Does Web Computing Play a Major Role in Addressing Some of Big Geospatial Data Challenges?

Hassan A. Karimi

Geoinformatics Laboratory
School of Information Sciences
University of Pittsburgh

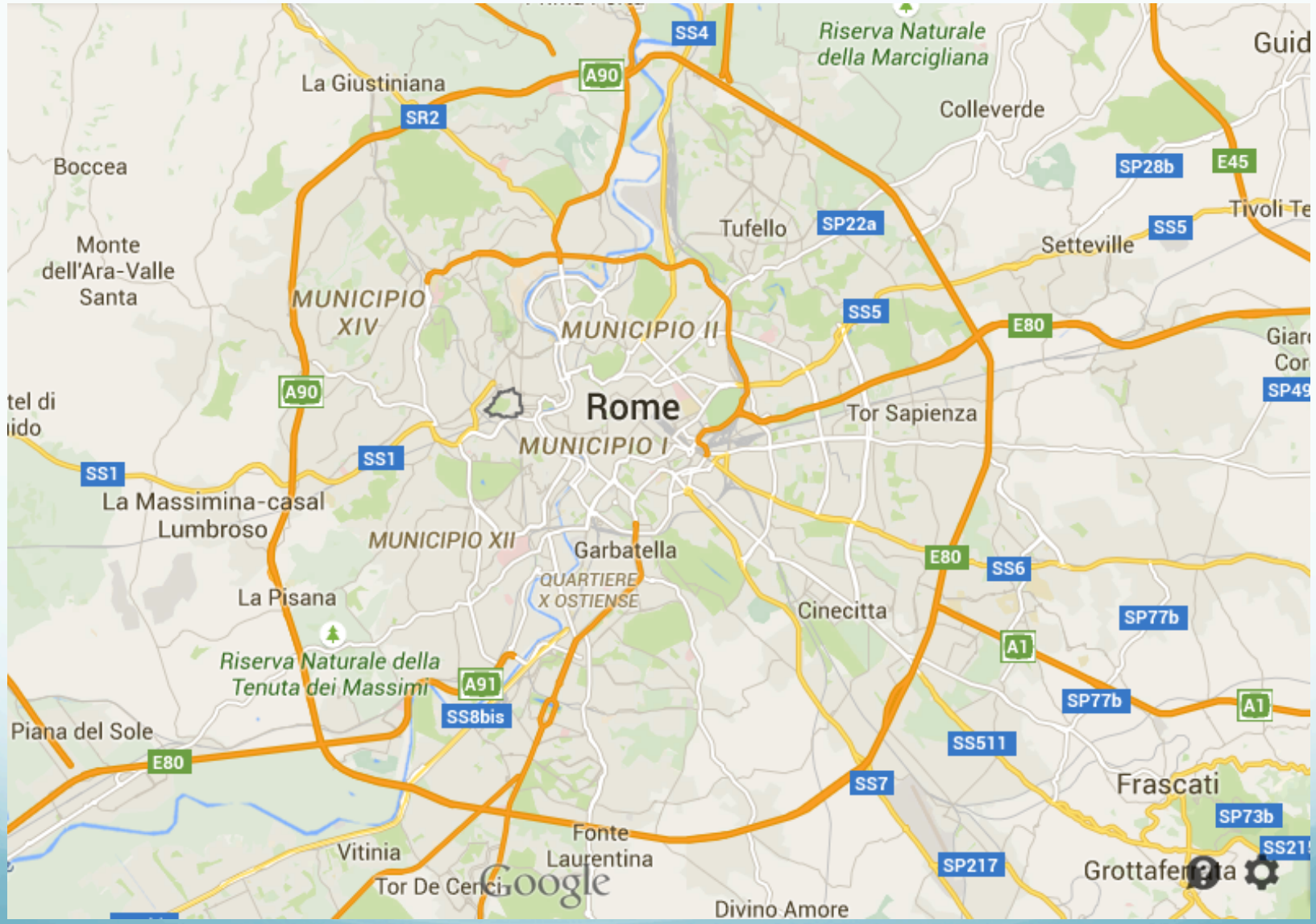
Geospatial Data

- What is geospatial data?
- What are geospatial data models?
- What are some common geospatial data sources?
- What are some geospatial applications?

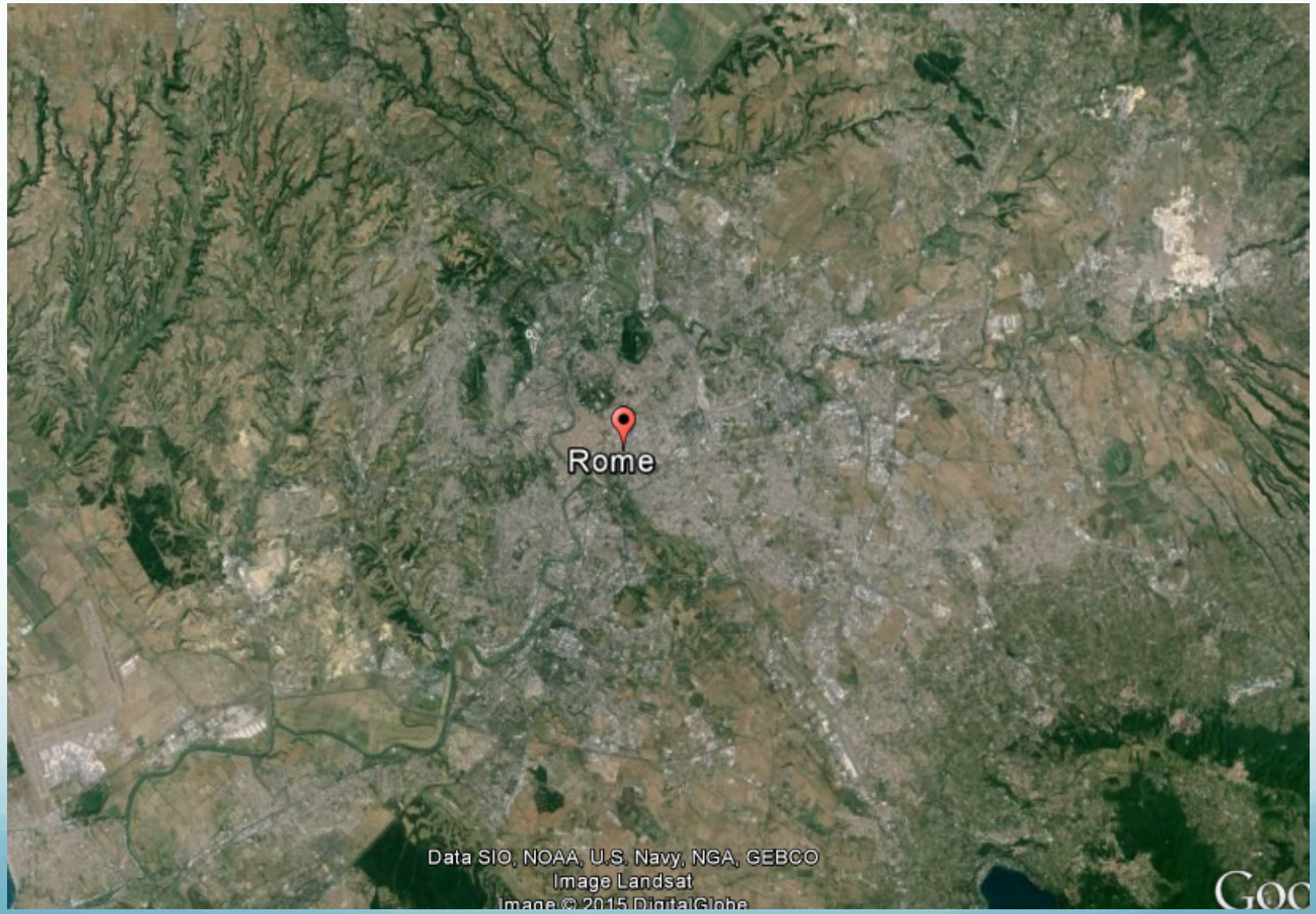
Geospatial Data?

- Location, location, location
- Two models
  - Vector
  - Raster
- Vector model or object model
  - Points, lines, areas (polygons)
- Raster model or field model
  - Pixels

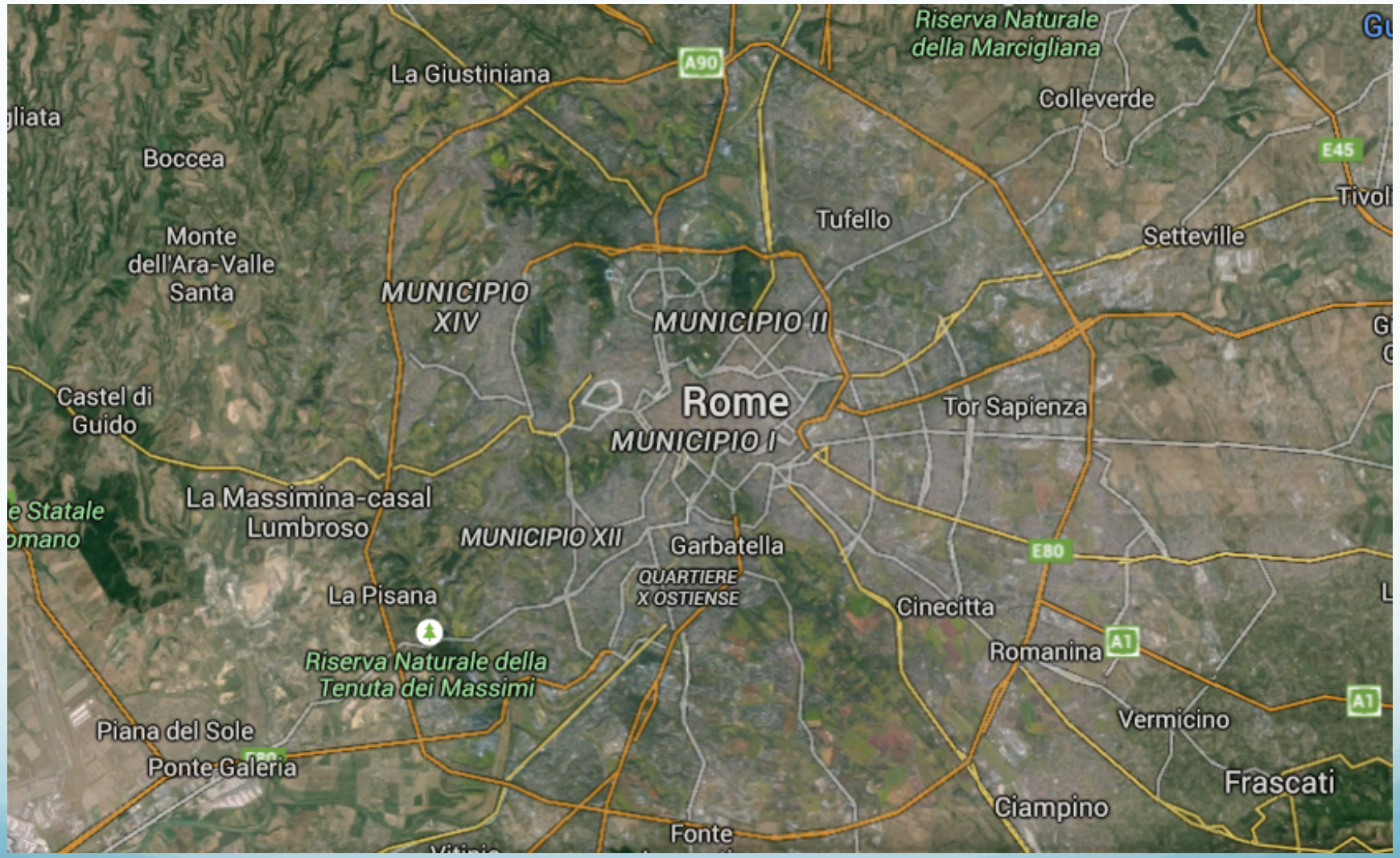
Rome (Google Maps: vector data)



Rome (Google Maps: raster data)



Rome (Google Maps: vector+raster)



Geospatial Data Are Special?

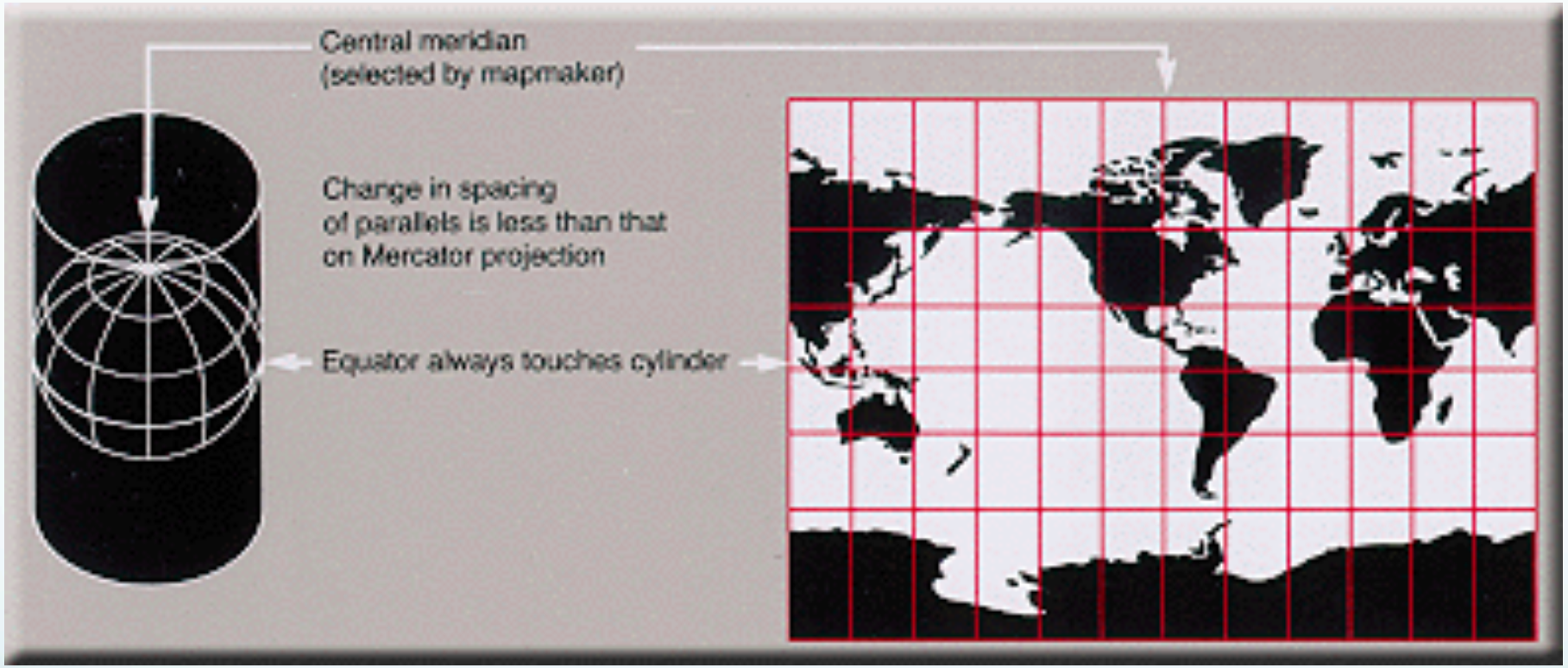
- Geometry
  - Location of an object or field
  - Object may be a point, line, or an area in vector model
    - E.g., latitude/longitude of a point (house)
  - Locations of pixels in raster model
- Coordinate system, projection
- Topology
  - Relationship between objects in a geographic area
    - E.g., a road is adjacent to a house
  - Topology can be explicit in vector DBs
  - Topology is implicit in raster DBs
- Scale
  - Small, medium, large

Multiple-Representation of Objects

- A geospatial database can store an object differently at different scales
  - Small scale → building → a point
  - Large scale → building → an area (polygon)

Spatial Data Manipulation

- GIS supports functions such as:
  - Merging
  - Map overlay
  - Buffering
  - Proximity analysis
  - Point-in-polygon



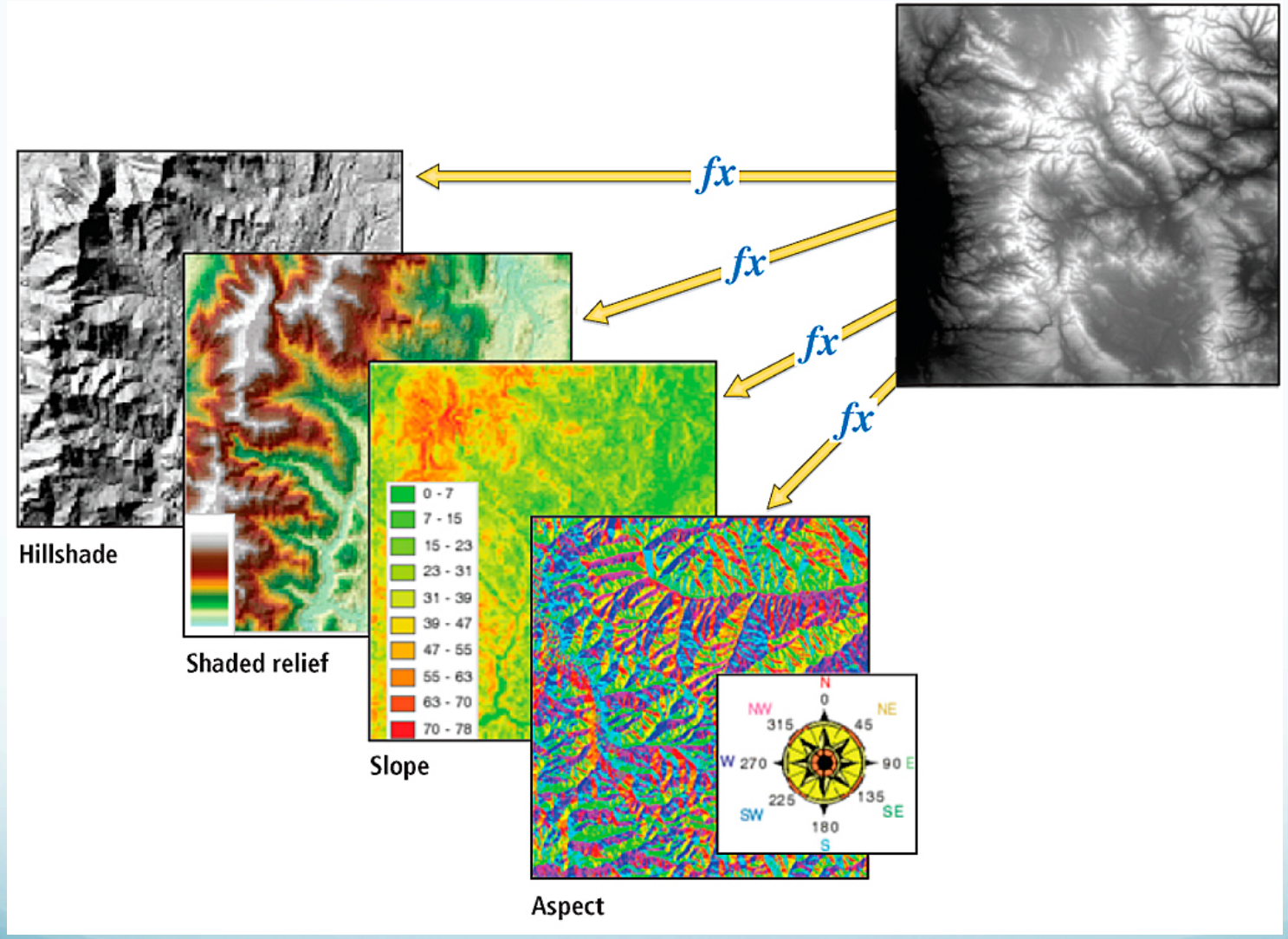
Source: http://en.wikipedia.org/wiki/Map_projection#/media/File:Usgs_map_miller_cylindrical.PNG

Example Computations on Vector Data

- Overlay
  - Polygon intersection
- Buffering
- Proximity analysis
- Network analysis
  - Routing

Example Computations on Raster Data

- Image classification
- Feature extraction
- Object recognition
- Temporal analysis



Source: http://www.esri.com/news/arcuser/0610/graphics/mosaicdataset_5-lg.jpg

Nepal: After Earthquake

Source: https://commons.wikimedia.org/wiki/File:Hires_150508-M-WN441-131A_An_aerial_image_of_damages_after_earthquake_in_Nepal.jpg

Applications

- Environmental engineering
- Natural resources
- Transportation
- Public health
- Economic development
- Sustainable development
- Climatology
- Agriculture
- Forestry
- Insurance

Geospatial Data Sources

- Digitization of existing paper maps → vector or raster
- Global Navigation Satellite System (GNSS) → vector
- Imagery (satellite, LIDAR) → raster
- Existing databases (clearinghouses) → vector and raster
- Field surveying → vector
- Crowdsourcing through volunteers → vector and raster

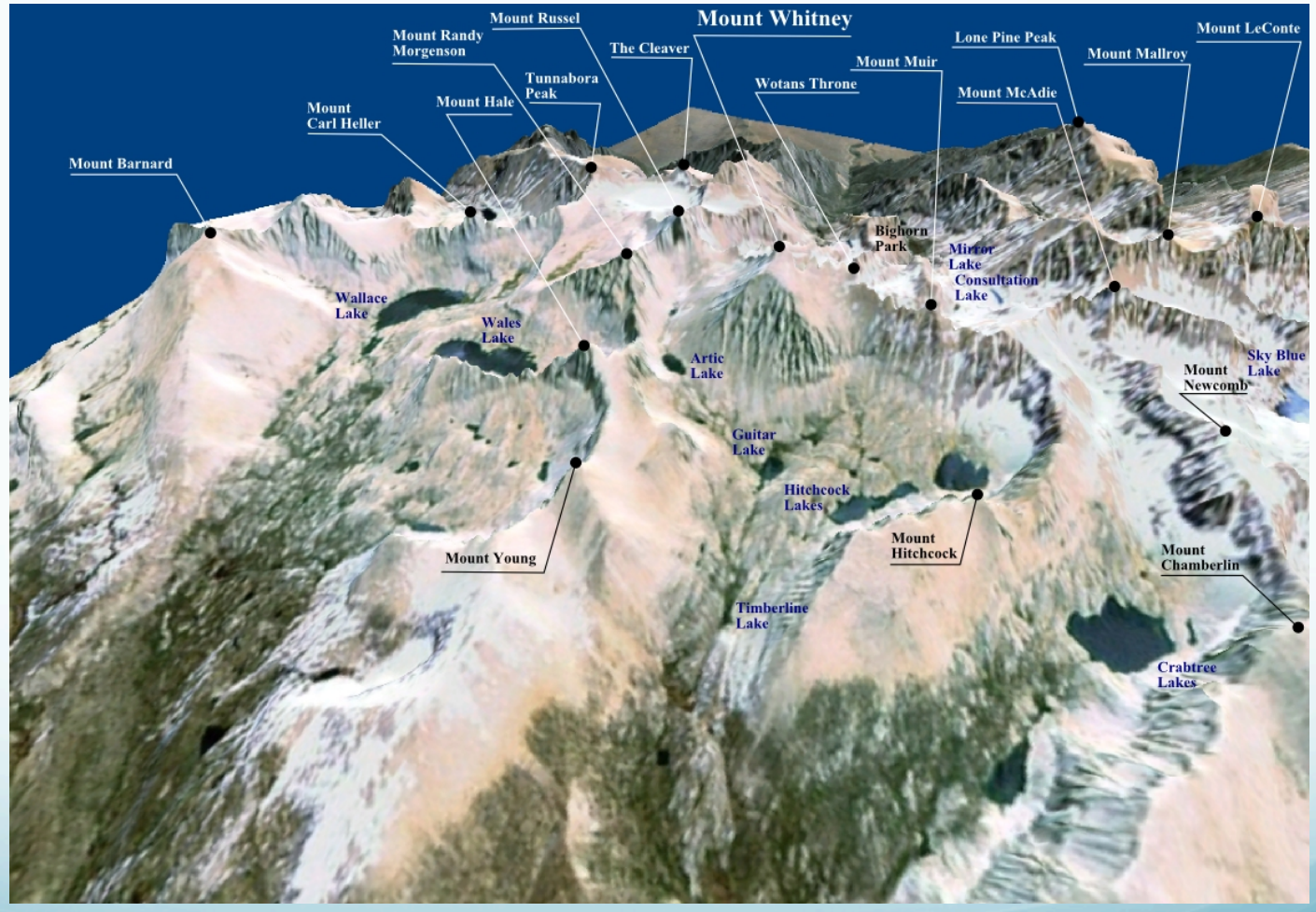
Geospatial Databases

- Vector data
  - Dimension
    - 2D, 2.5D, 3D, 4D
    - Topologically structured
      - Less topology → less storage → more computation
      - More topology → more storage → less computation
- Raster data
  - Resolution
    - Spectral, spatial, temporal
      - Low spatial resolution → less details → less storage
      - High spatial resolution → more details → more storage

2D Vector Data

Source: http://webhelp.esri.com/arcgisserver/9.3/java/geodatabases/feat_c-1166806282.gif

3D Vector Data



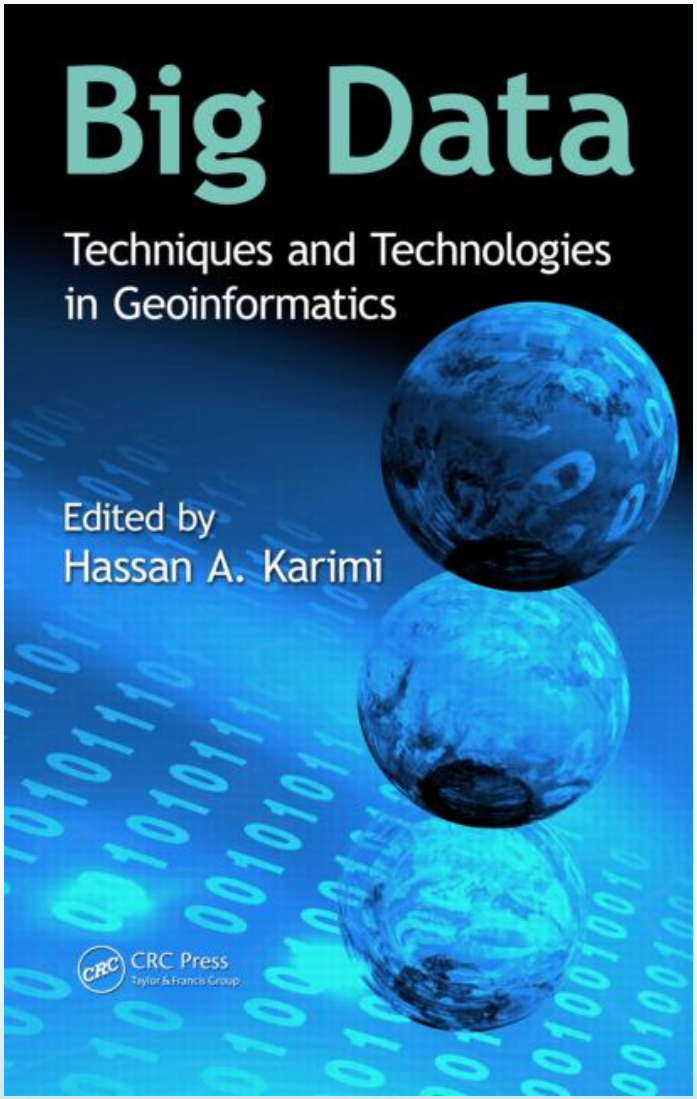
Source: https://commons.wikimedia.org/wiki/File:Mount_Whitney_3D_map_version1.jpg

High Resolution Satellite Image

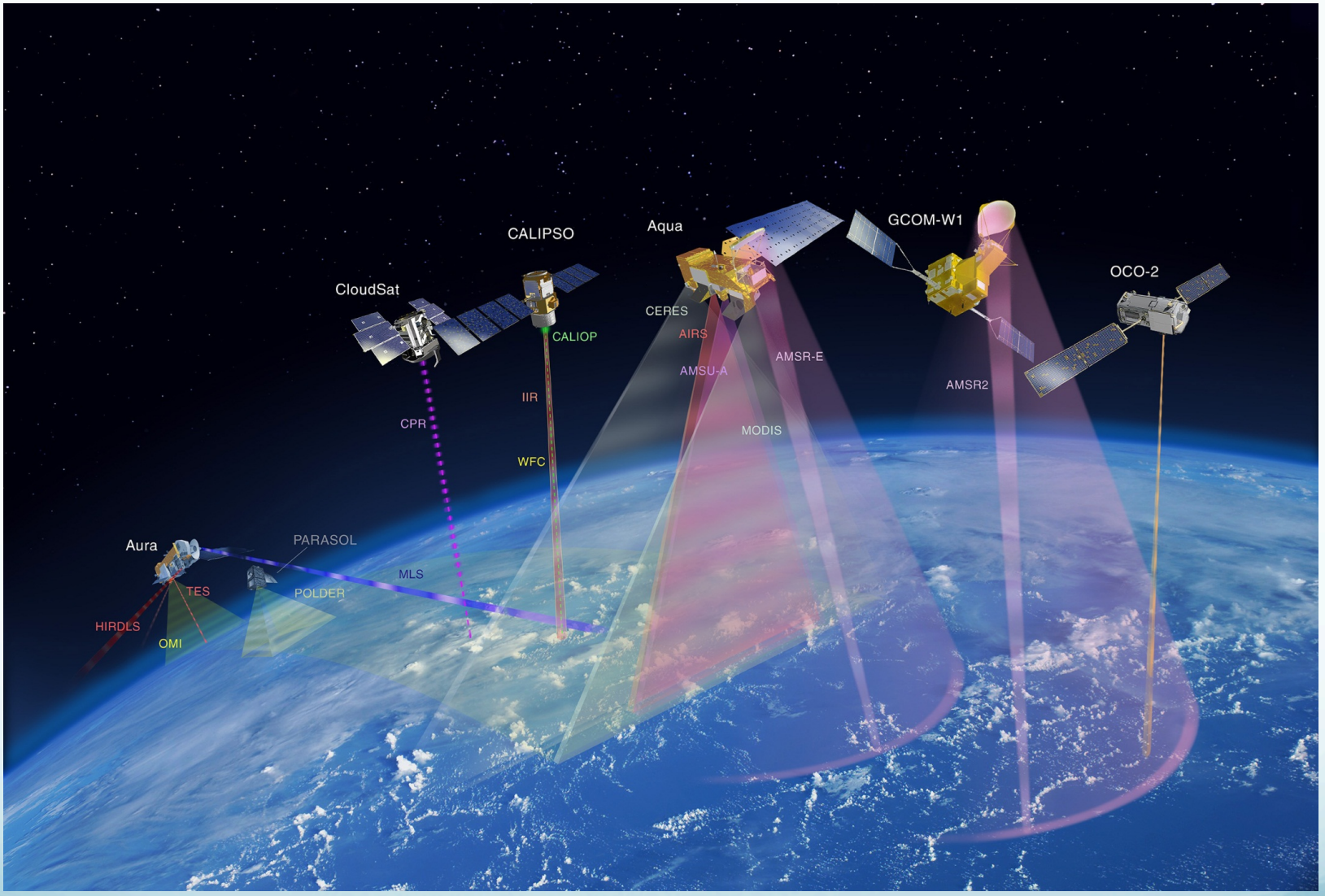
Source: http://modis.gsfc.nasa.gov/gallery/individual.php?db_date=2015-06-03

Big Geospatial Data

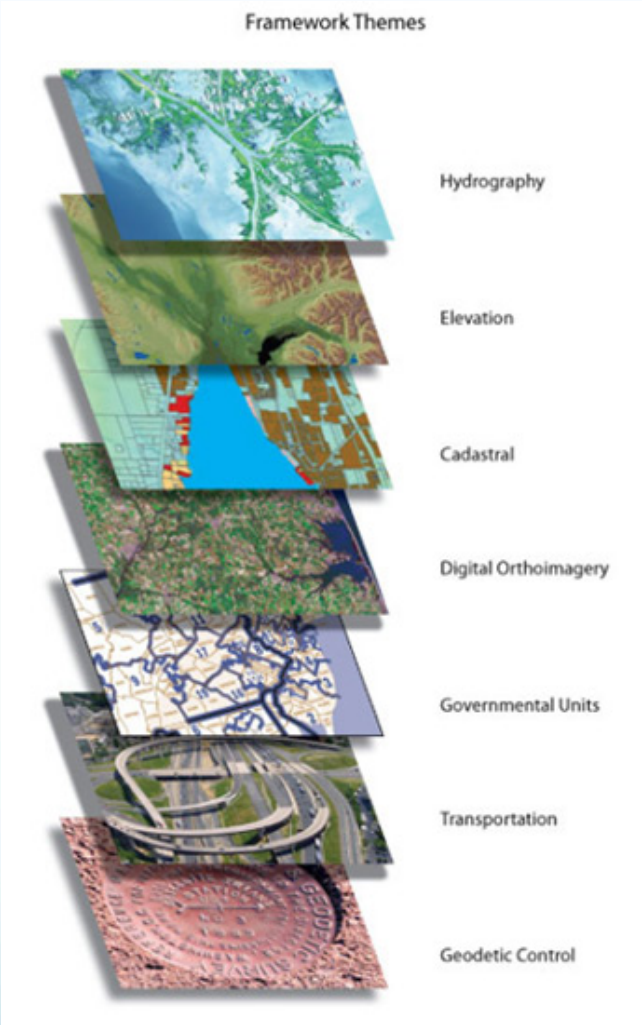
- Vector data
  - Very large databases
    - Different dimensions
    - Different scales
    - Topologically structured
- Raster data
  - Resolution
  - Some satellites revisit every day (terabytes data per day)



Source: https://www.crcpress.com/product/isbn/9781466586512



Source: http://atrain.nasa.gov



Source: http://www.fgdc.gov/images/largegraphics/fgdcgraphic.jpg

Web Mapping

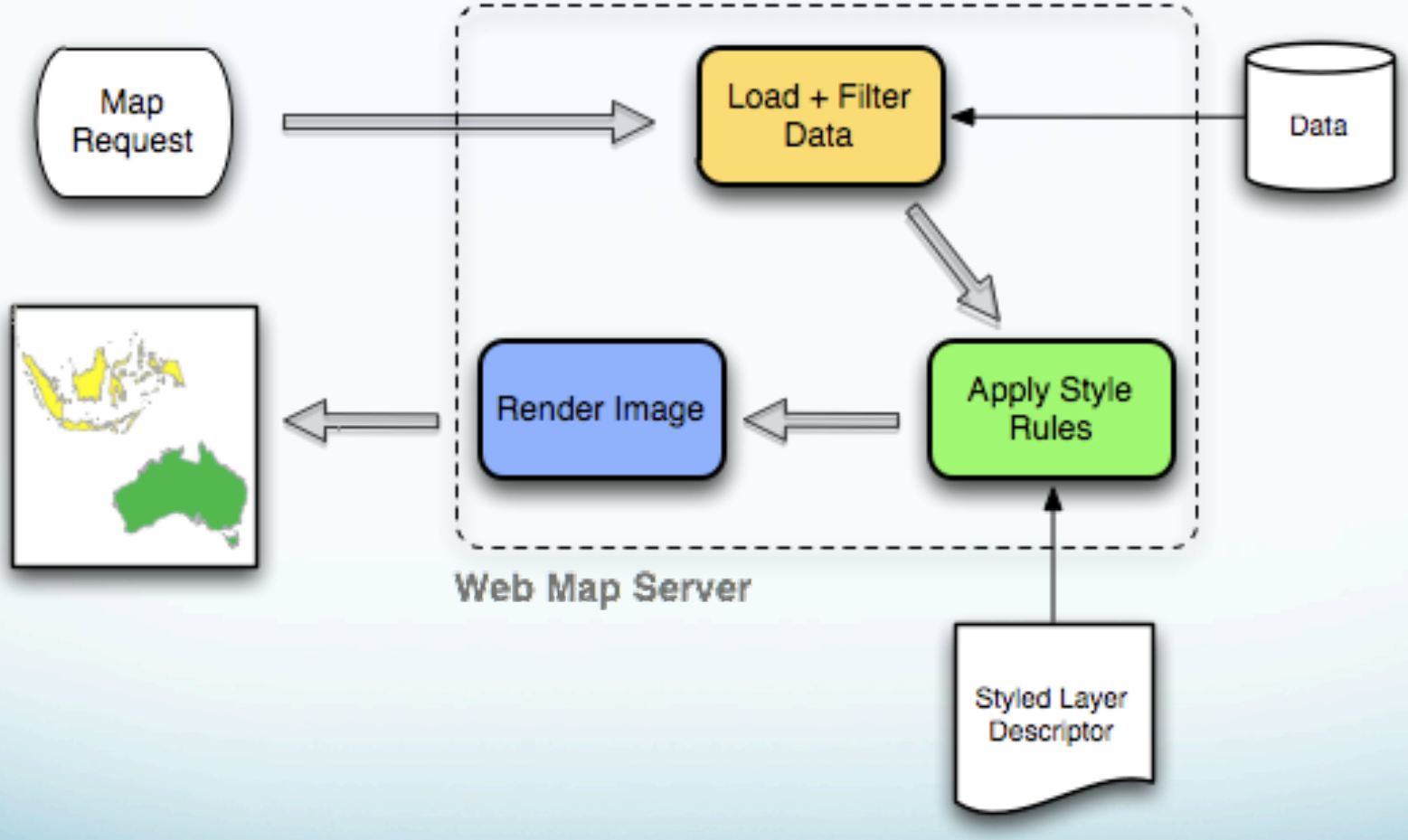
Source: http://resources.arcgis.com/en/communities/mapping/home/GUID-22E6CB53-87B4-4567-8439-1E257515E66E-web.jpg

"E. W. Gilbert's version (1958) of John Snow's 1855 map of the Soho cholera outbreak showing the clusters of cholera cases in the London epidemic of 1854"

Source: http://en.wikipedia.org/wiki/Geographic_information_system#Applications

Topographic Map (Google Maps)

Web Mapping Services



Source: http://presentations.opengeo.org/2012_FOSSGIS/suiteintro/_images/wms.png

Geospatial Data Analytics

- Pitfall
  - Standard statistics do not apply
- Techniques
  - Geometry (e.g., distance)
  - Topology (e.g., adjacency)

Geospatial Data Analytics

- Autocorrelation
  - Point and area
- Point patterns
  - Quadrants
- Cluster detection
  - Scan statistics
- Geometry
  - Proximity polygons
- Spatial interpolation
  - Inverse Distance Weighting (IDW)
  - Kriging
- Trend surface analysis
  - Digital Elevation Model (DEM)
  - Triangulated Irregular Network (TIN)

Cyberinfrastructure for Big Geospatial Data Problem Solving

- Web
  - Recognize different relevant resources (data, people)
  - Connect different resources (HW, SW)
  - Allow platform-independent access and computation
  - Facilitate distributed computing

Web Computing

- Crowdsourcing geospatial data
  - Collaborative mapping
- Smartphones
  - Geo-tagging
  - Gateway to cyberinfrastructure
- Personalization of maps
- Web services for handling complex tasks
- Collaborative problem solving
- High-performance computing on Web browsers
  - E.g., 3D rendering and animation (4D)

Thank You