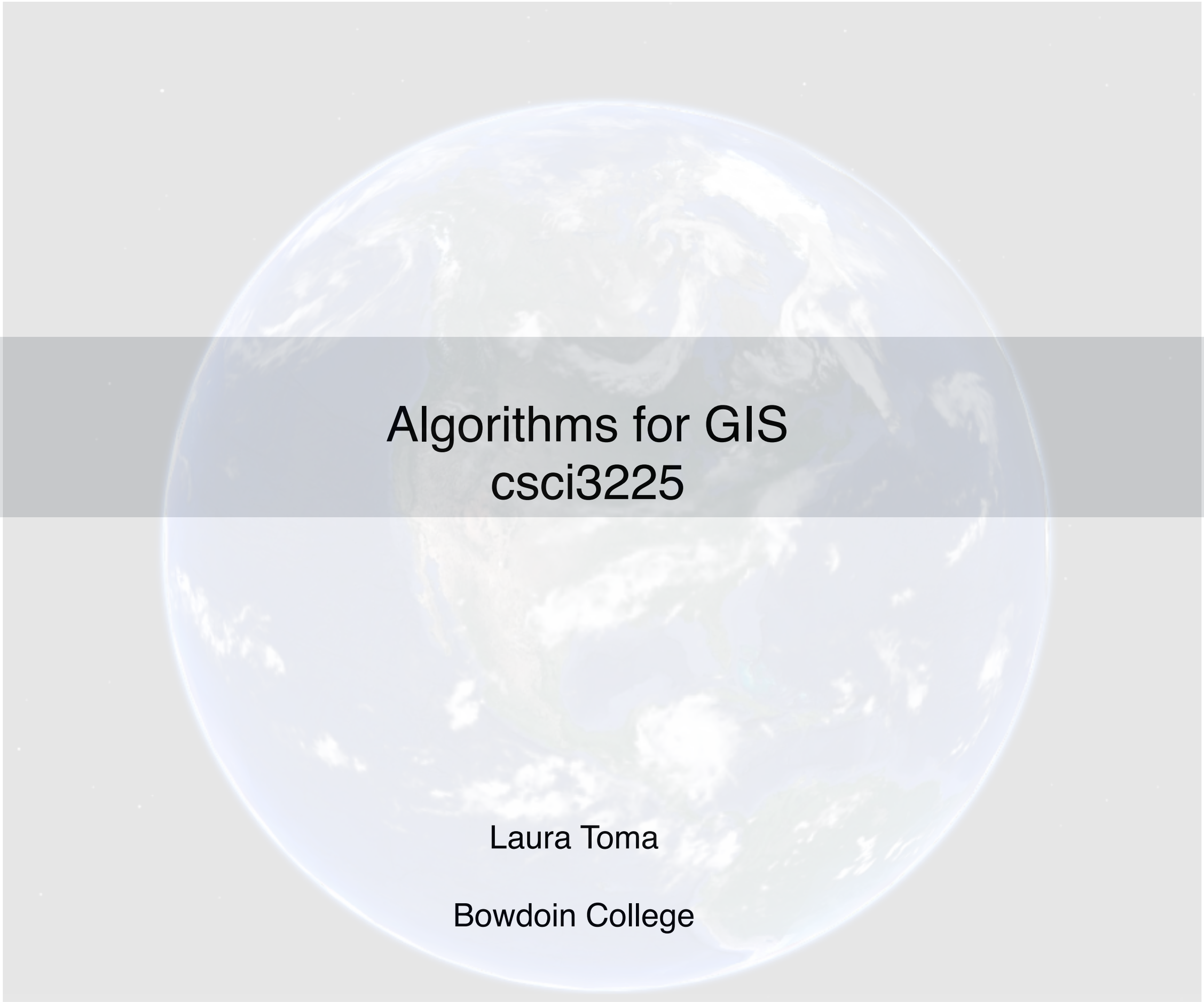
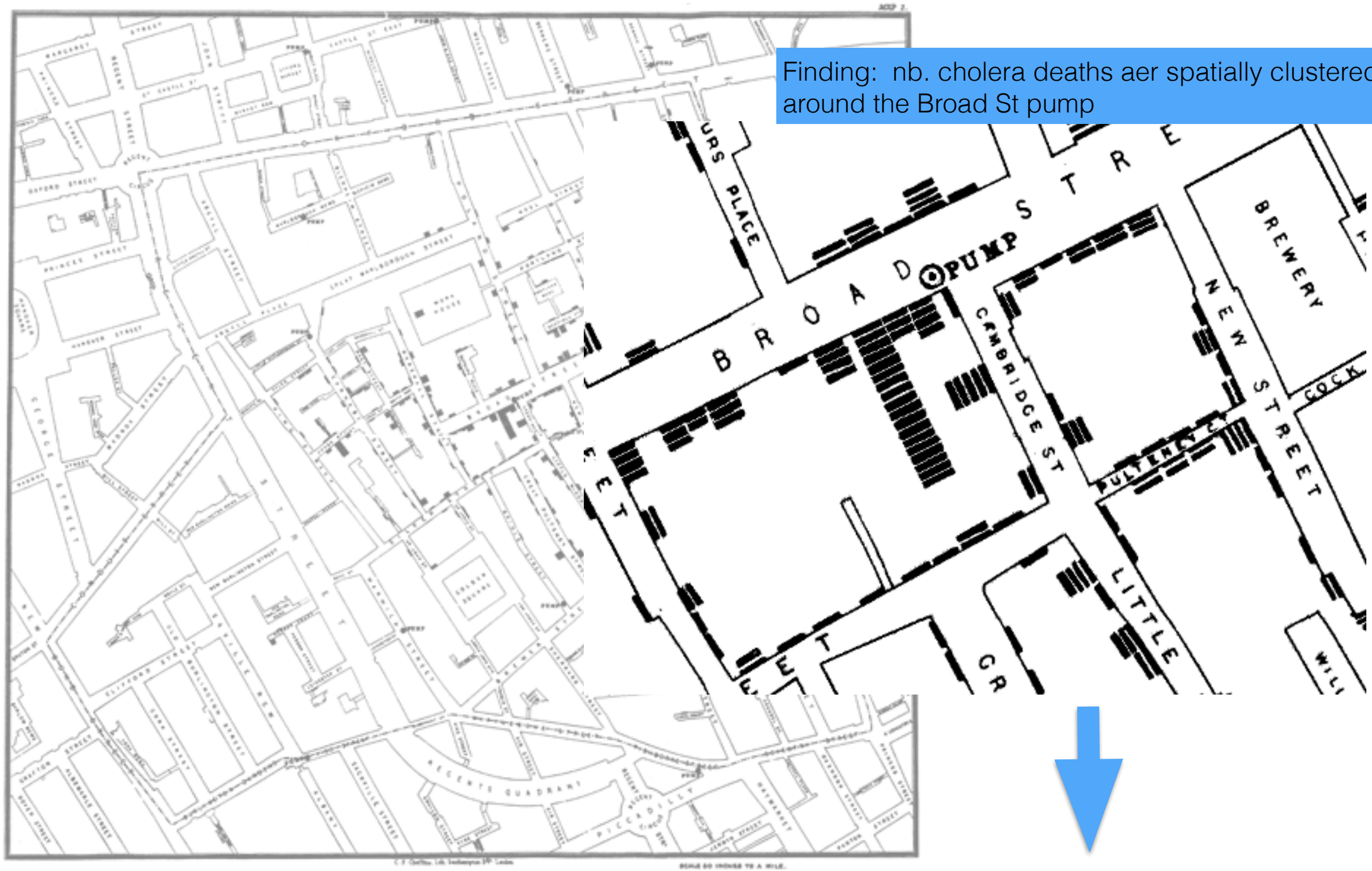


Algorithms for GIS
csci3225

Laura Toma

Bowdoin College

Spatial analysis: the beginnings



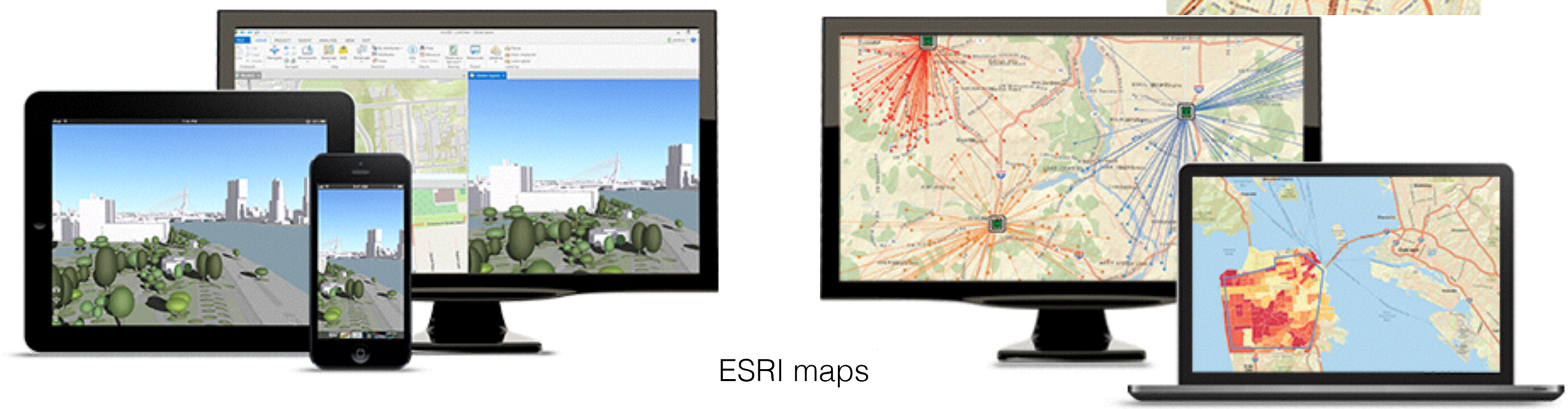
Finding: nb. cholera deaths aer spatially clustered around the Broad St pump

1848, London, John Snow

Claim: Cholera is a water-transmitted disease

GIS (Geographic Information Systems)

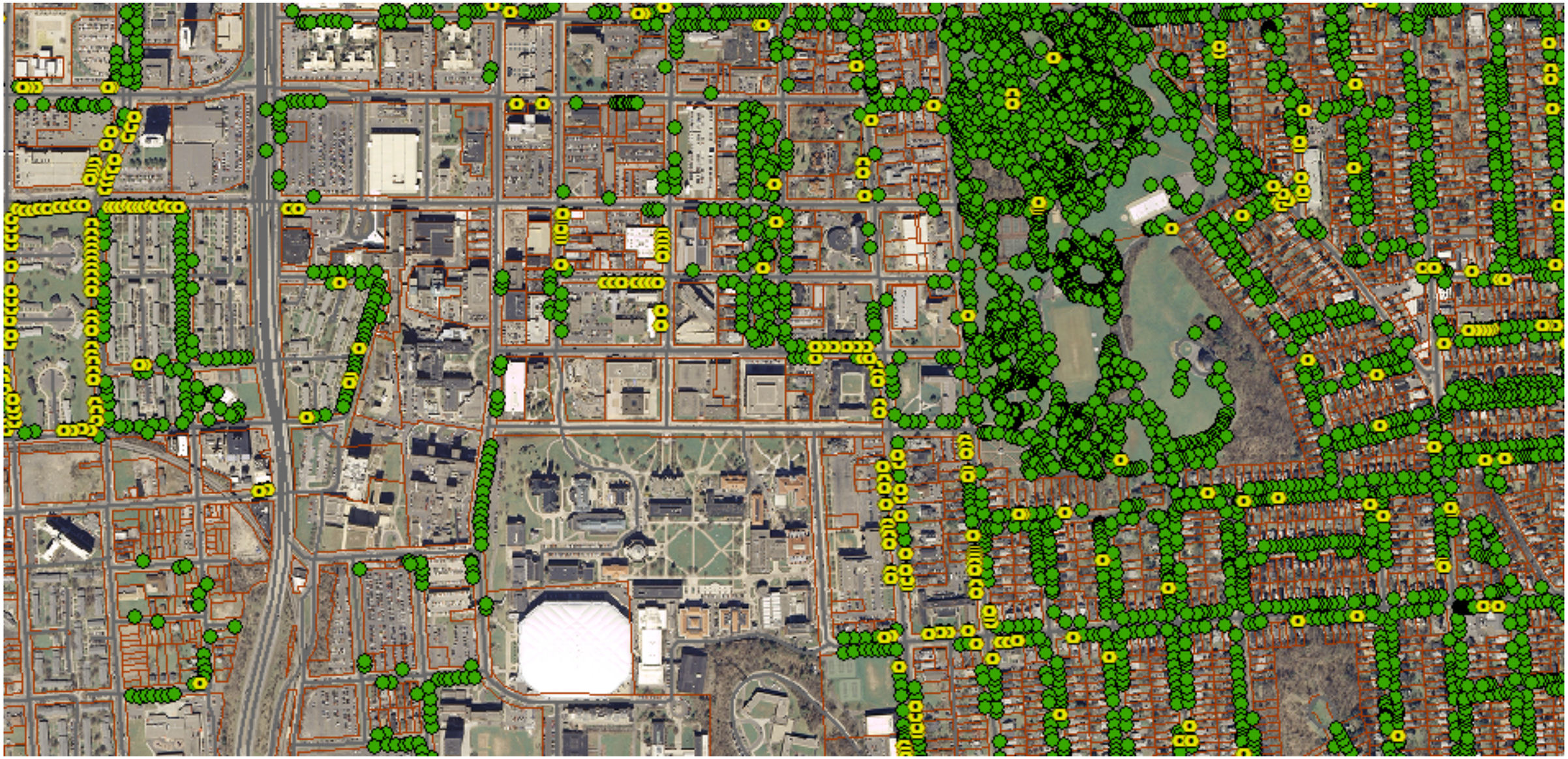
- Systems for storing, visualizing and analyzing geospatial data
- Started in 1970s as an extension of traditional cartography
- First use: Mapping and visualization
  - Display different types of data, all on same location (layers), turn layers on and off, zoom in/out, etc
  - Create beautiful, interactive maps
  - Combine data from many sources



ESRI maps

GIS: Vizualization

http://researchguides.library.syr.edu/c.php?g=258118&p=1723814



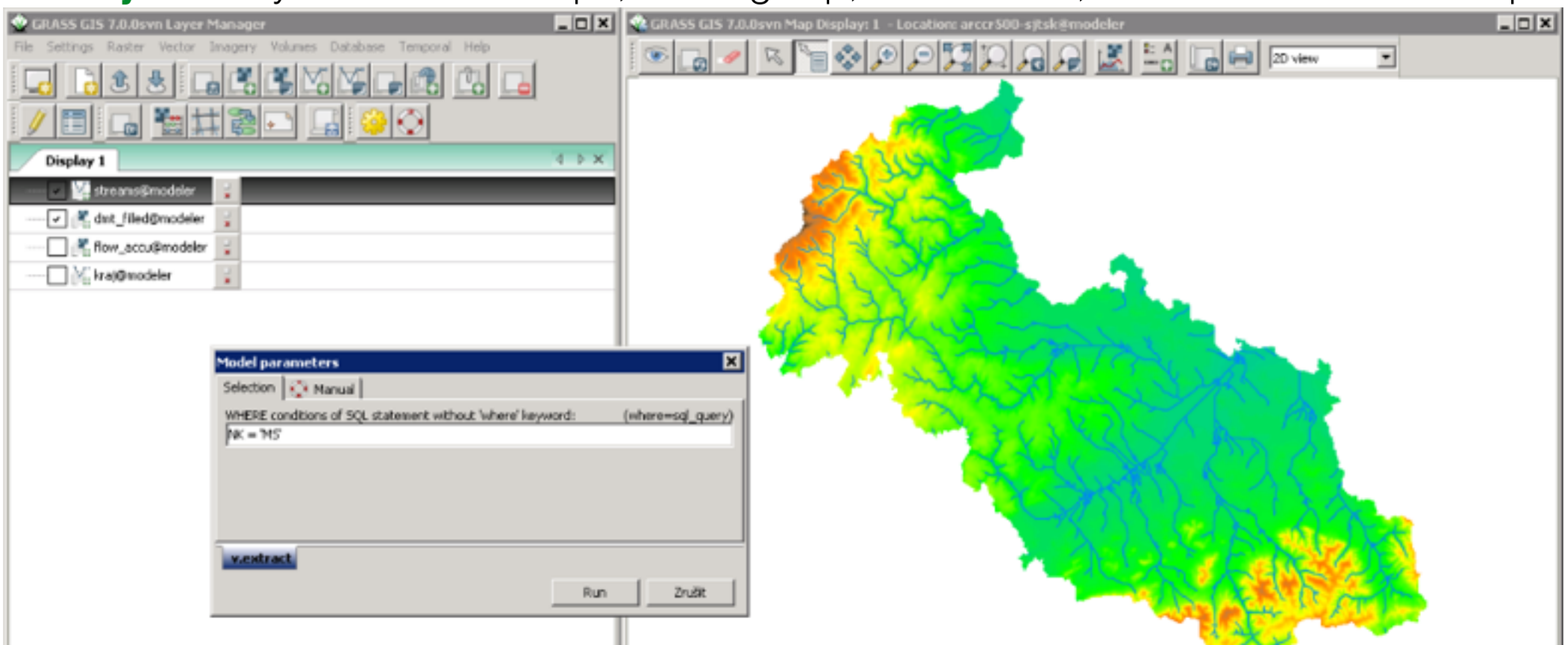
Polygons (properties), lines (streets), points (trees) and raster images (air photo) are integrated into one map.

GIS: Spatial analysis

- Spatial operations
  - e.g. What lies within 5 miles of a dump site?
  - e.g. What other crimes have occurred in this selected region?
  - e.g. What is the total length of the river network?
  - e.g. Find shortest routes, connectivity

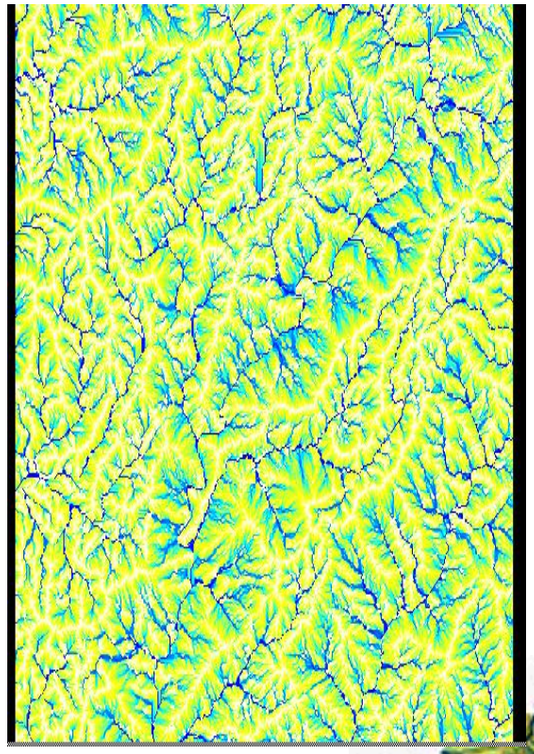
GRASS: module of the day

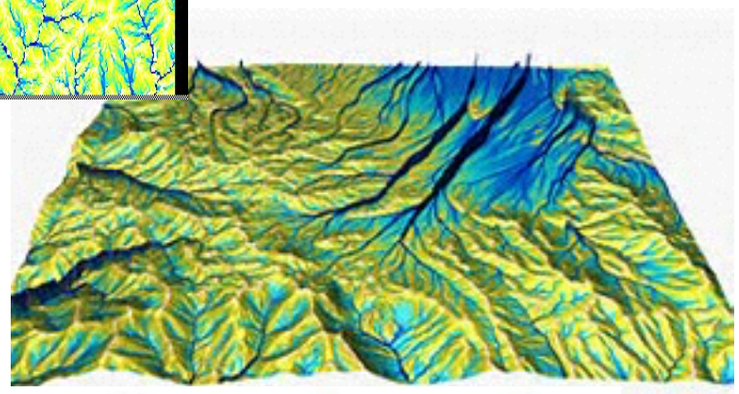
v.overlay: overlays two vector maps, offering clip, intersection, difference and union operators

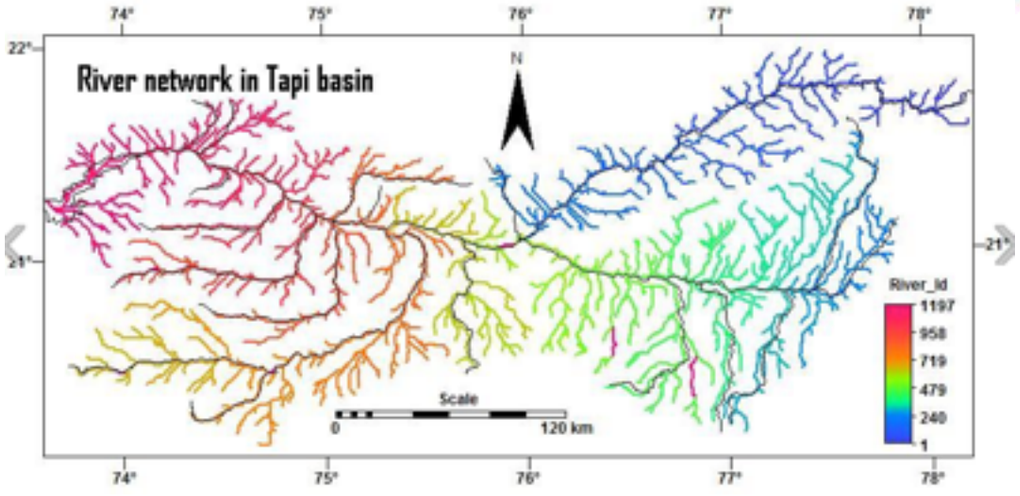


GIS: Terrain analysis

e.g. modeling flow

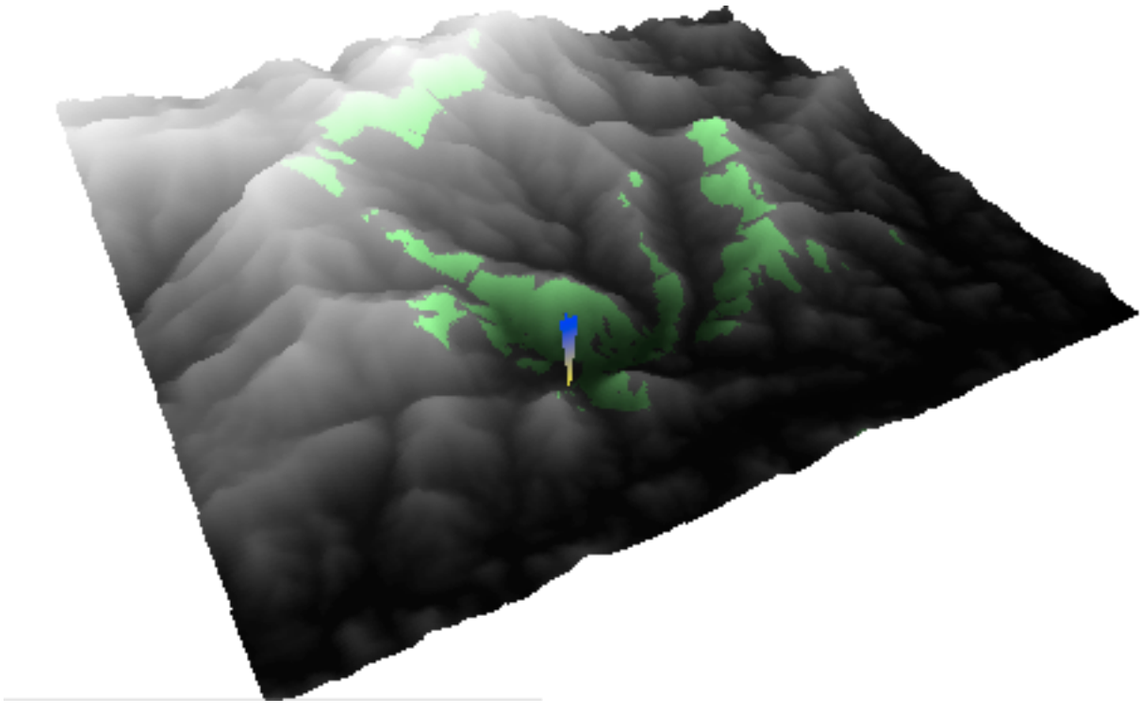


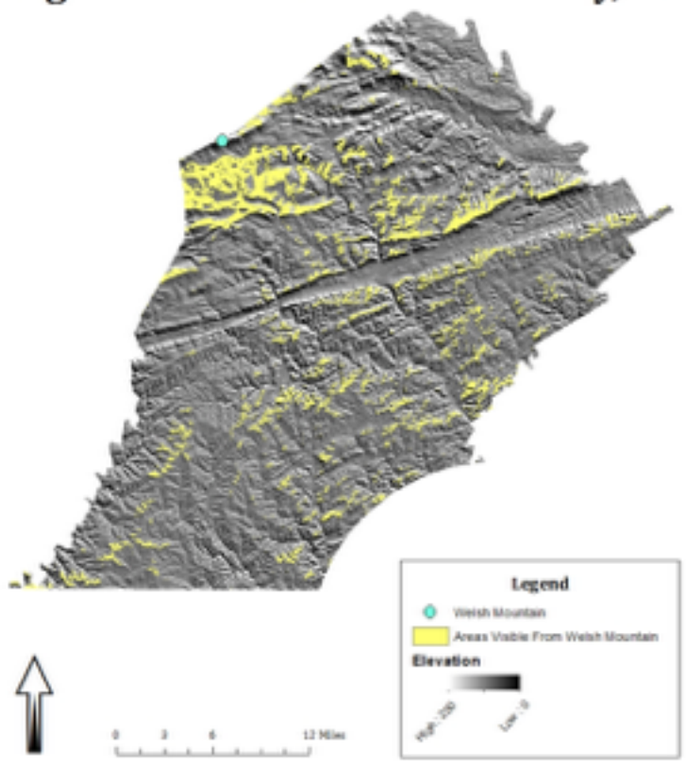




GIS: Terrain analysis

e.g. modeling visibility



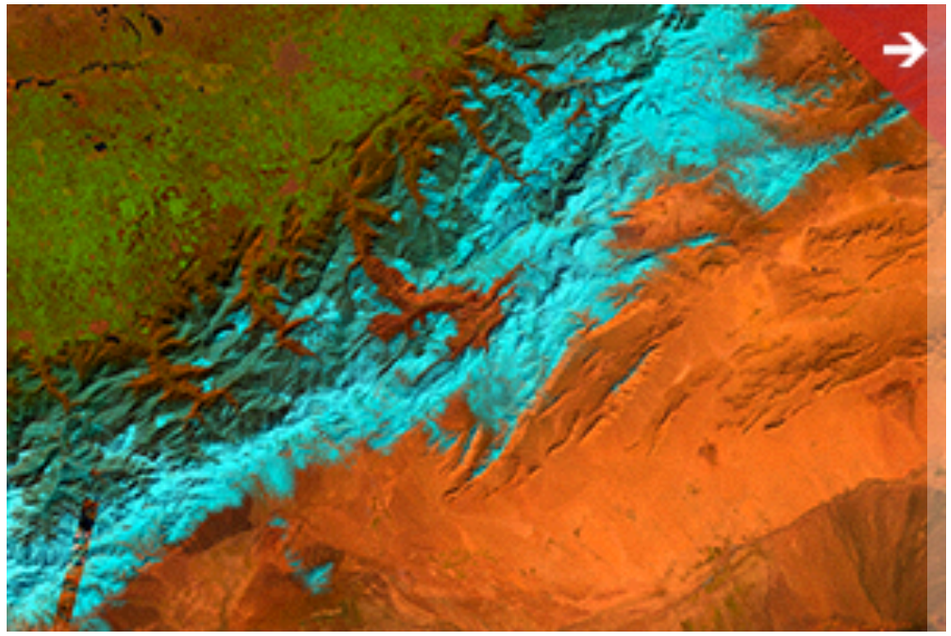


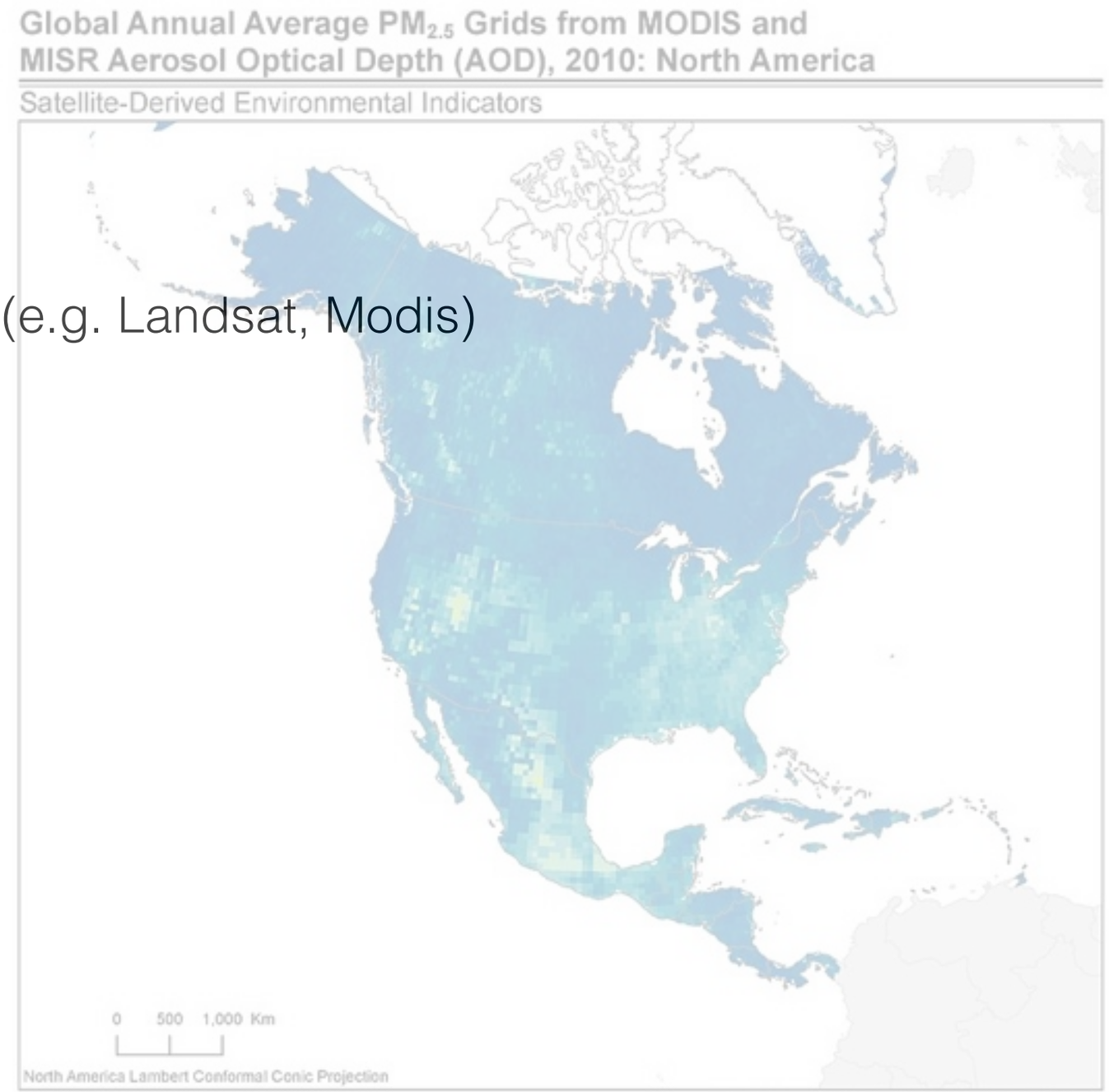
GIS: Terrain analysis

e.g. modeling flooding

GIS: Satellite imagery

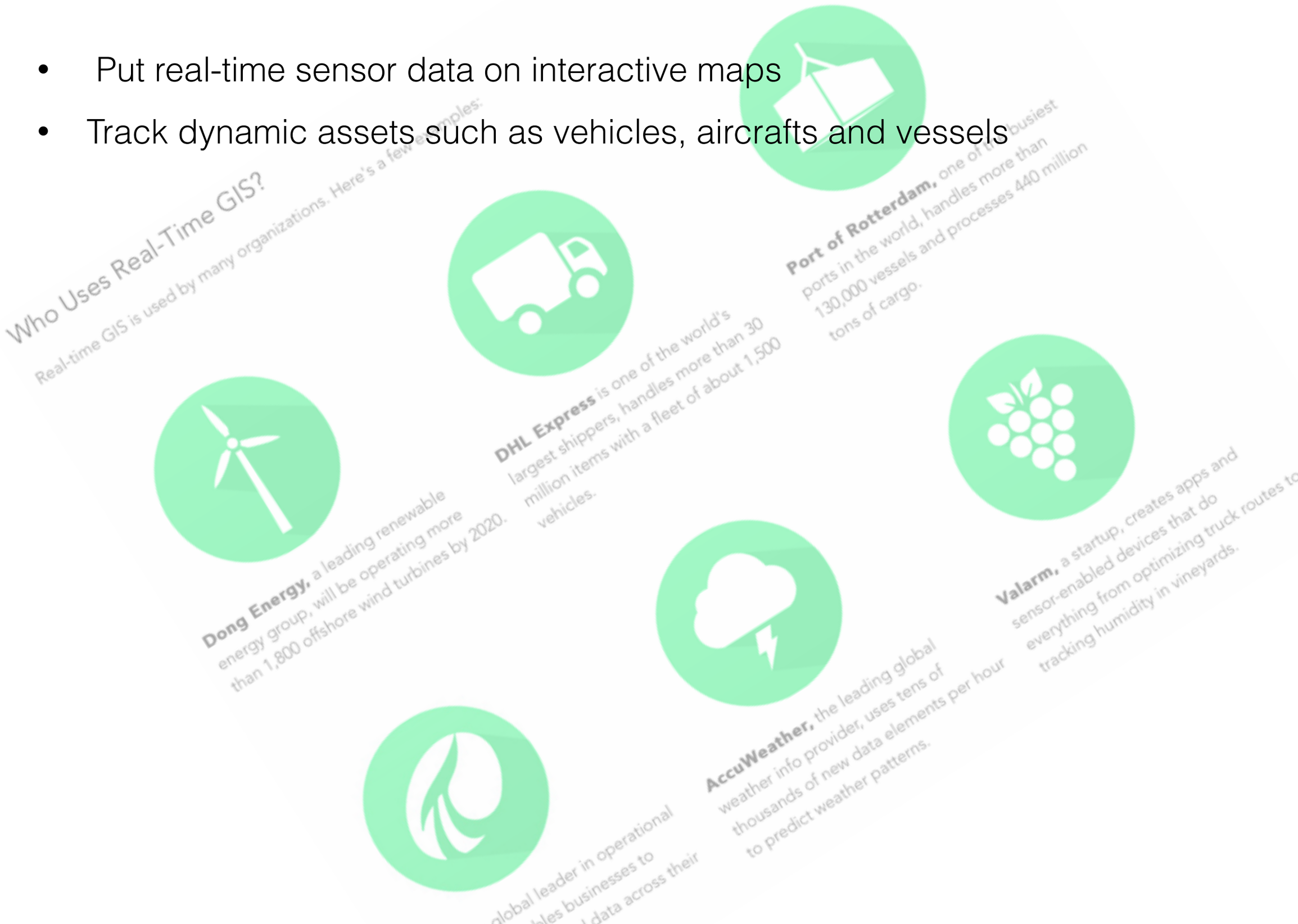
- Access existing data collection (e.g. Landsat, Modis)
- Visualization and analysis





Real-time GIS

- Put real-time sensor data on interactive maps
- Track dynamic assets such as vehicles, aircrafts and vessels



GIS

- Used by a growing number of disciplines
    - earth, atmospheric and oceanographic sciences
    - environmental studies
    - digital humanities, …
- Also used by city planners, government, …

Explosion of digital data ==> GIS has seen tremendous growth

GIS software

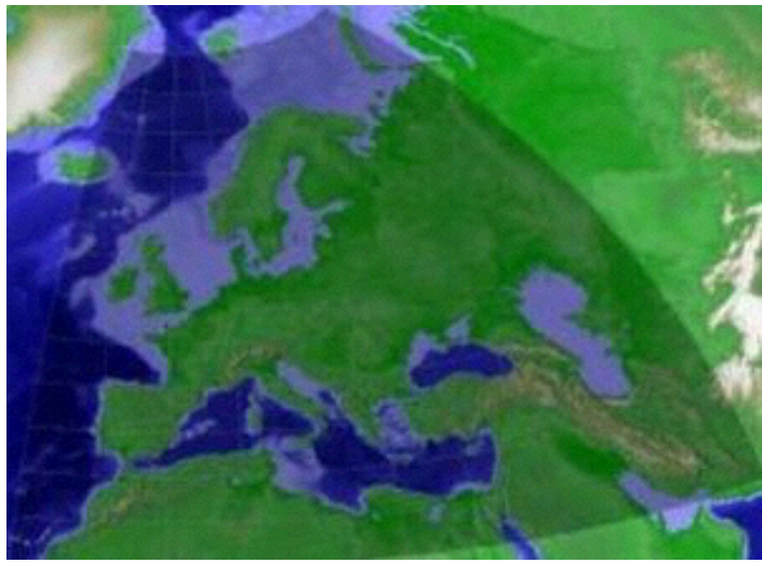
- **ArcGIS**
  - developed by ESRI
  - probably most comprehensive system
  - complex interface
  - available in Bowdoin labs ; IT/ES offer tutorials

- https://grass.osgeo.org/# Open source systems
  - e.g. **GRASS GIS**, **QGIS**

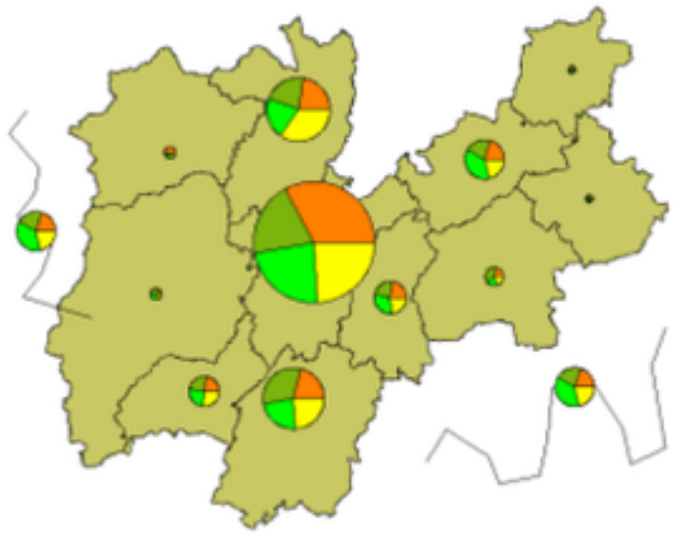
- Other proprietary modules, with specialized functions
  - e.g. **LAStools**

e.g. GRASS screenshots

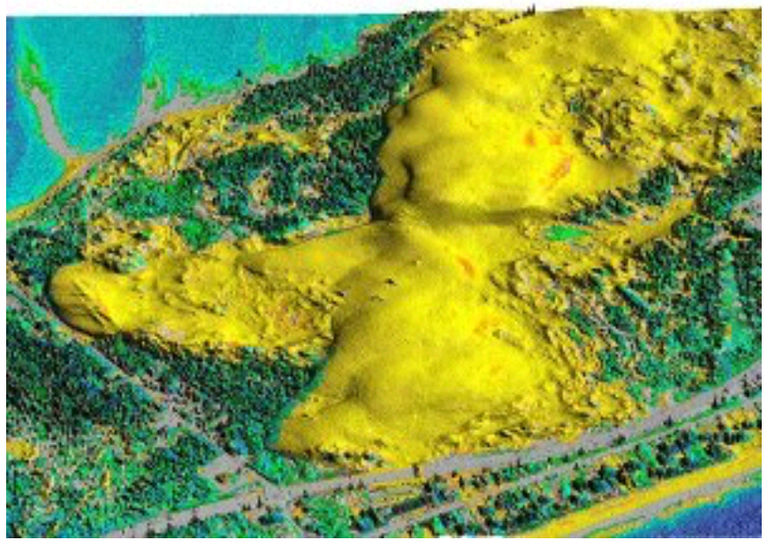
Raster map operations



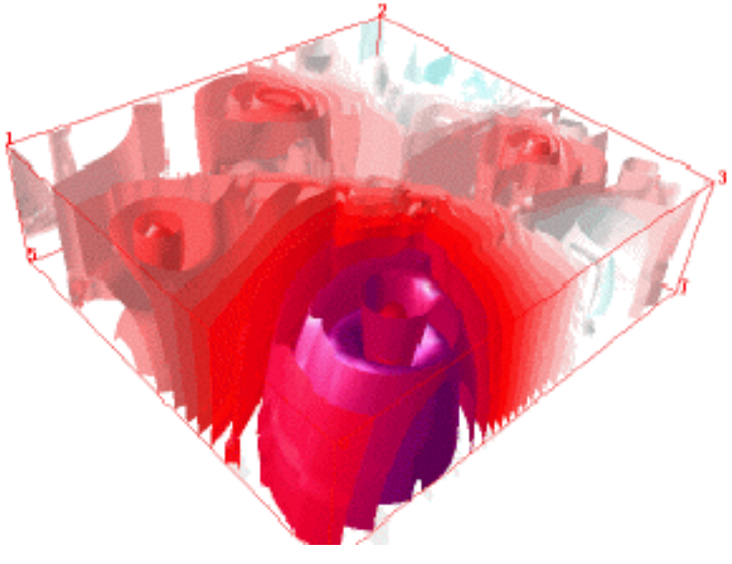
Vector map operations



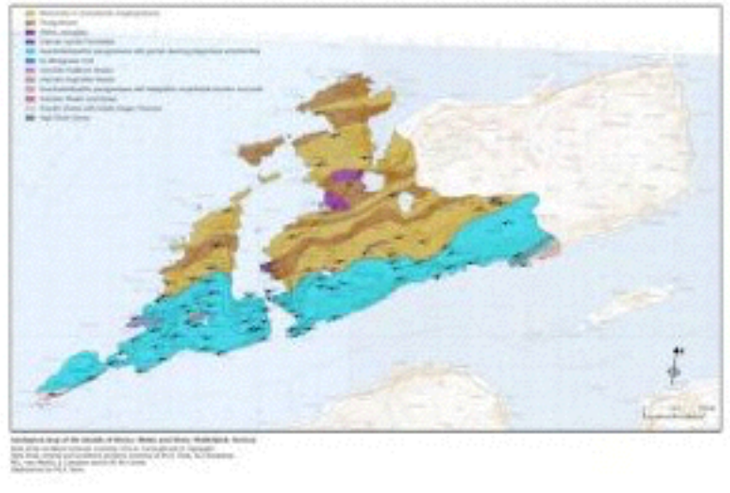
LiDAR data Processing



3D Visualization



Cartography



GIS and geospatial analytics

- Rich source of problems in CS
  - algorithms, databases, interfaces, visualization, cloud computing, systems

- LOTS of geospatial data available
- What can one do with it?

geospatial “data science”

Descartes Labs raises $30 million Se
brand of geospatial analytics

Posted Aug 24, 2017 by John Mannes (@JohnMannes)

Descartes Labs is announcing a $30 million Series B this mor
led by March Capital. This comes just three months after one c
competitors, Orbital Insight, closed its own $50 million Series C

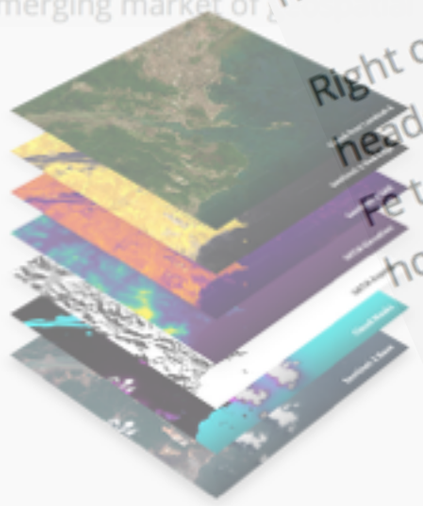
Descartes Labs

Both Descartes and Orbital are playing in the emerging market of

A data refinery, built to understand our planet

Instant access to science-ready imagery and intelligence from multiple data sources.

Read more about our Series B raise



FOUNDED
2014

OVERVIEW
Descartes Labs launched as a spin-out from Los Alamos National Laboratory in April of 2015. Th
underlying technology uses computer vi
learning, and cloud-based infrast
computers to see and un
them. Initially, the
an agricult
pro

Both Descartes and Orbital are playing in the emerging market of geospatial analytics. These companies use machine learning to produce insights from satellite imagery and other data. This capability has proved exceedingly popular with hedge funds where images of store parking lots, for example, can be used to project out revenue numbers. But companies running the gamut from agriculture to logistics see the value in having an extra pair of trained eyes in the sky.

Most of Descartes previous investors participated in today's growth round. Crosslink Capital and Cultivian Sandbox participated though Data Collective did not. The company previously raised $8.3 million in funding.

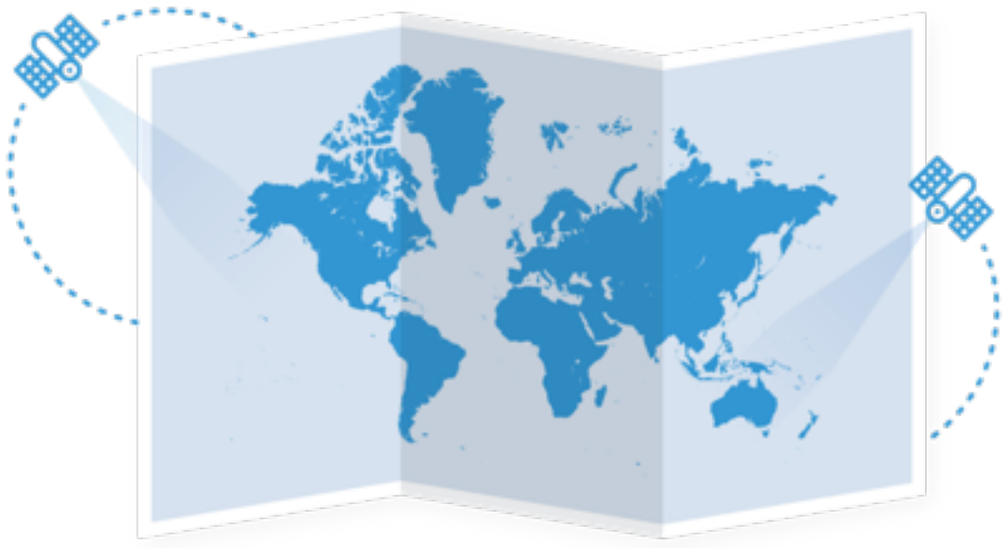
Right out of the gate, one of the more unique things about Descartes is that it is headquartered in Santa Fe, New Mexico. Co-founder Mark Johnson pitched Santa Fe to me as a place where an engineer and his or her family can actually buy a house. Today's raise is one of the largest in the history of New Mexico.

"Our original founding thesis was that lots of money was going to satellite hardware but not to the equivalent software," Johnson explained to me in an interview. In the three years since, more money than ever has been spent on geospatial software technologies. Johnson noted the number of startups he meets on a basis trying to do interesting things combining deep learning and the myriad taken from different satellites every day continues to grow.

Descartes Labs

Descartes Maps

For decades, satellites have been collecting data about our planet. Yet complex global systems like agriculture, deforestation, water cycles, and many others, affect billions but are still poorly understood. Not only are these systems critical to understand for the good of humanity, they're also important for businesses as they adapt to — and attempt to thrive in — an ever-changing environment.



With the price of satellites dropping from over a billion dollars to tens of thousands, new constellations with new sensor characteristics are coming online, producing ever more data. Combining multiple sources provides a far richer view than that of any single constellation. For organizations that want to harness this power and enable computation at global scale, it's necessary to have a data refinery that combines data from multiple sources, cleans it up and makes it ready for science. That's exactly what we've built at Descartes Labs.

They build global composites using their cloud-based parallel computing infrastructure

Landsat 8

RGB bands

15 meters per pixel resolution

3.1 trillion pixels per band (Red, Green, Blue)

lt from 70 trillion pixels per band captured from 2013 to 2017 (320 TB)

Sentinel-2

Red Edge bands

20 meters per pixel resolution

1.8 trillion pixels per band (RE1, RE2, RE3)

Built from 22 trillion pixels per band captured from 2015 to 2016 (120 TB)

Descartes Labs

for United States

As of October 11th 2016

Corn yield

171.2

bu/a (± 2.1)

+1.7% from 2015

Soy yield

48.8

bu/a (± 1.4)

+1.6% f

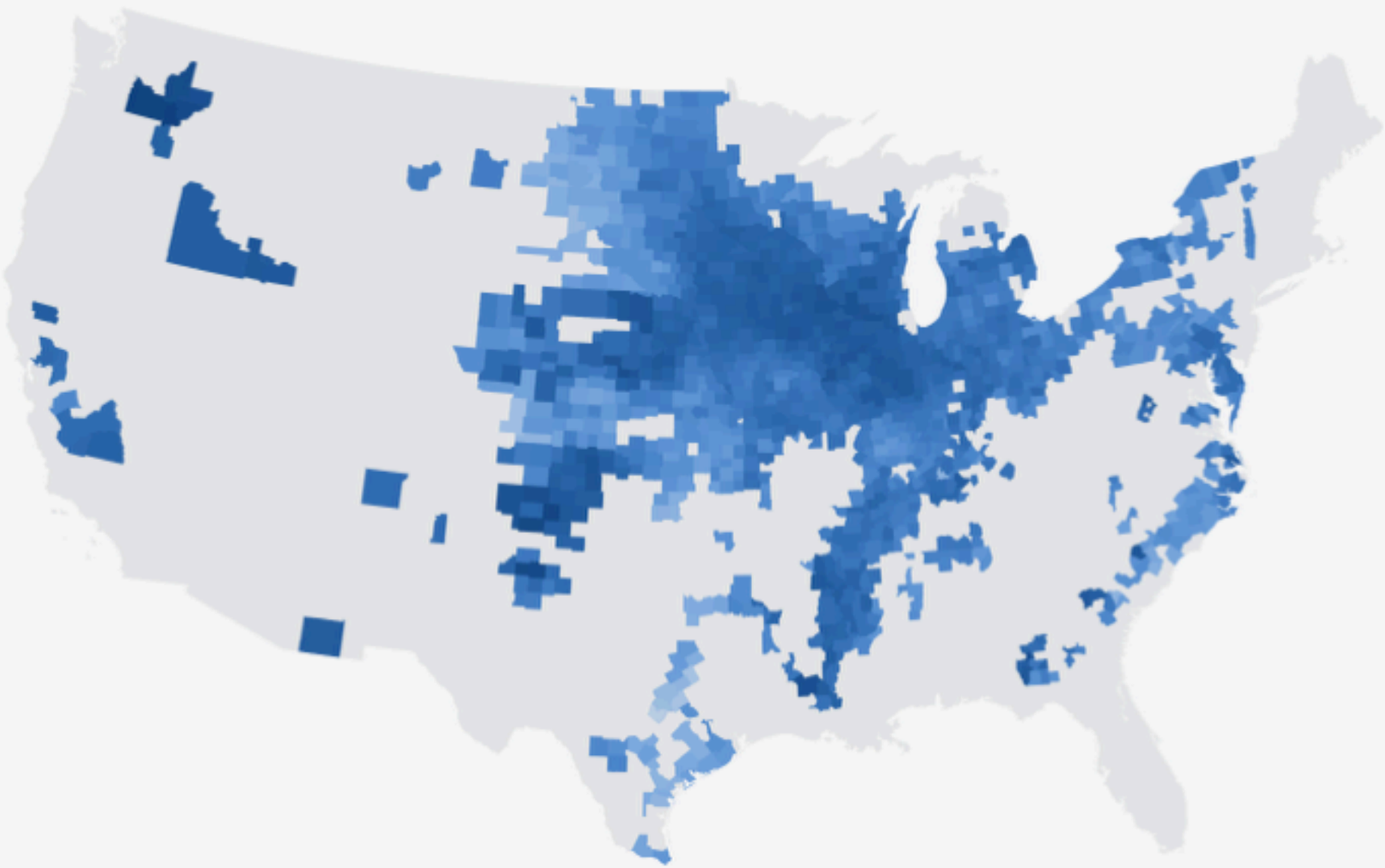
Agriculture Forecasts

Descartes Labs offers a growing collection of global agriculture forecasts and intelligence for trading, finance, insurance, and government.

Learn more

We generate a new forecast every day because we collect new data every day. A lot of it. Nearly 5 terabytes. This data consists of satellite imagery from multiple NASA, ESA, and commercial constellations. We also collect the latest weather readings and additional agronomically significant signals in real time. To enable data collection and daily processing at this rate, we've developed our own hyper-scalable, machine-learning platform. Forecasting at this scale simply cannot be accomplished by sampling individual farms.

Descartes Labs releases a weekly forecast for every corn and soy producing state and county in the United States. This visibility is unprecedented. USDA NASS provides monthly forecasts at the national level and for select states. With Descartes Labs, you can see your county's end of year yield months in advance of the final January production report.



Descartes Labs

GeoVisual Search

By using machine learning at a massive scale, we've built a computer vision tool that lets you search for any object over the entire globe.

Learn more

Orbital Insight

Platform Solutions About Us Blog Login

A Macroscope for Understanding Earth

Satellite Image Spot6 Baku Azerbaijan

Geospatial data follows Moore's Law

Machine vision comes of age

Parallel computing moves to the cloud

Actionable intelligence for the world

Poverty Mapping

Assessing socio-economic status of a region by combining insights from building development, crop yields, infrastructure and more.

Agriculture

Estimating harvest yields throughout the growing season for various crops in countries of interest.

Orbital Insight

US Retail Traffic

Identifying and quantifying retail traffic patterns in parking lots of over 90 major retailers

Global Water Reserves

Weekly estimates of surface water reserves for any region in the world.

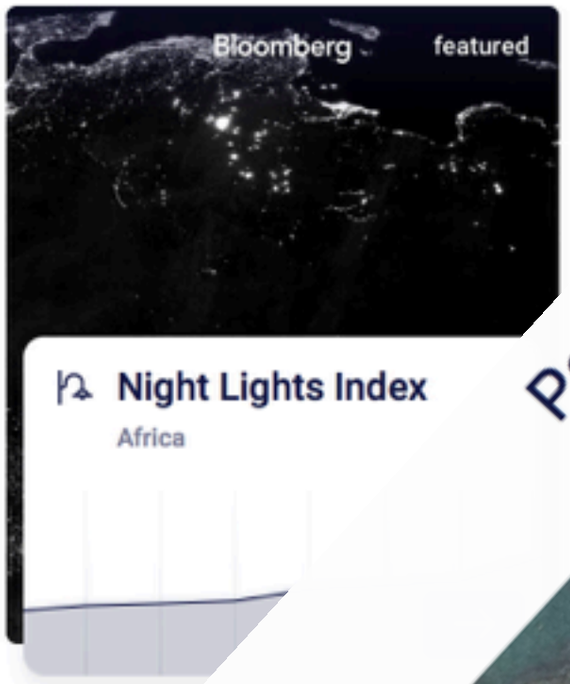
World Oil Storage Index

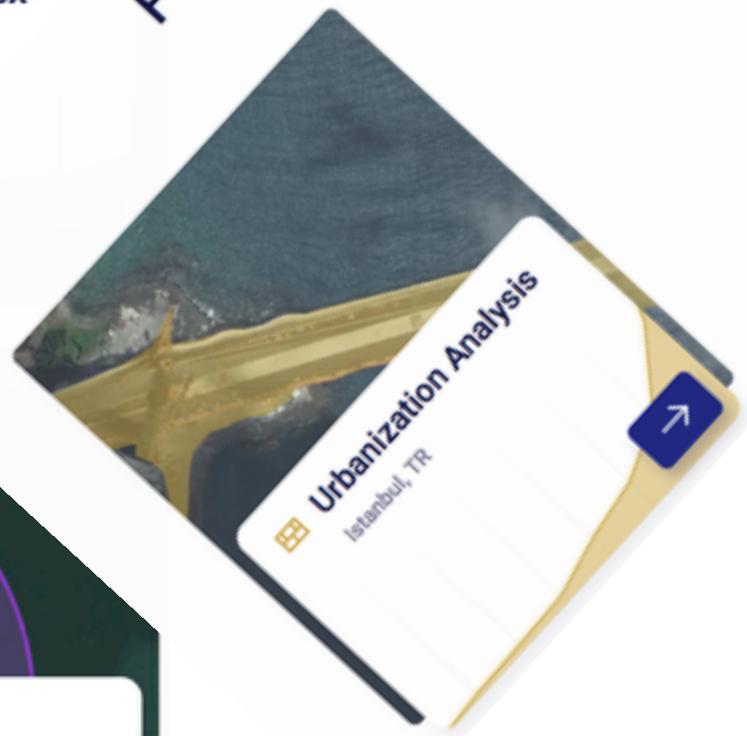
Continuous monitoring of global crude oil inventories

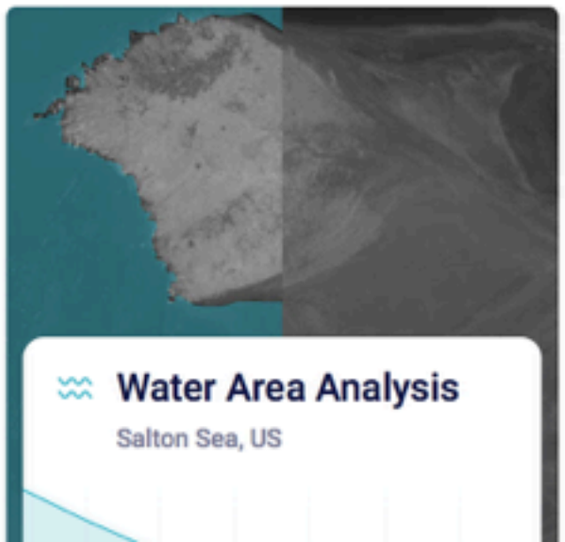
SPACEKNOW

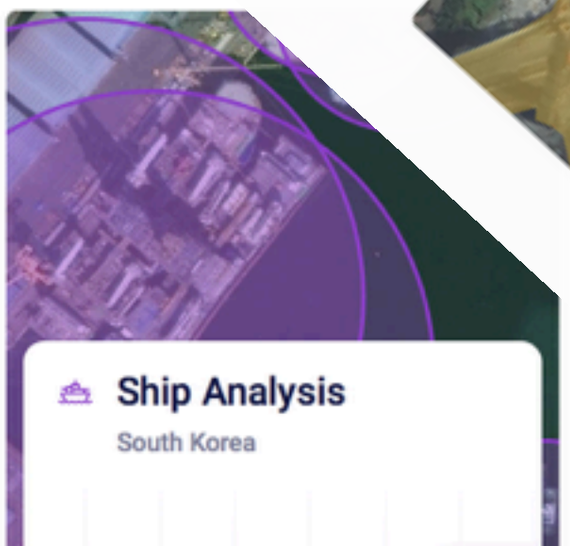
Use satellites to get actionable data

Popular Analyses









Logistics

Class overview

- Explore selected GIS applications and the algorithms/data structures involved.

Some basic GIS topics

- Data models (raster, vector, TIN) and representations
- Shortest paths and least-cost path surfaces
- Flow
  - river network, watersheds, flooding, sea-level rise
- Visibility
  - viewshed, total viewshed, guarding, and approximation.
- Simplification
  - 2D (line simplification) and 3D (terrain simplification).
- Spatial data structures: B-trees and quadtrees.
- LiDAR data

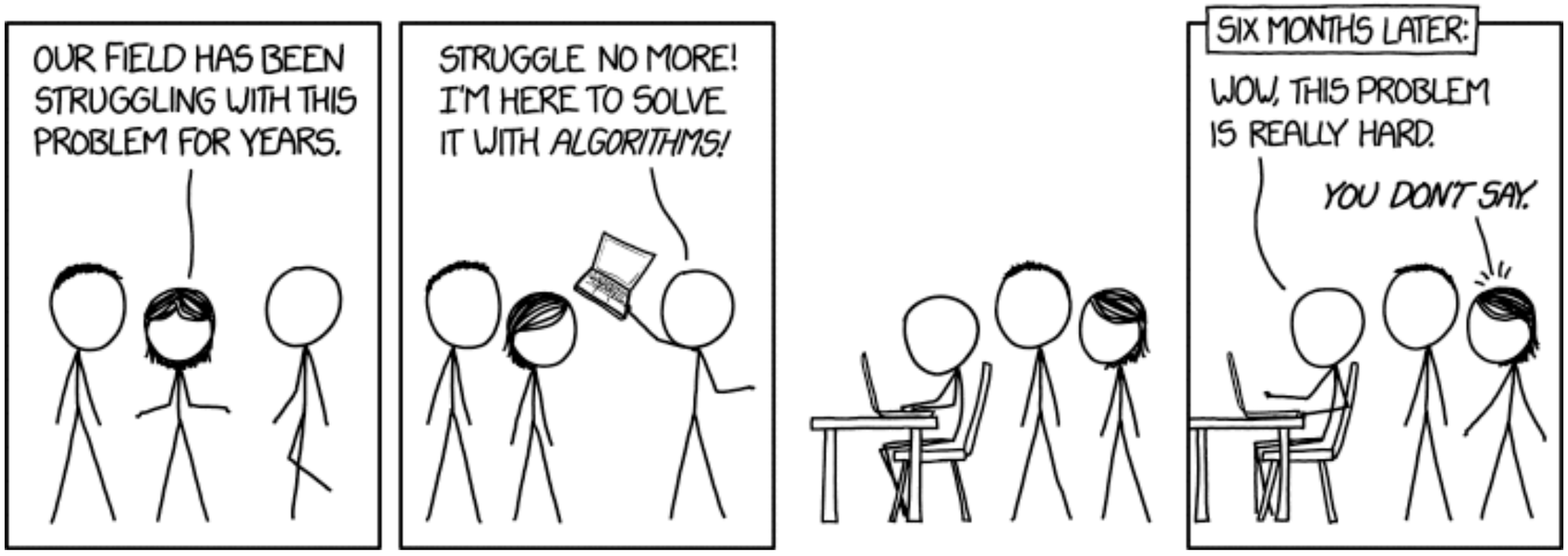
Visualization

- programming with OpenGL

Dealing with large data

- Parallel programming with OpenMP
- cache- and I/O-efficiency
- Space-filling curves

Focus: the interplay between theory and practice and scalability to large data



Class info

- Pre-requisites
  - 1101, 2101 (data structures) and 2200 (algorithms)
- CS curriculum
  - Satisfies the “projects” requirement
  - Satisfies the “theory” requirement
- No textbook
  - Papers, slides, and other online materials
- Discussion forum
  - Piazza
- TAs
  - Jason Nawrocki

Work

- The work for the class consists of
  - programming assignments & project
  - research papers reviews and class discussions
  - presentations (such as project proposal, updates, and final presentation)
  - final project report, final project demo and presentation
    - date in polaris: Dec 18, 2pm
- class participation