ONE STORE HAS ALL?

The Backend Story of Managing Geospatial Information
Toward an Easy Discovery
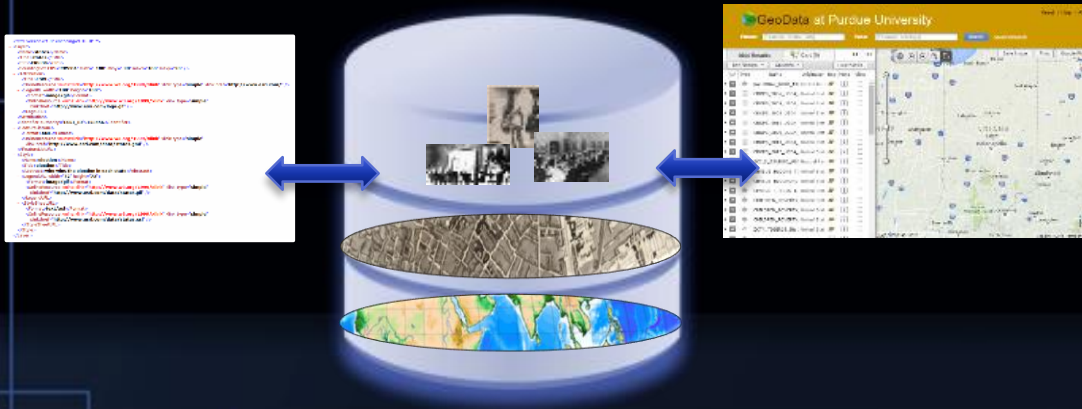
Nicole Kong
Assistant Professor, GIS Specialist
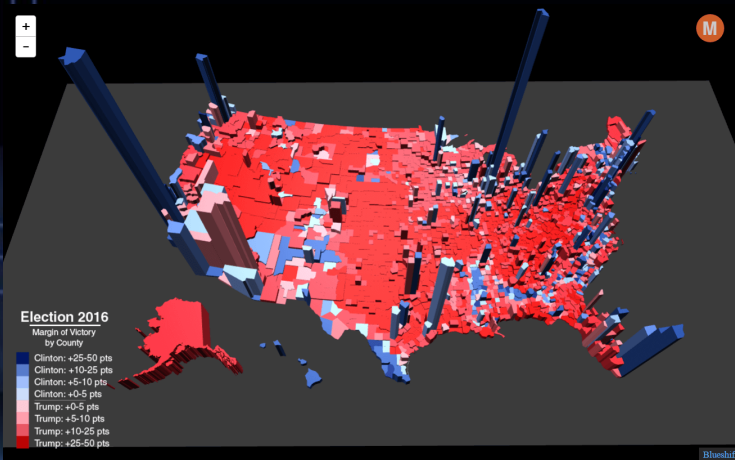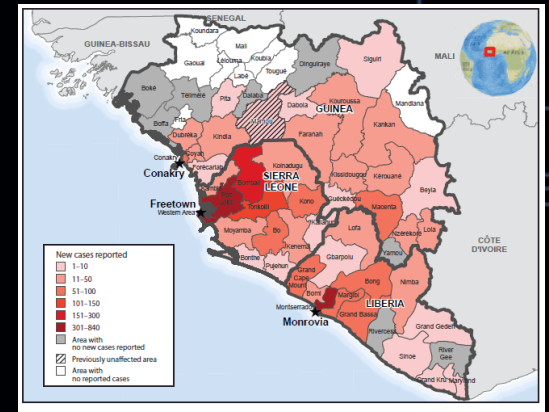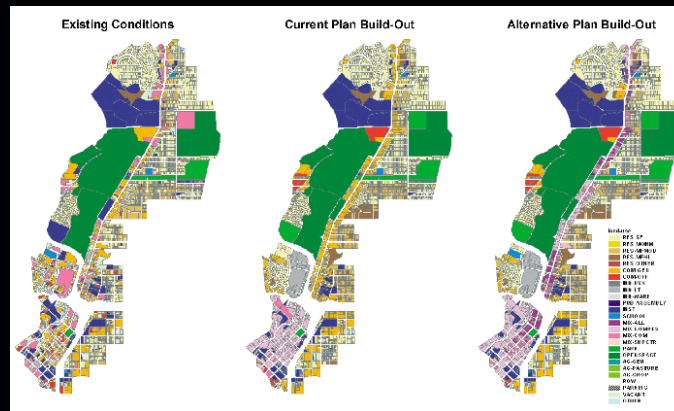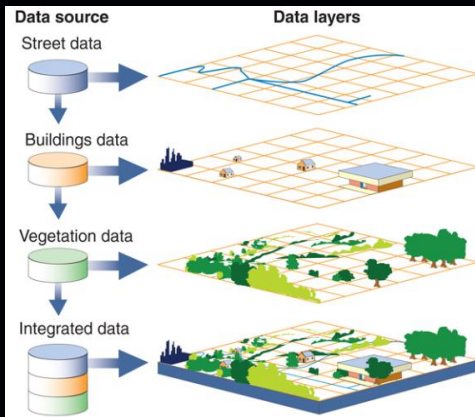Purdue University Libraries

Geospatial Information

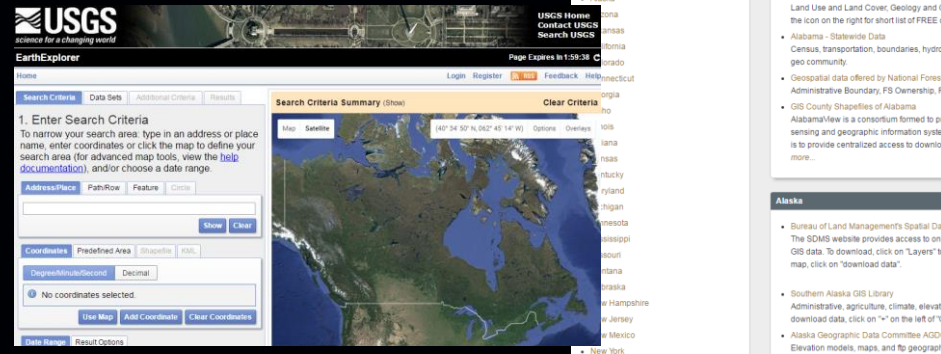- Geospatial data are used across disciplines
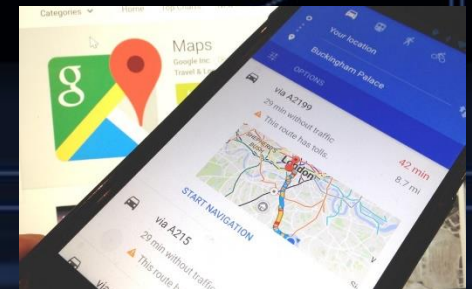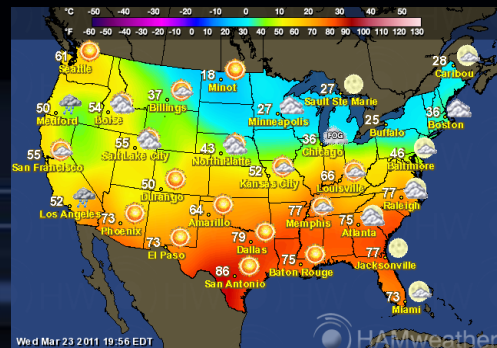
Geospatial Data in Libraries
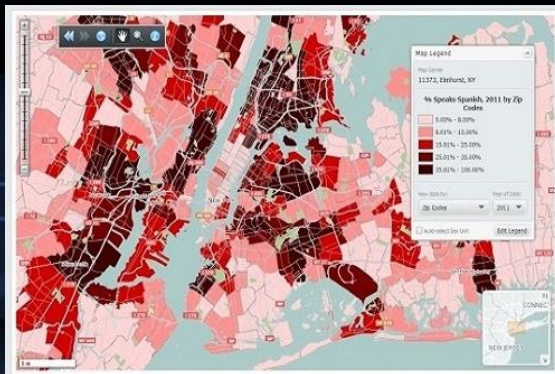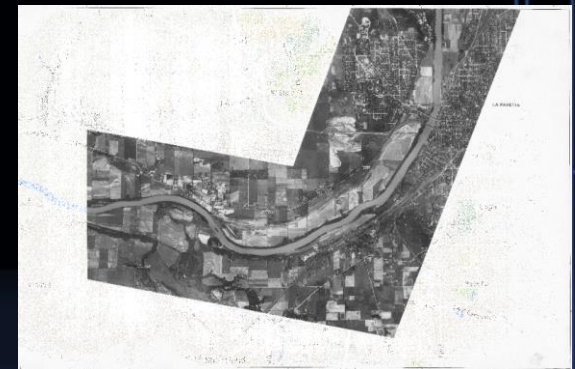

Maps

Digital Maps

Public Domain Data

Proprietary Data

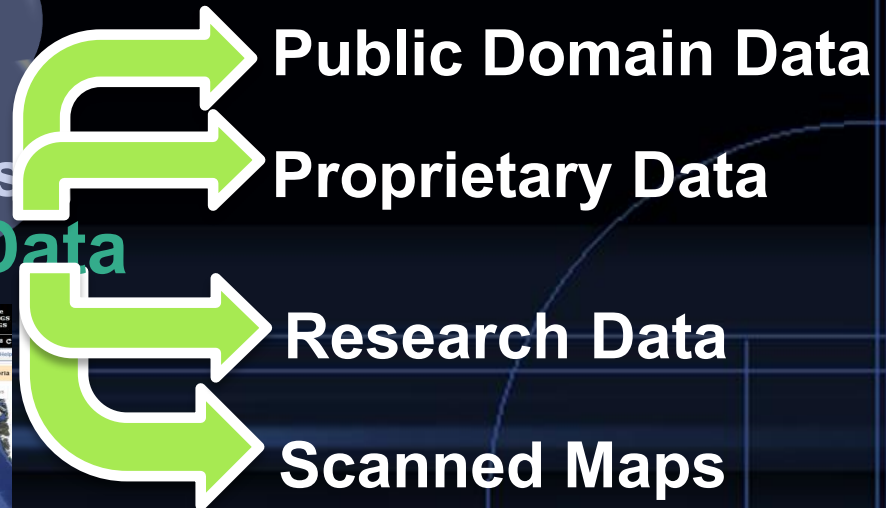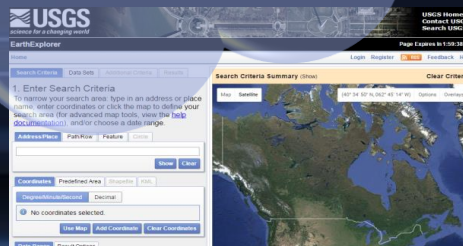Research Data

Scanned Historical Maps

Geodata Portal

- Geodata portal is a discovery layer which allows users to search various geospatial data, and access the information by:

  a web map + data downloading

Public Domain Data

Proprietary Data

Research Data

Scanned Maps

Digital Maps

Geospatial Data

Maps

Geodata Portal in Libraries

Existing efforts:

- Different data providing organizations have their own data portal.

- From the library's world:

    - OpenGeoportal
    - GeoBlacklight

Open Geoportal Interface

GeoBlackLight Interface

Challenges of Building A Portal

- Choose and create the **cyberinfrastructure** Software + Hardware

- Organize geospatial **metadata**

- Manage library's **datasets**

- Interface design and **usability**

Research Question

- How to manage the ever increasing datasets in the library's environment?
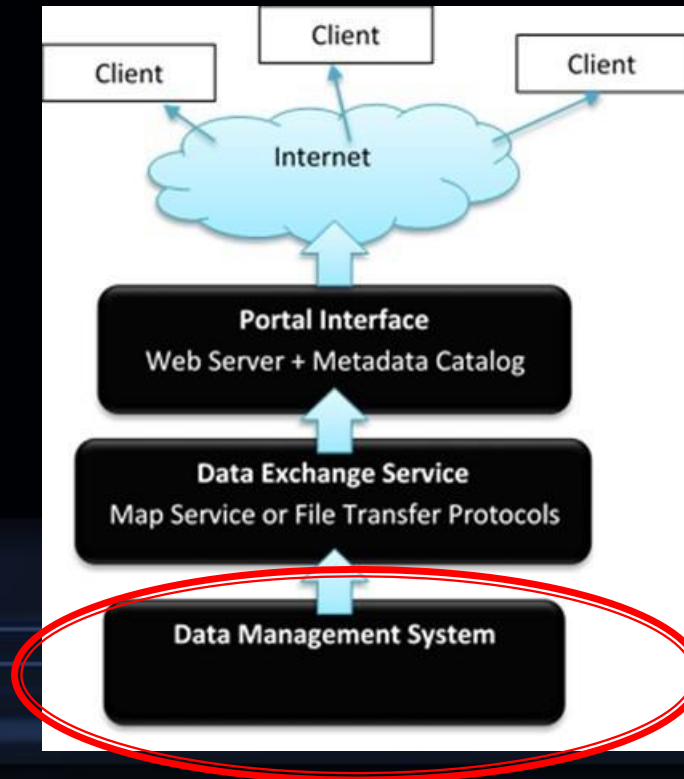
Study Methods

- Project review – data management portion or the existing project
  - Non-library setting
  - Academic library approaches
- Metadata review - records shared by Open Geoportal community

Results

- Designing the Data Management System
- Spatial Data Management Options
- Solutions in Non-Library Projects
- Solutions in Library Projects
- Practices in OpenGeoportal Community

Results

- Designing the Data Management System
- Spatial Data Management Options
- Solutions in Non-Library Projects
- Solutions in Library Projects
- Practices in OpenGeoportal Community

Data Management Design

Needs assessment

- ➤ Data properties
  - ➤ Data formats, data organization, metadata schema, content standard, license agreement, update mechanism.

- ➤ Expected functions
  - ➤ Data download/visualization/mapping

- ➤ Long-term preservation
  - ➤ Data repository – format registry, right management and contracts, metadata documentation, collection management

Results

- Designing the Data Management System
- <u>Spatial Data Management Options</u>
- Solutions in Non-Library Projects
- Solutions in Library Projects
- Practices in OpenGeoportal Community

Spatial Data Management Options

Various options to deliver data via Internet

- OGC (Open Geospatial Consortium) web services

- Spatial DBMS (Database Management System)
    - Oracle 10g Spatial
    - Esri enterprise geodatabase
    - PostGIS
    - Other FOSS: MySQL spatial extension, SQLite

- Other: FTP, HTTP protocols.

Results

- Designing the Data Management System
- Spatial Data Management Options
- <u>Solutions in Non-Library Projects</u>
- Solutions in Library Projects
- Practices in OpenGeoportal Community

Solutions in Non-Library Projects

- ➤ **File management system**
  - ➤ Implemented by a metadata inventory
  - ➤ File could be accessed by FTP, ODBC, HTTP.
- ➤ **Spatial DBMS**
  - ➤ PostGIS + Geoserver
  - ➤ Esri ArcGIS Server + SQL database
- ➤ **Data management system**
  - ➤ Fedora

Results

- Designing the Data Management System
- Spatial Data Management Options
- Solutions in Non-Library Projects
- Solutions in Library Projects
- Practices in OpenGeoportal Community

Solutions in Library Projects

- Spatial DBMS + OGC services
  - OpenGeoportal community
    - Or, direct data download
- Physical map collection can be integrated into geoportal
- Data repository
  - More choice of data formats, data management solutions.
  - GeoHydra

Results

- Designing the Data Management System
- Spatial Data Management Options
- Solutions in Non-Library Projects
- Solutions in Library Projects
- <u>Practices in OpenGeoportal Community</u>

Practices in OGP Community
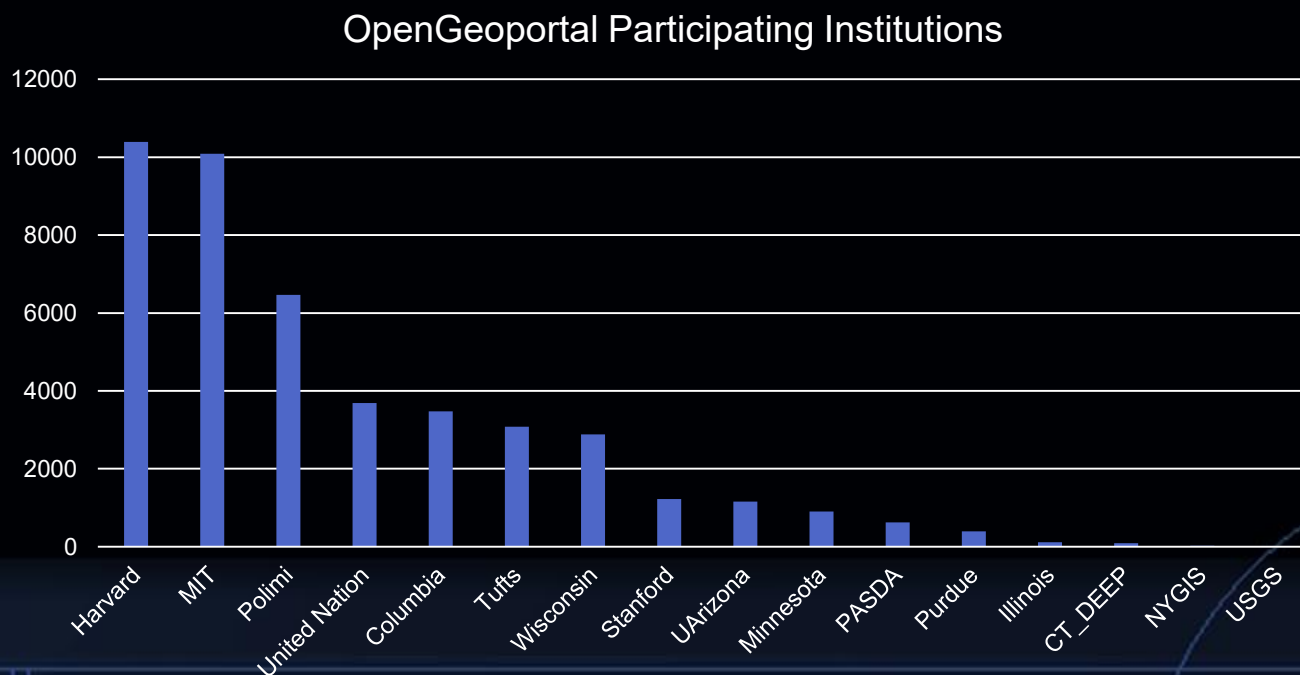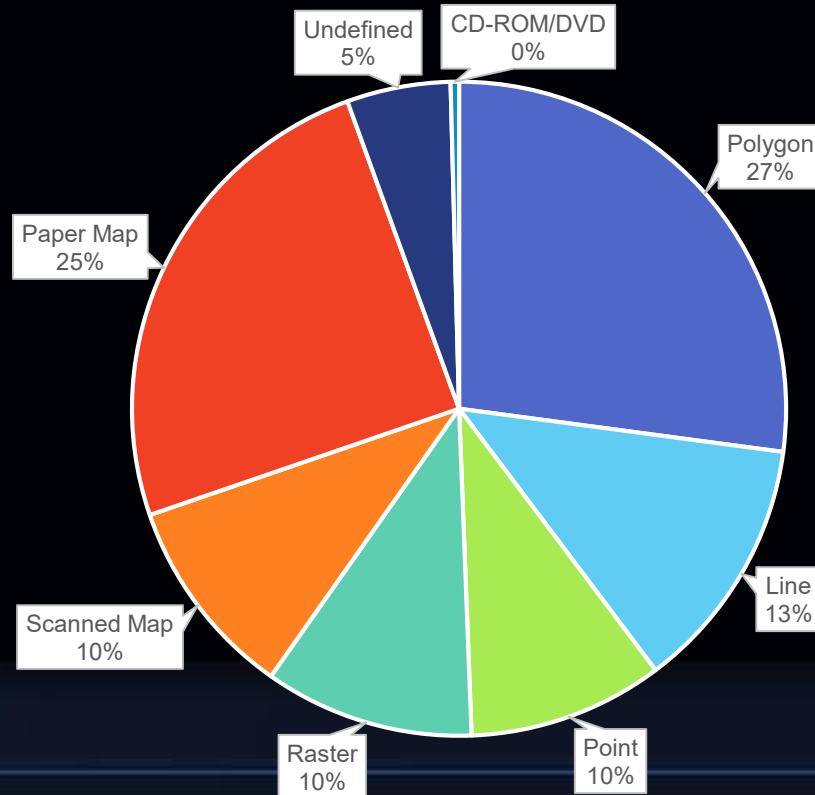
- Metadata shared by OpenGeoportal community.
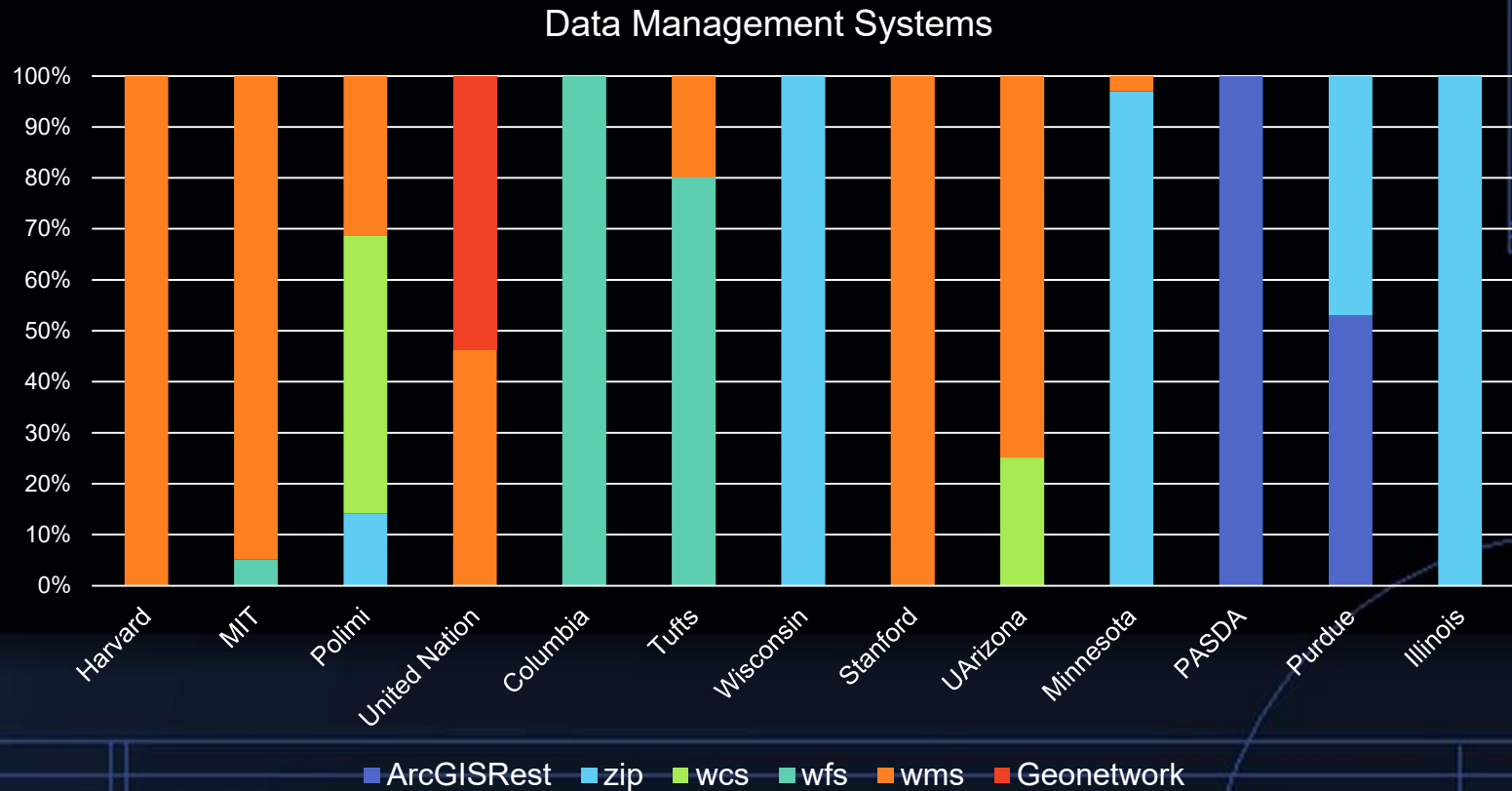- 44,603 records with online link information.

OpenGeoportal Participating Institutions

Practices in OGP Community

Data Type within the OpenGeoportal

Practices in OGP Community

Data Management Systems

Discussion

➤ A well-designed data management system along with good metadata practices are essential for academic libraries to manage the great variety of geospaital datasets.

➤ Not every library data repositories was designed for geospatial data.

➤ It will need a long-term discussion and effort between librarians in order to preserve geospaital datasets