MYGEOHUB SEMI-ANNUAL

Coordinates to Move Your Science Forward

December 2021

Inside This Issue

Message from the Principal Investigator

As we end 2021 by having served 12,000 users, we would like to share some of the recent developments on MyGeoHub and hear your input on making our geospatial data framework more usable for your research workflows.

Carol Song, Purdue University

Meet the GABBs Family

MyGeoHub.org is a public science gateway supporting broad geospatial modeling, data analysis and visualization needs of research and education communities through hosting and management of groups, datasets, tools, training materials, and educational contents. Built on HUBzero, it features geospatial data capabilities developed under the GABBs project efforts.

G·A·B·B·S
geospatial data analysis building blocks

GABBs is an NSF-funded project to create a powerful Web-based system that will allow researchers worldwide to manage, curate, share, analyze and visualize geospatial data for purposes ranging from predicting damaging floods to projecting climate change effects on the global population.

Project Spotlight – GABBs 2.0

The GABBs team joins two large NSF projects to harness complex data and accelerate convergence sciences.

In GABBs 2.0, we are developing an extensible geospatial data framework (GeoEDF) to add novel geospatial data capabilities making large scientific and social geospatial datasets directly usable in scientific models and tools. The key building blocks are reusable *data connectors* and *data processors* that implement data acquisition from various repositories using various data access protocols, and a range of domain-agnostic or domain-specific geospatial processing operations, respectively. The cyberinfrastructure team is working with our domain research partners in using GeoEDF to generalize and automate their data-driven workflows to make them more accessible and reusable.

We are excited to be part of two new projects: (1) NSF funded AccelNet GLASSNET project to build a global network of science teams and communities to collaborate in tackling global land and water sustainability challenges; (2) NSF funded HDR Institute I-GUIDE to harness complex geospatial data for convergence research on resilience and sustainability issues at multiple scales. The GABBs team will contribute its cyberinfrastructure capabilities and resources to advance the research and training goals of these projects.

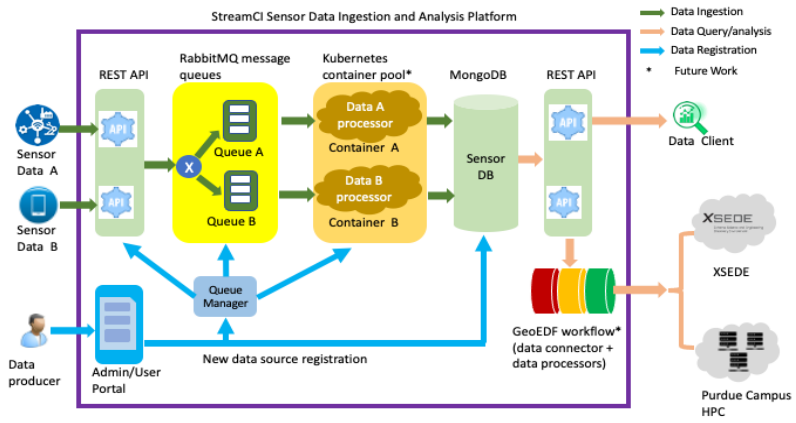
Cyberinfrastructure in Focus

StreamCI *enables researchers to easily collect, process, manage, analyze, and access streaming data from devices and sensors on the cloud*

Developed as part of the GABBs project, ***StreamCI***, a scalable and flexible sensor data collection and analysis platform, enables individual researchers to easily collect, process, store, and access large volumes of streaming data on the cloud.

With the front end deployed on MyGeoHub and the backend on XSEDE JetStream, StreamCI provides (1) a web interface for easy registration of new streaming data sources, (2) real-time data ingestion, (3) processing pipelines, (4) a data query and access API, (5) Read/Write access control, and (6) a data management portal.

The current system design is illustrated below:



Get Involved

Interested in learning more, please contact: **info@gabbs-project.org**

Follow Us on Twitter:

@MyGeoHub_

@gabbs_project

GABBs Project:

mygeohub.org/groups/gabbs/

StreamCI:

streamci.org

The beta version of ***StreamCI*** has been released with its user portal accessible on MyGeoHub. It is being used in multiple research and education projects as shown in the figure below:

- Collect and analyze **crop sensor data** to monitor plant health in real time
- Manage **environmental sensor data** and teach students real world big data analytics in an ecology course
- Ingest **manufacturing sensor data** via MTConnect protocol and access the data for analysis