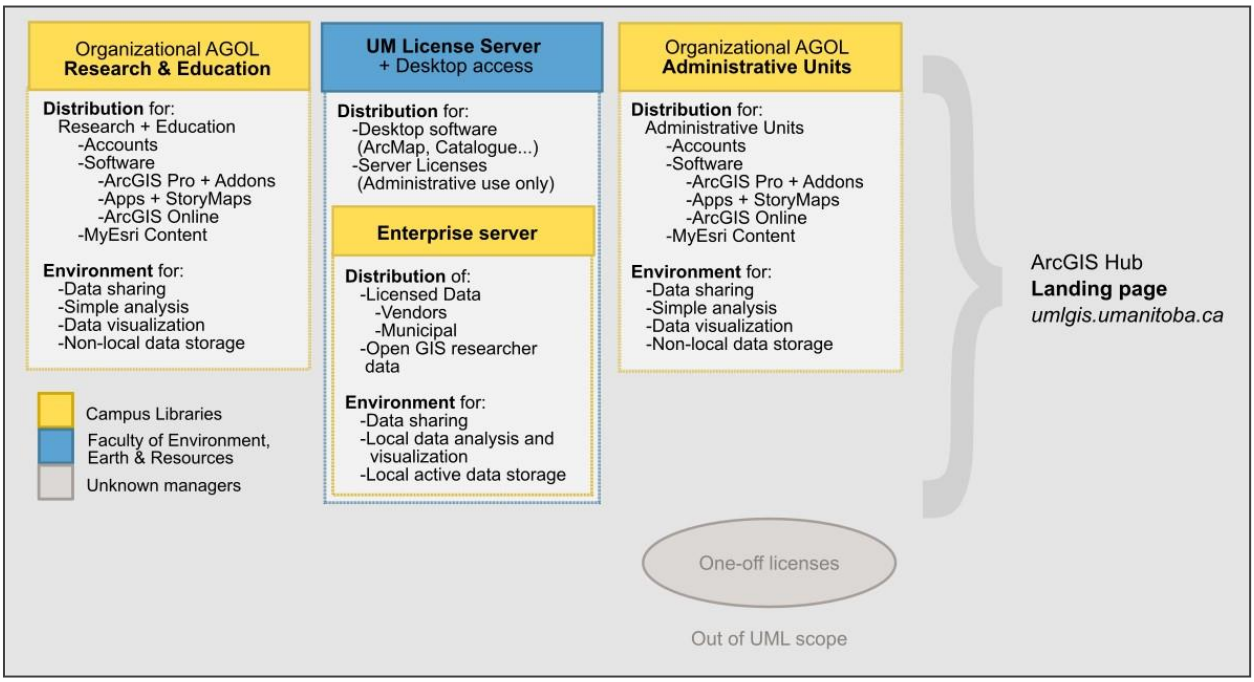

Figure 2: Software and managers at the University of Manitoba

To summarize, Libraries is responsible for all software/systems that are tied to ArcGIS Online accounts as well as the Enterprise Server. Users can access these systems by authenticating with university credentials. To tie everything together ArcGIS Hub was used to create a landing page for the service.

Implementation details

UM-ESRI Integration

Our end users are UM patrons and MEGIC consortium users. These two groups access the GIS Hub through Desktop apps, Web browsers or the Field Data Collector tools. Based on the methods they use to access; institutional users are authenticated by LDAP or SAML and other MEGIC users are provided with manually created individual logins.

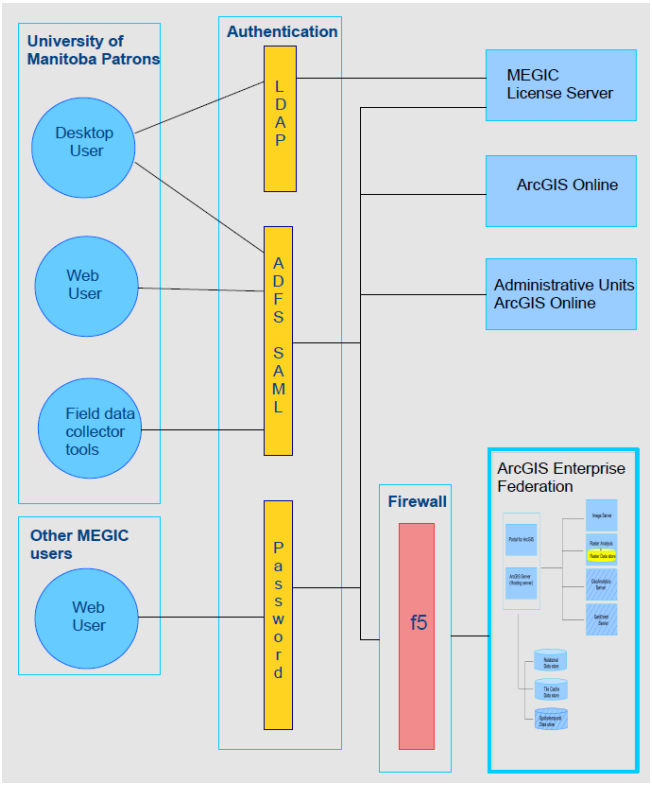

Figure 3: University of Manitoba- ESRI integration

When a university patron signs in for the first time, they would use the SAML method, this creates an account automatically using university credentials and by default, they are assigned a Viewer role. The user then must contact Libraries to obtain any other role other than the default Viewer role in the system. Additionally, with the SAML method login the user login session is preserved between ArcGIS Online and the Enterprise Federation. This also facilitates a single sign-on experience with other UM provided web services. Implementing this process has freed

up a significant portion of staff time which was previously spent on creating and managing accounts.

ArcGIS Enterprise Federation:

Our current ArcGIS Enterprise architecture consists of one Portal for ArcGIS, one GIS Server which is also the hosting server, an Image Server, Raster Server, Raster Datastore, Relational Datastore and one Tile Cache Datastore. All components are at ArcGIS Enterprise version 10.7.1 for Linux. RHEL7 is the current operating system and the F5 campus firewall.

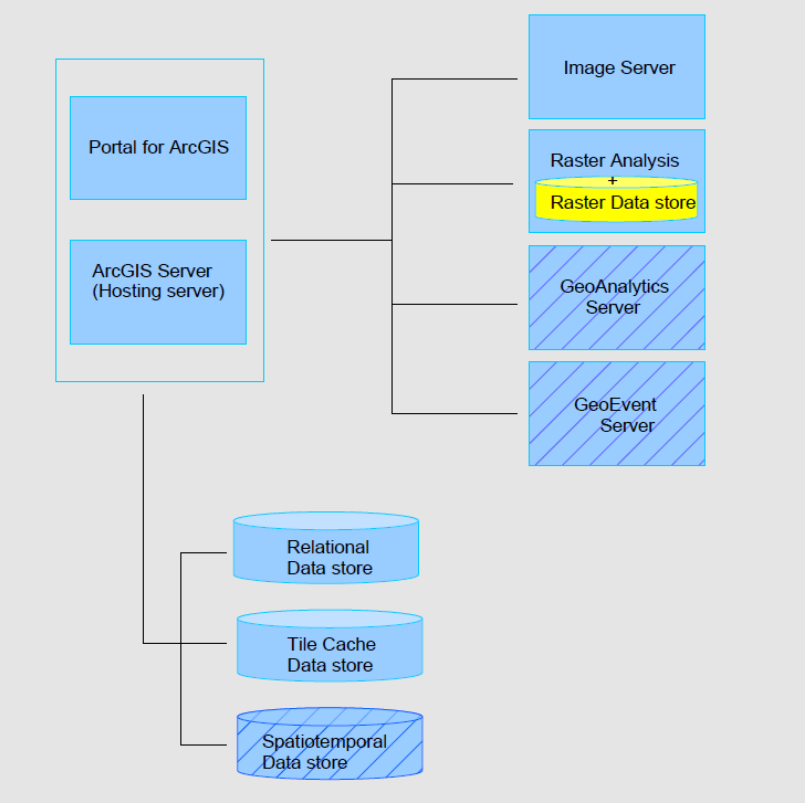

Figure 4: ArcGIS Enterprise federation

Except for the Raster Datastore, all the above components are installed on separate virtual machines. In terms of resource allocation, the hosting server which is the workhorse of the federation is given 16 CPUs and 64GB of RAM, whereas other virtual machines are given 8 CPUs and 32GB of RAM and 50GB of local disk space each. All VMs share a common network volume of 5TB for the uploaded files.

ArcGIS Enterprise Components	Virtual Machine Resource Allocation
ArcGIS Server (GIS Server + Hosting Server)	16 CPU, 64Gb Ram, 50Gb local storage, 5Tb of shared storage volume.
Portal for ArcGIS	8 CPU, 32Gb Ram, 50Gb local storage, 5Tb of shared storage volume.
Image server	8 CPU, 32Gb Ram, 50Gb local storage, 5Tb of shared storage volume.
Raster server	8 CPU, 32Gb Ram, 50Gb local storage, 5Tb of shared storage volume.
Relational Datastore	8 CPU, 32Gb Ram, 50Gb local storage, 5Tb of shared storage volume.
Tile Cache Datastore	8 CPU, 32Gb Ram, 50Gb local storage, 5Tb of shared storage volume.

For automating the installation, deployment, and configuration process, ESRI recommends using Chef Infra. In terms of how the UM computing environment is set up, it ended up being necessary to take a hybrid approach to this recommendation. Chef Infra cookbooks were referenced for deploying the major software like ArcGIS packages, web adaptors, licenses, iptables and Tomcat. For other components like SSL certificates, LDAP & SAML login, Hostname aliases etc. individual configuration was necessary. Also, instead of deploying all components with a single Chef Infra automation script, one VM at a time was targeted.

For the most part, ESRI documentation was helpful, however, putting it all together as a federation and making them work in our campus environment was a challenge. ESRI support was able to give some guidance when approached for clarification along the way. Good knowledge of the campus IT environment is an asset if the reader is considering a similar setup at their local institution.

Users

The Hub is being piloted as a connected whole over the summer semester. The four groups were identified based on their needs:

1. Faculty to student data sharing
2. Campus administrative users
3. Researcher data collection/sharing
4. New GIS users

The Faculty to Student group uses a workflow that is very similar to the one Libraries use to share proprietary data with the rest of campus, but instead is for a small subset of users for a finite amount of time. These users identified having adequate documentation and training for themselves and their students as being the main priorities. The hands-on portion for this type of user for Libraries moving forward will be to set up the data sharing group from Enterprise, the rest can be administered by the Faculty member.

The administrative unit group came from Campus Planning. They require a dynamic connection to update their inventories as well as had more interest in reporting functionality than other user groups. Functionality and training were identified as being critical for this group to adopt the platform. Aside from initial training and set- up this group will not require much ongoing library support.

The team selected to trial the Hub for data collection purposes brought out what the author considers to be the greatest weakness of the system. This specific group is involved in many

indigenous knowledge projects, and while the Hub is appropriate for some of this work, this proprietary system does not allow for the First Nations Principles of OCAP to be adhered to in terms of much of the collection, ownership and control process. The ethics of this extends beyond this group to any researcher who is conducting work with external (especially marginalized) groups. These users identified having adequate documentation and training for themselves and their students as being their main priorities. Libraries will be heavily involved providing regular training sessions.

The final pilot group is from Nursing and were new to GIS as a whole. They will require the most support out of all groups but have been extremely useful in terms of being able to articulate what starting documentation is necessary, and the overall user experience of navigating the system.

Each pilot group is working through a small project over the course of the semester and changes and suggestions are being implemented as they come up. These tasks cover the range of basic activities that are expected to take place in the GIS Hub. In the coming semester the system will be opened to all university accounts, with feedback being solicited regularly.

Scope

As this is a new piece of infrastructure on campus, one of the main priorities has been to articulate and communicate the scope of the project to our stakeholders. This is broken down in terms of capacity, licensing, and policy.

Discussing the GISHub in terms of computing capacity allows researchers to more easily understand if it is the correct solution for their work. We are framing it as a service for active research and teaching, not a platform for long-term data hosting and archiving.

In terms of policy, discussion and documentation are centered around data management planning and sharing. While ESRI vocabulary includes the word ‘open’, having a dataset in the GISHub with this option selected does not mean the researcher is satisfying tri-agency open data requirements. Additionally, it is up to the researcher to put appropriate sharing restrictions on their research data, University of Manitoba Libraries staff does not do housekeeping in this area.

The final scope note is ensuring that we communicate the licensing terms of the software to our users. When manually setting up accounts we could ensure researchers were aware use of the product was for research and education only (no profit generating activities). When campus authentication gateway is open to UM members in the fall, having this documentation in a variety of places including with the central IT branch will ensure we are doing our due diligence to not violate the terms of our license.

Conclusion

There are many future plans associated with the GIS Hub, but current priorities include adding the GeoAnalytics, GeoEvent and Spatiotemporal servers to Enterprise for increased analytics capacity. Also, growing the GIS Hub as a repository for open geospatial research data from the University of Manitoba has been identified as a priority.

For other institutions considering ways of integrating their site license, in our experience main points to consider are:

- *Understand your campus needs*: who are your users, what are they doing? If you only have one program with 30 users who make use of the software, this might not be necessary for you.
- *Have a team member with in-depth knowledge of the campus IT environment*: this is more valuable than any ESRI experience.
- *Start small and build a strong base*: by stepping back and breaking things into smaller components, we cleaned up legacy problems, and were able to demonstrate consistent progress to management.
- *Communication- you are not an island-:* Keeping stakeholders in the loop builds trust and goodwill and improves the chances of adoption.

This project has been a massive undertaking and both authors (Librarian and Systems Analyst) have many other components to their jobs. Taking the time to step back and maintaining good relationships with contacts in central IT and the Faculty of Environment Earth & Resources have been critical to the success of the project.

Meg Miller
GIS & Data Visualization Librarian, University of Manitoba
meg.miller@umanitoba.ca

Mullai Manickavalli
Systems Analyst, University of Manitoba
mullai.manickavalli@umanitoba.ca