

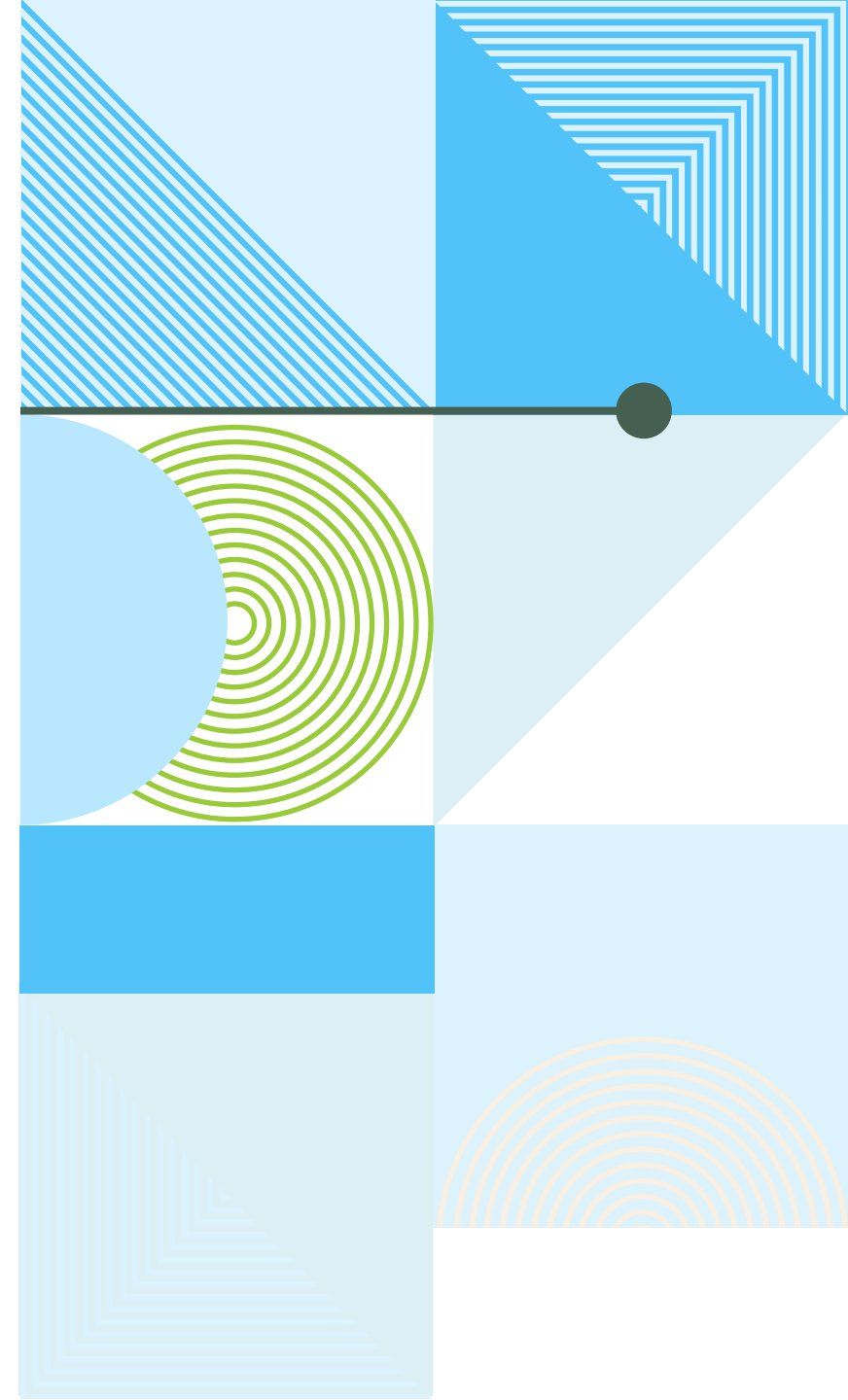
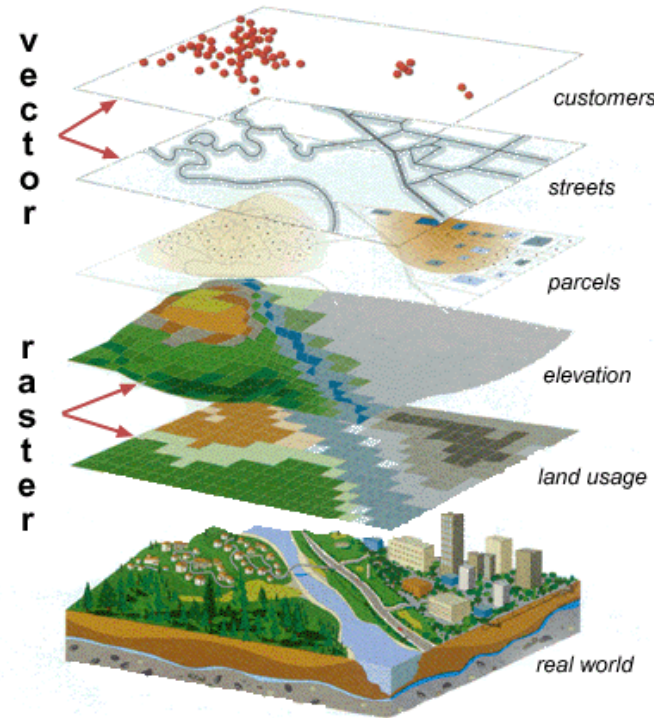


PRESERVING GEOSPATIAL DATA

Meagan A. Snow, PhD
Library of Congress

GEOSPATIAL DATA

- Spatial data can be broadly understood as a geometric feature (a point, line, polygon, or cell) tied to a table of attribute data and a spatial location
- Maps are typically created by *mixing and matching* these spatial data layers and applying symbolization based on values stored in the attribute table
- Spatial analysis allows for operating on and transforming spatial data layers



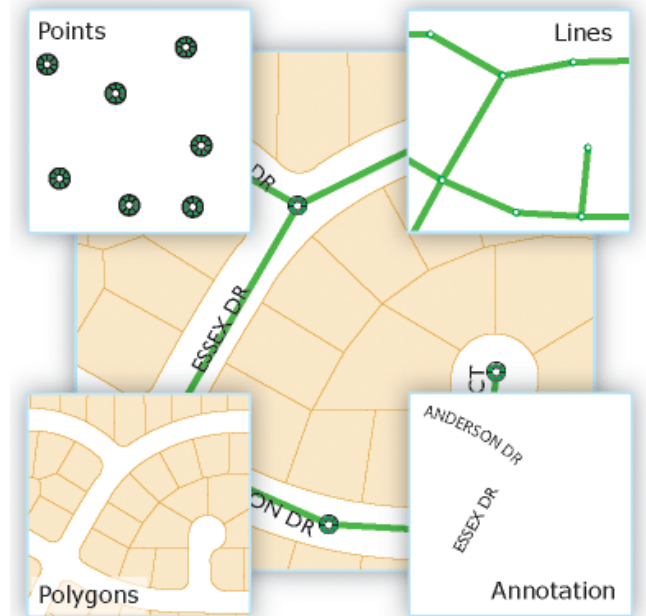
GEOSPATIAL DATA FORMATS

RASTER FORMATS:

- GeoTIFF
- JPEG 2000
- GeoPDF
- NetCDF
- Cloud Raster
- Legacy Formats

VECTOR FORMATS:

- Shapefile
- Geodatabase
- Geopackage
- KML/KMZ
- GML
- OpenStreetMap (OSM)
- Legacy Formats





UPDATING 2009 REPORT

- The new 2023 report is an update to the original Geospatial Technology Report authored by Guy McGarva, Steve Morris and Greg Janée
- In authoring an updated report, much of the content of the 2009 report held true

ENDURING CONSIDERATIONS

DATA STRUCTURE/FORMAT

- Geospatial data formats are diverse, file types are numerous, and not are openly specified
- Transformation between formats requires specialized software/tools and can result in losses
- Geospatial datasets can be large in file size, research projects can generate input and output data

DATA, MAPS, & PROJECTS

- Preservation decisions often need to be made not just for individual data sets (or data layers) but also cartographic outputs and GIS software project files

RIGHTS/LICENSING

- Proprietary software and file formats
- Geospatial data often does not have clear rights or licensing,
- Projects also often use multiple data sources which could have conflicting rights

METADATA

- Geospatial data often lacks metadata or gets detached from original metadata
- Spatial reference information is unique to spatial data and must be properly defined or recorded

NEW CHALLENGES



VOLUME OF DATA

- Real-time monitoring, location-based services, and earth observation is creating huge amounts of data very quickly
- Wide-spread use and adoption of GIS software and spatially-enabled devices means more people are creating spatial data

WEB SERVICES AND MAPS

- Data often continuously updated
- Difficult or impossible to preserve complex interactions across data layers, including complex cartography that changes at scale or with interaction
- Data or maps can now be native to web-based applications or platforms

RESOURCES



RECOMMENDED FORMAT STATEMENT

- Organizes digital file formats by content type, then further groups subject-specific file formats into 'preferred' and 'acceptable' collecting categories based on the assessed long-term sustainability of the format.

SUSTAINABILITY OF DIGITAL FORMATS

- An online resource documenting digital content formats which both identifies emerging file formats, and describes the long-term sustainability of individual file formats

GEOSPATIAL SOFTWARE AND FILE FORMATS WEB ARCHIVE

- Web archive focused on documenting geospatial software and file format documentation, expected to debut in 2024.