Geostreaming Data Framework

The Geostreaming Data Framework provides data management capabilities and web application interfaces for the management and visualization of geostreaming data.

To maximize flexibility in supporting heterogeneous data sources, the framework includes four components:

1. A geo-temporal web service API to store and serve the normalized data
2. A geodashboard web application providing web interfaces to visualize, interact and retrieve the data
3. Data parsing software libraries written in Python to normalize the data from different data sources into one common schema
4. Clowder, a web-based data management system to store, curate and analyze raw files and associated metadata.

The four components interact to provide pre-processing, cleaning, and visualization of geospatial earth science time series data such as water health data. The raw data from various sources are ingested into the geo-temporal web service API using a variety of data parsers. The parsers organize raw data into an information model composed of three main entities: sensors, streams, and datapoints. The geo-temporal API web service provides methods to query the ingested data by different software clients, including the geodashboard web application.

Projects currently using and developing the software:

- Great Lakes Monitoring
  - https://greatlakesmonitoring.org/
  - Wiki
- Great Lakes to Gulf Virtual Observatory
  - http://greatlakestogulf.org/
  - Wiki
- Intensively Managed Landscapes Critical Zone Observatory
  - http://data.imlczo.org/geodashboard/
  - Wiki
- TERRA-REF
  - http://terraref.org/

Source code:

- https://opensource.ncsa.illinois.edu/bitbucket/projects/GEOD

Task management:

- https://opensource.ncsa.illinois.edu/jira/projects/GEOD/issues