

Geography 360: GIS & Mapping

Data Models and Databases

Data Collection II

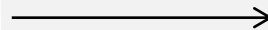
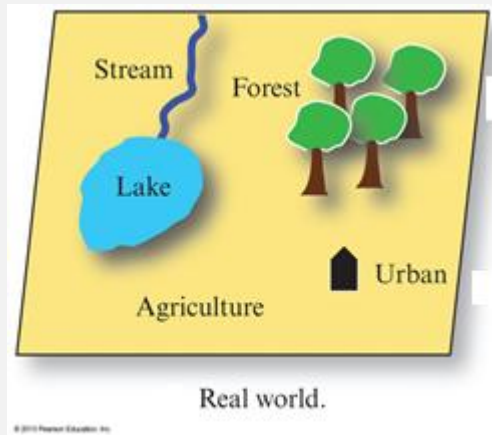
Vaishnavi Thakar



UNIVERSITY *of* WASHINGTON

Review

Raster Data Collection



1	1	1	1	2	1	1	1	1	1	1
1	1	1	1	2	1	1	1	1	1	1
1	1	1	2	2	1	1	4	4	4	1
1	1	1	2	1	1	4	4	4	4	1
1	1	3	3	3	1	4	4	4	4	1
1	3	3	3	3	1	1	4	4	1	1
1	3	3	3	3	1	1	1	1	1	1
1	1	3	3	1	1	1	5	1	1	1
1	1	1	1	1	1	1	5	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1

Raster data model.

© 2013 Pearson Education, Inc.

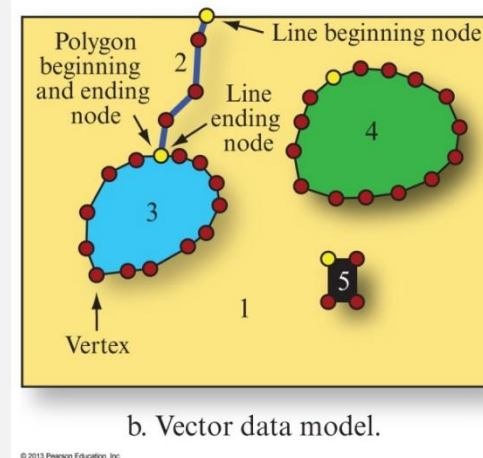
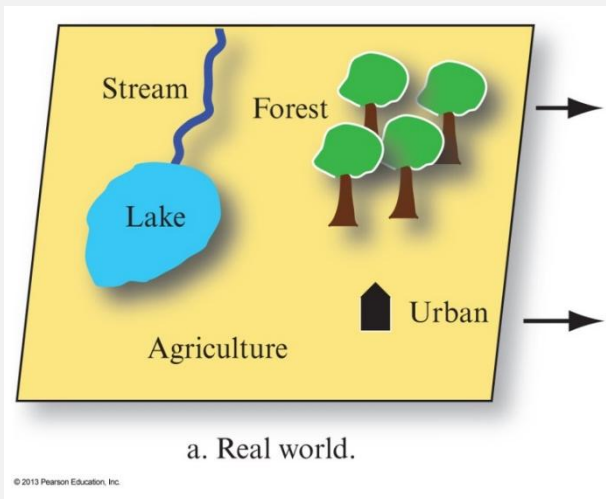
- Remote sensing
- Vector to raster conversion
- Raster data capture using scanners
- Interpolation

Learning Objectives

- ◆ Understand the primary and secondary techniques of data capture
- ◆ Understand concepts in remote sensing for raster data capture
- ◆ **Be familiar with techniques of vector data capture e.g. scanning, manual digitizing, vectorization.**
- ◆ **Familiarize with new sources of spatial data**

Vector Data Collection

- In Situ Surveying
 - GPS
 - LiDAR
- Primary Geographic Data Capture
- Raster to vector conversion
 - Digitization of historical hard-copy maps and other types of spatial data
- Secondary Geographic Data Capture



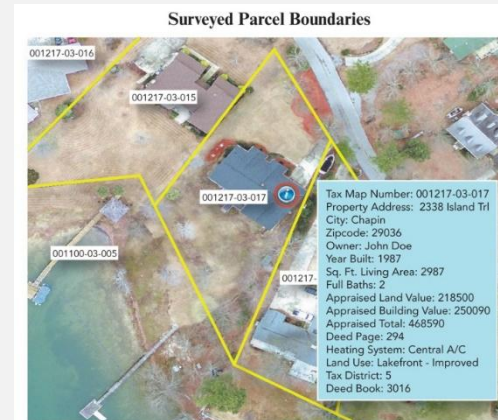
Vector Data Collection

■ Surveying

- Ground survey is a very **time-consuming** and **expensive** activity.
- Most accurate method for **large scale i.e small areas**.
- Surveying is typically used for **objects** that need to be **located accurately**.
 - E.g., capturing buildings, land and property boundaries
 - E.g., **ground control points** to georeference satellite images



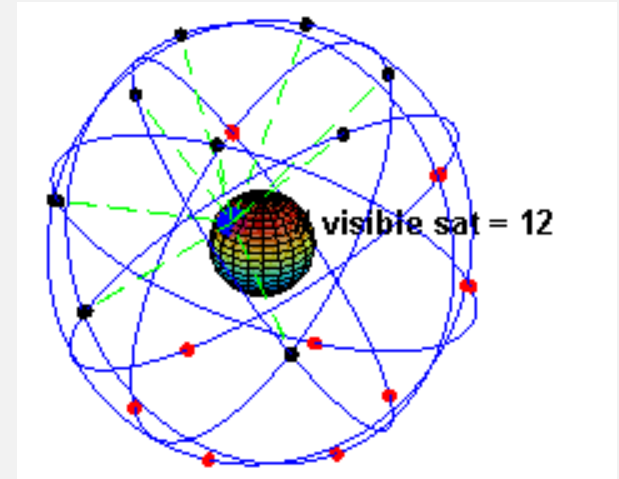
A tripod-mounted Total Station



Vector Data Collection

■ GPS (Global Positioning System)

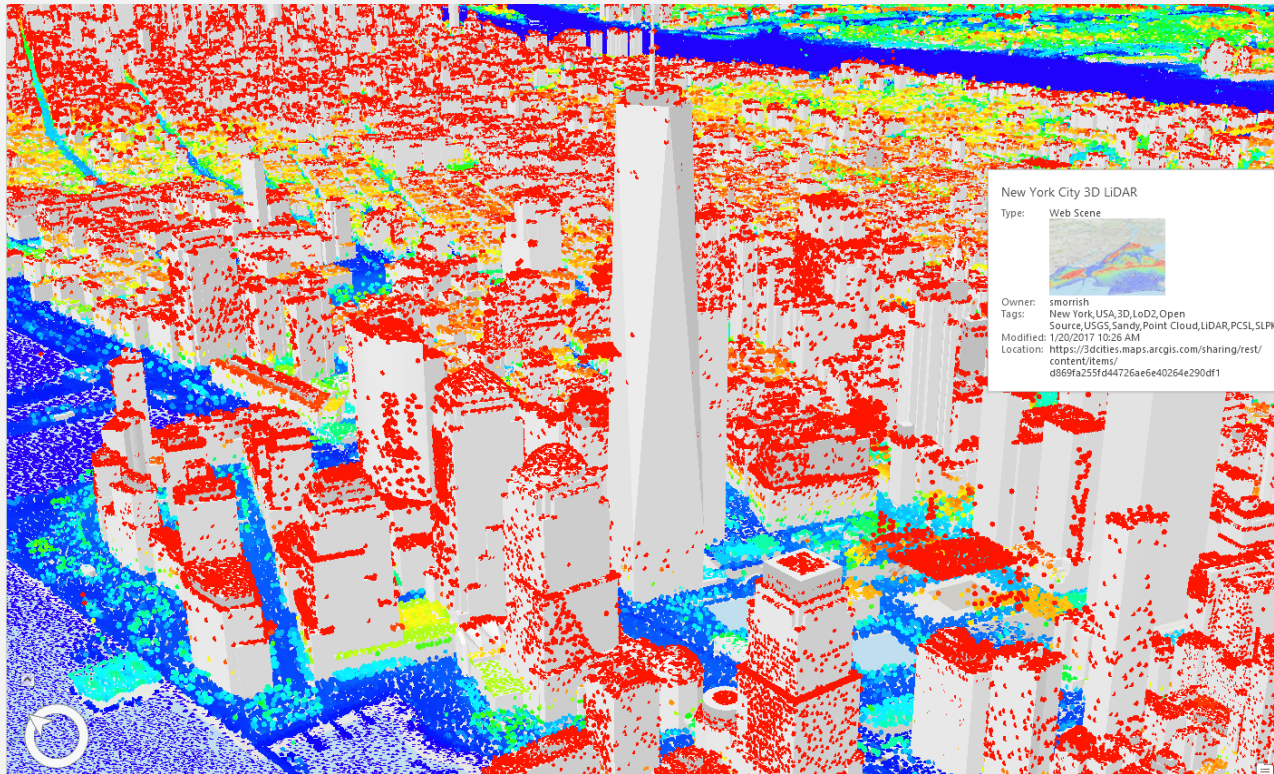
- Collection of satellites used to fix locations on Earth's surface
- Consists of **three** components
 - A space segment
 - A ground control segment
 - A user segment



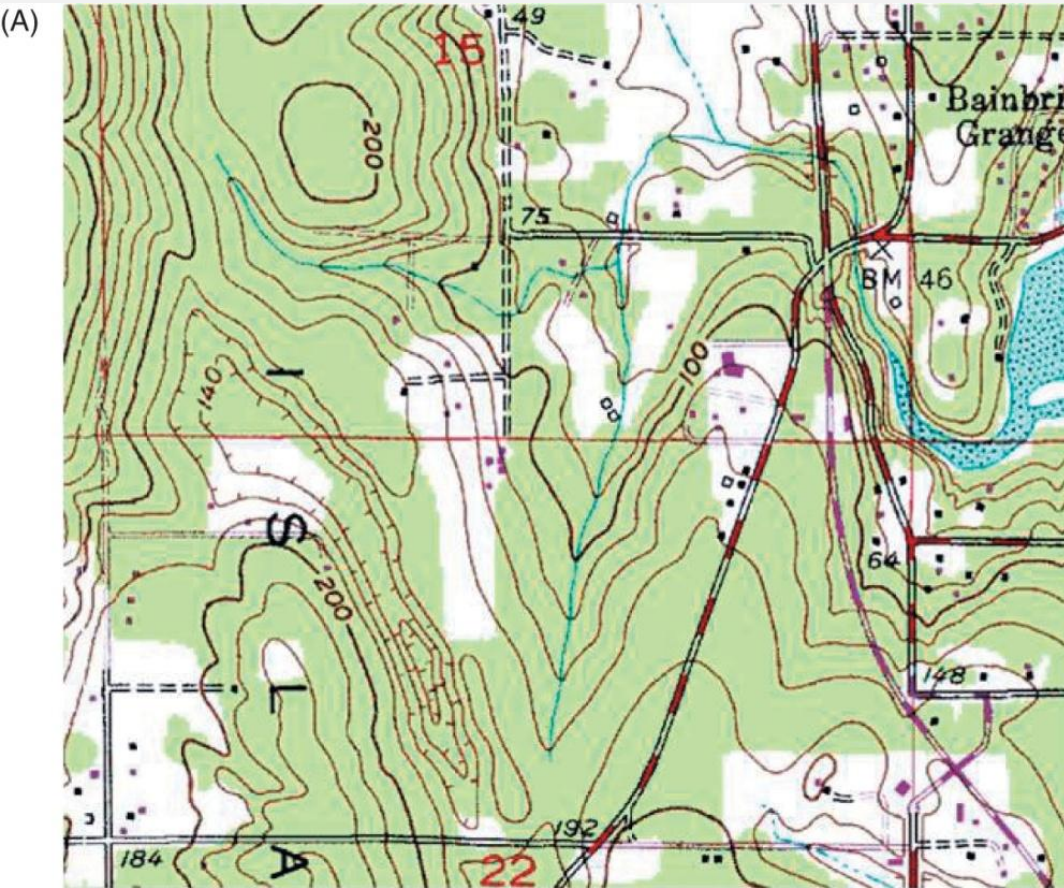
Source: <https://en.wikipedia.org/wiki/File:ConstellationGPS.gif>

Vector Data Collection

- **LiDAR** (Light detection and Ranging)
 - Active remote sensing
 - Point cloud (x,y, and z)
 - Remove tree canopies, buildings to create a bare earth point dataset

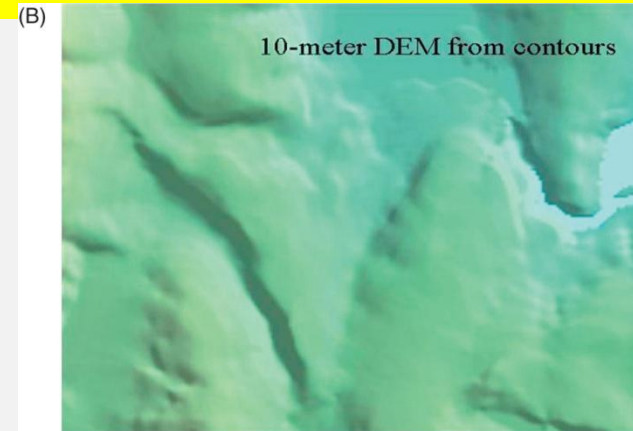


Comparison of **three datasets** for 1 square mile of Bainbridge Island, Washington State

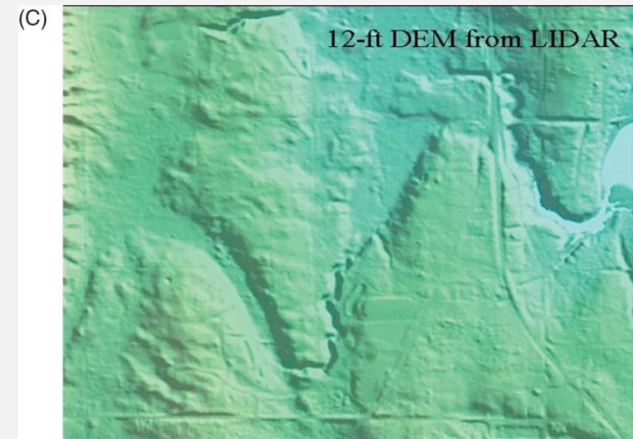


(A) scanned USGS 1:24,000 topographic map sheet

(From www.pugetsoundlidar.ess.washington.edu/About_LIDAR.htm)



(B) 10 m digital elevation model (DEM) derived from contours digitized from a map sheet;



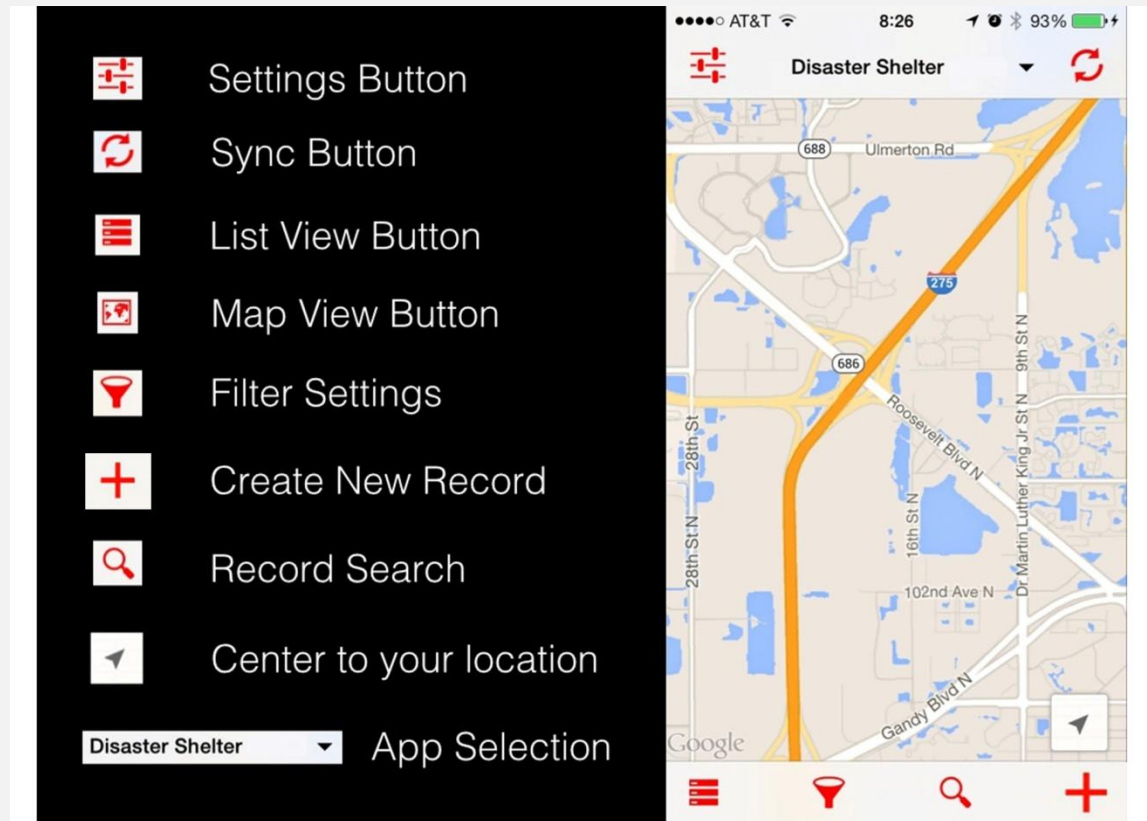
(C) 12 ft (365-cm) resolution DEM derived from a LiDAR survey

Vector Data Collection

- Mobile phone applications

Fulcrum

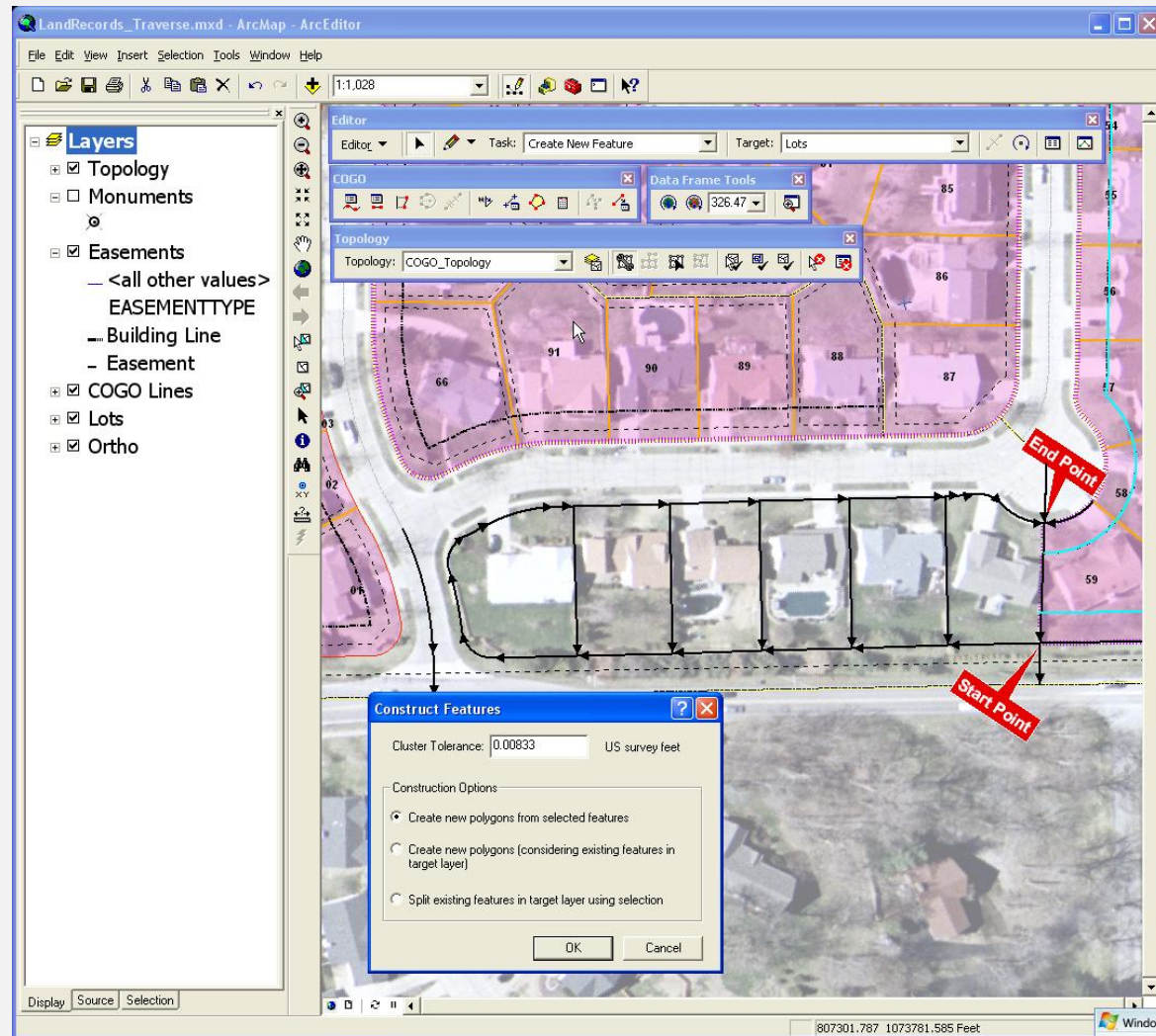
CitySourced



Vector Data Collection

- Collection of vector objects from **maps, photographs, and other hard-copy documents.**
- **Digitizing**
 - Digitizing tables and tablets
 - Heads-up digitizing and vectorization - Manual digitization by tracing a mouse over features displayed on a computer monitor

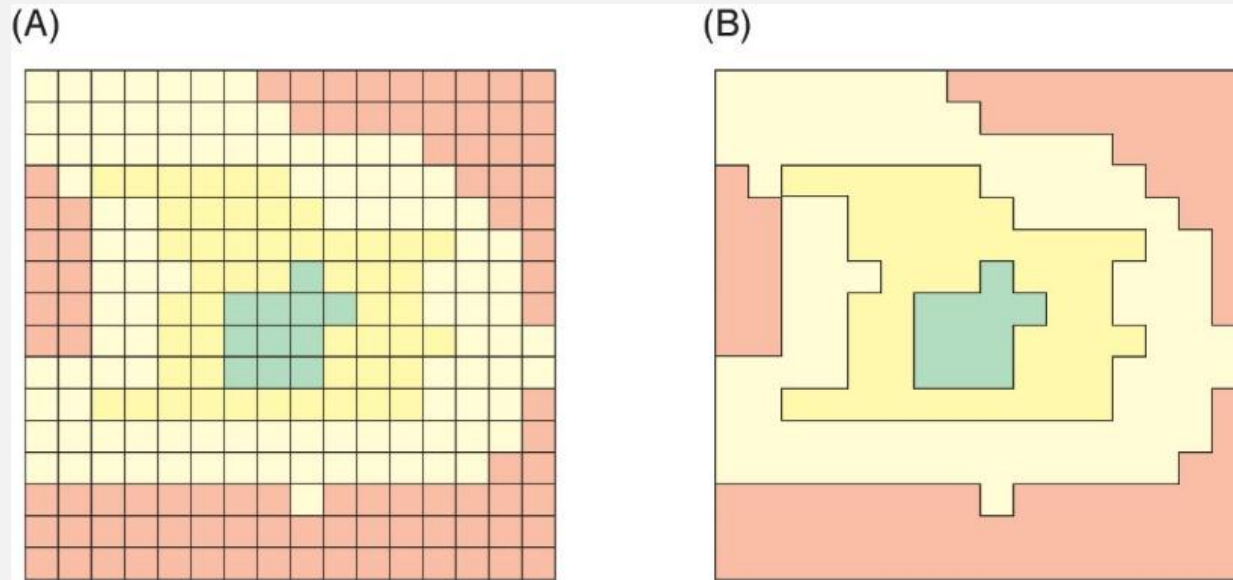
Heads-Up Digitizing



An example of raster background data (color aerial photography) underneath vector data (land parcels) that are being digitized on screen.

Vectorization

- Vectorization is the process of converting raster data into vector data.



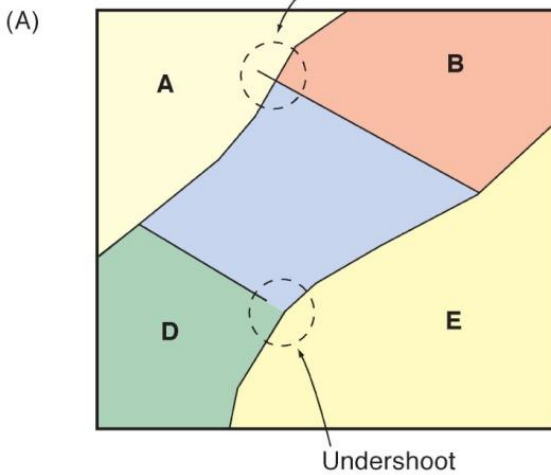
Batch vectorization of a scanned map: (A) original raster file; (B) vectorized polygons. Adjacent raster cells with the same attribute values are aggregated. Class boundaries are then created at the intersection between adjacent classes in the form of vector lines

Heads-up on-screen digitization using scanned maps or images

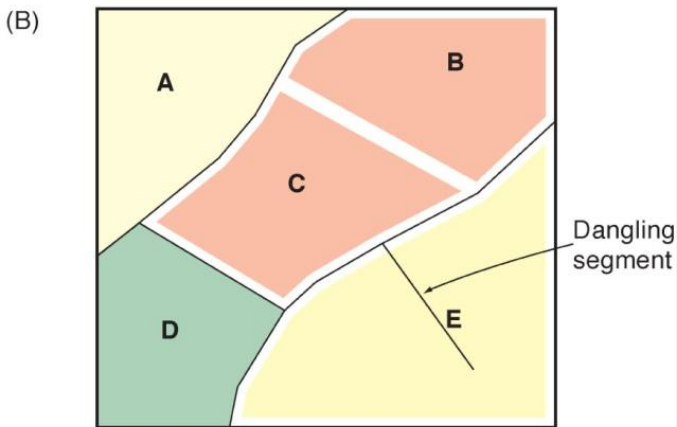
■ Steps

1. The **map** or image that is going to be used to extract the x, y-coordinates of point, line, and polygon features is **scanned**.
2. The scanned map or image is then **displayed on the computer screen**.
3. Georegistration.
4. Digitizing.

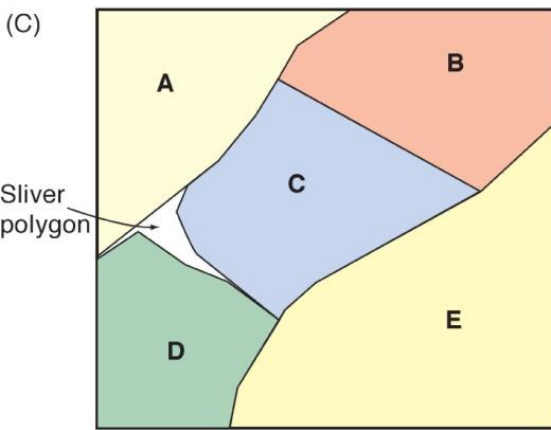
Examples of human errors in digitizing



(A) undershoots and overshoots



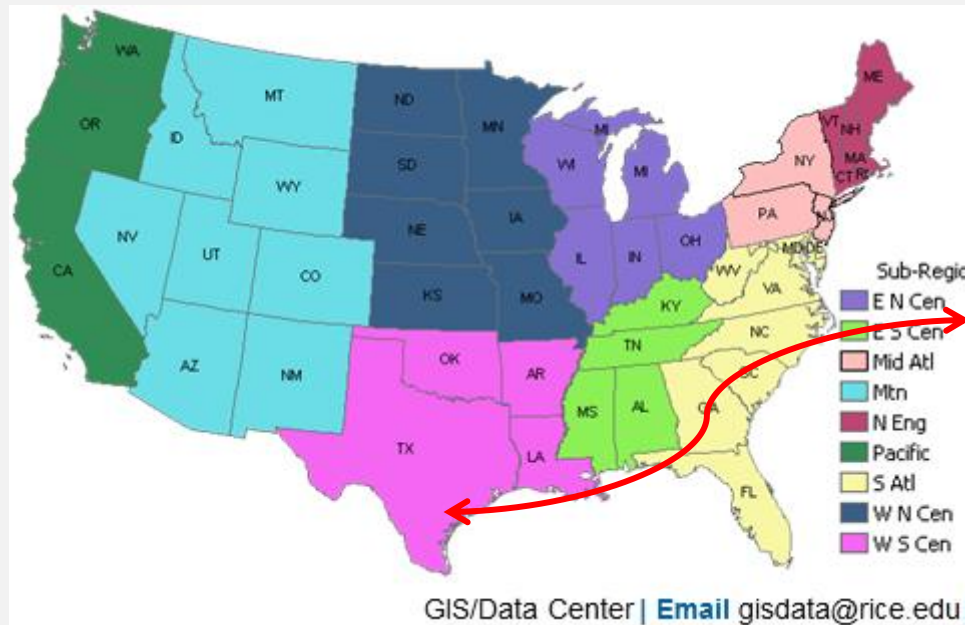
(B) invalid polygons



(C) sliver polygons

Capturing Attribute Data

- Attributes can be entered by direct data loggers, manual **keyboard** entry, **optical** character recognition (OCR) or, increasingly, **voice** recognition.
- An essential requirement for separate data entry is a **common identifier** (also called a **key**) that can be used to relate object geometry and attributes together following data capture.



STATE_NAME	DRAWSEQ	STATE_FIPS	SUB_REGION	STATE_AE
California	25	06	Pacific	CA
Ohio	26	39	East North Central	OH
Illinois	27	17	East North Central	IL
District of Columbia	28	11	South Atlantic	DC
Delaware	29	10	South Atlantic	DE
West Virginia	30	54	South Atlantic	WV
Maryland	31	24	South Atlantic	MD
Colorado	32	08	Mountain	CO
Kentucky	33	21	East South Central	KY
Kansas	34	20	West North Central	KS
Virginia	35	51	South Atlantic	VA
Missouri	36	29	West North Central	MO
Arizona	37	04	Mountain	AZ
Oklahoma	38	40	West South Central	OK
North Carolina	39	37	South Atlantic	NC
Tennessee	40	47	East South Central	TN
Texas	41	48	West South Central	TX
New Mexico	42	35	Mountain	NM
Alabama	43	01	East South Central	AL
Mississippi	44	28	East South Central	MS
Georgia	45	13	South Atlantic	GA
South Carolina	46	45	South Atlantic	SC
Arkansas	47	05	West South Central	AR
Louisiana	48	22	West South Central	LA
Florida	49	12	South Atlantic	FL
Michigan	50	26	East North Central	MI
Alaska	51	02	Pacific	AK

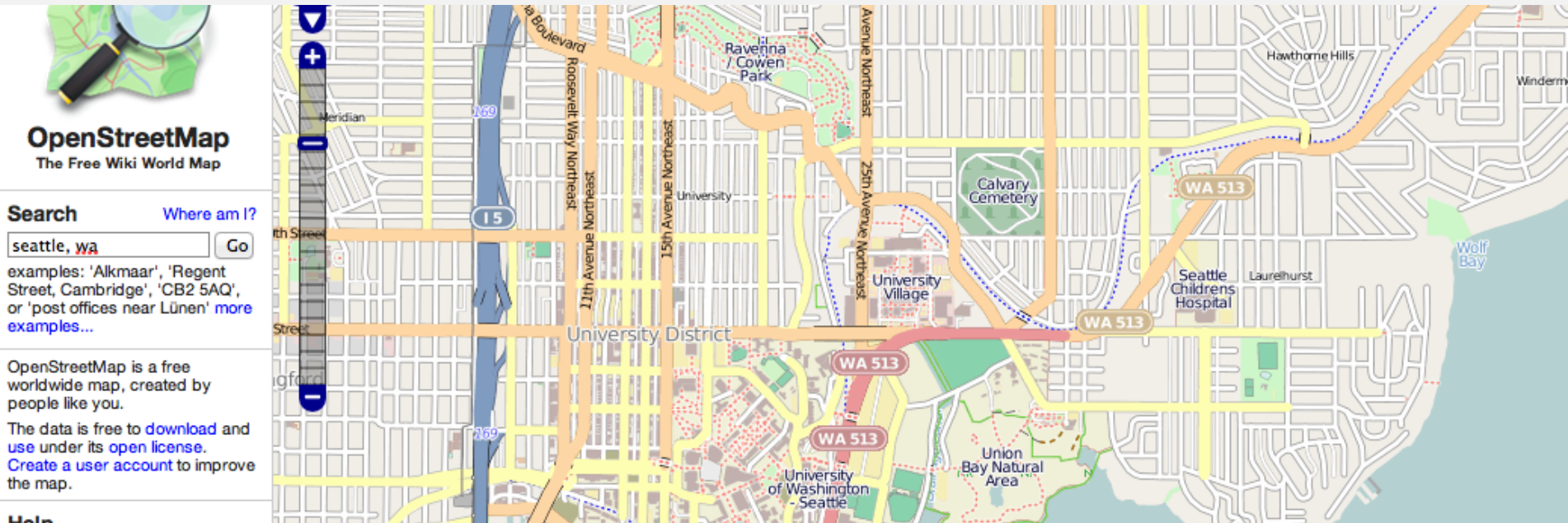
New Sources of Spatial Data

- Volunteered Geographic Information (VGI)
- Crowdsourcing
- Social Media

New Sources of Data

- Today, spatial data aren't just coming from formal organizations and authorities
- People without formal geographic training are creating it
 - Volunteered Geographic Information (VGI)
 - Neogeography
 - Crowdsourcing
 - Social media, internet logs, etc.
- **Raises issues of ethics and accuracy**
 - Can you trust this data?

New Sources of Data



OpenPaths



Get your data.

Access your location data from mobile devices, social services, and more.



Own your data.

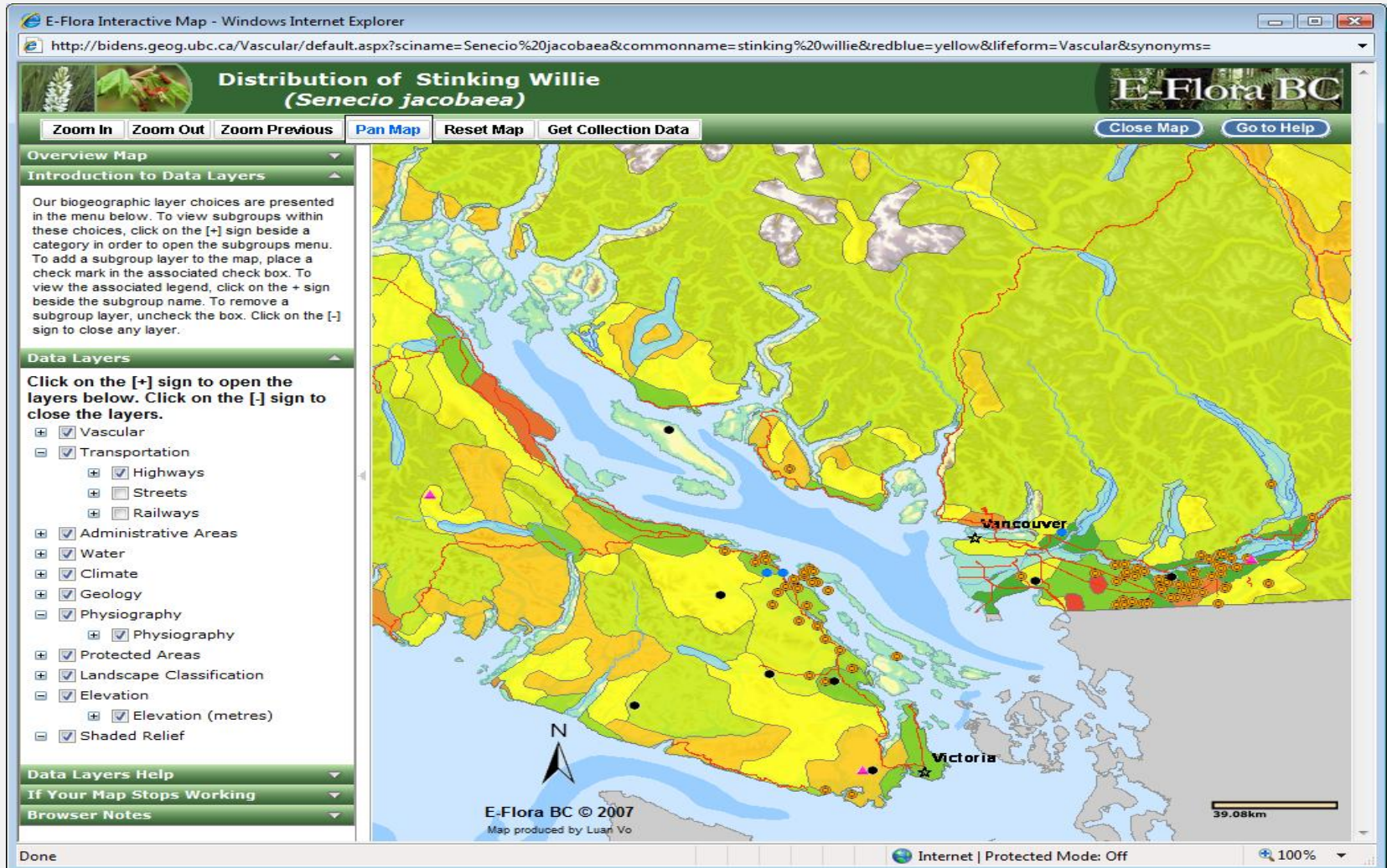
Your data is protected with the highest level of security so that you control access to it.

Sign in

New Sources of Data

Citizen-Centric Web-Based Data Collection

■ Web 2.0 technologies



New Sources of Data

- These new data sources can be both very exciting and powerful, and dangerous and deceptive.
- Think carefully about what your dataset might actually mean!

Authoritative Data Sources

- **Federal, State, Local Government**

- Census Data – <https://factfinder.census.gov/>
- NASA – <https://data.giss.nasa.gov/> | <https://neo.sci.gsfc.nasa.gov/>
- USGS – <https://earthexplorer.usgs.gov/>
- King County GIS – www.kingcounty.gov/services/gis.aspx

- **Commercial Providers**

- ESRI – www.esri.com/data/esri_data/index.html

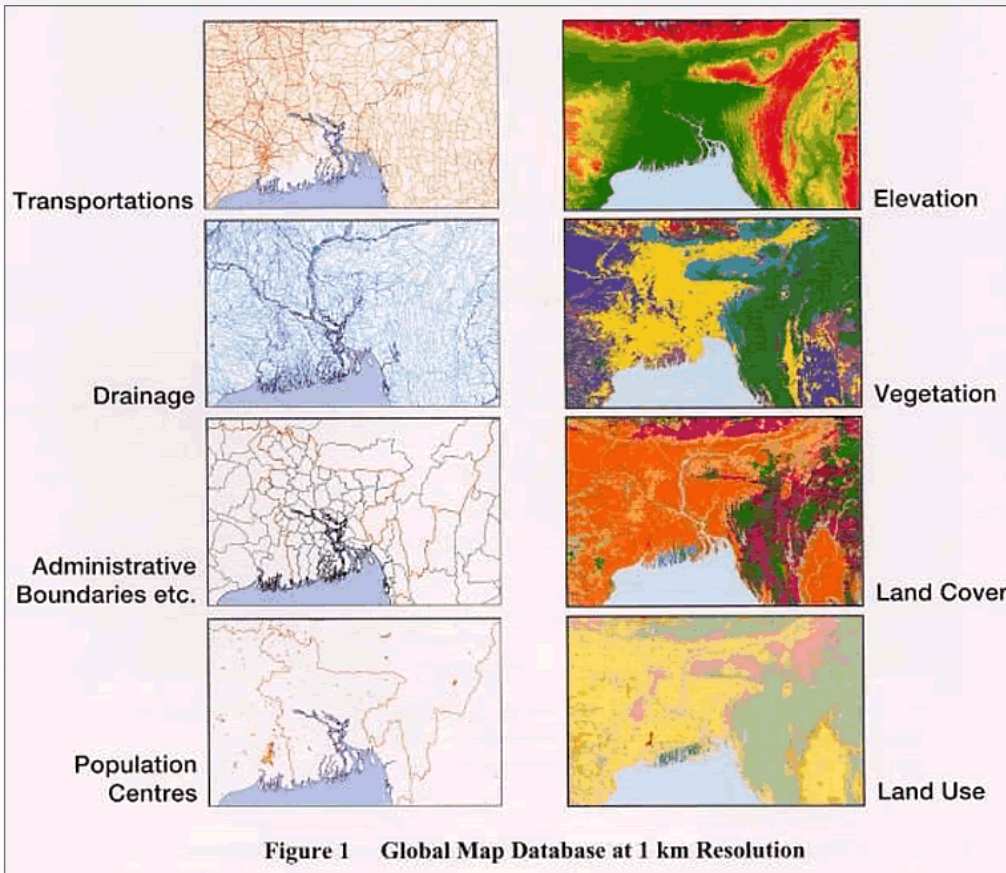
- **University-based Data Clearinghouses**

- UW – <https://wagda.lib.washington.edu>

- **International Organizations**


- Food and Agriculture Organization (FAO) – www.fao.org/geonetwork/
- United Nations Environmental Data Explorer – <http://geodata.grid.unep.ch/>

Authoritative Data Sources



- Natural Earth – North American Cartographic Information Society (NACIS)
- Global Map
- NCAR GIS Climate Change Scenarios – data used by IPCC
- Global 200 – ecosystem data from WWF
- Conservation GIS Data – The Nature Conservancy
- World Bank Geodata
- US National Atlas, NOAA

Authoritative Data Sources




Search

Data Groups

Sign in

Featured

9 Results




USA Counties

The layer, below 1:3m, presents the counties of the United States in the 50 states, the District of Columbia, and Puerto Rico. The layer, over 1:3m, presents the counties of the United States in the 50 states and the District of Columbia.

Layer Package by [esri](#) 3/16/2012

[Details](#) [Add](#)




USA Federal Lands

This layer package presents the federal and American Indian-owned land areas of the United States.

Layer Package by [esri](#) 11/2/2010

[Details](#) [Add](#)




USA Population by Zip Code

This layer presents the 2010 US Census population of the USA ZIP Code areas used by the U.S. Postal Service.

Layer Package by [esri](#) 3/19/2012

[Details](#) [Add](#)




USA Public Land Survey System (PLSS)

This layer presents the public land surveys of the United States.

Layer Package by [esri](#) 11/3/2010

[Details](#) [Add](#)




USA State Plane Zones

This layer presents the State Plane Coordinate System (SPCS) Zones for both the 1983 and 1927 North American Datum within United States.

Layer Package by [esri](#) 11/3/2010

[Details](#) [Add](#)



USA Zip Code Boundaries

This group layer presents five-digit ZIP Code areas used by the U.S. Postal Service to deliver mail more effectively. This detailed layer is approximately 146MB compressed.

Layer Package by [esri](#) 3/19/2012

[Details](#) [Add](#)

World Administrative Divisions


[Details](#) [Add](#)

World Cities

[Details](#) [Add](#)

Search: **population** Arranged by **Relevance** Results 1-100 of 721

Population Density for Africa in 2000 (people/




This documentation describes the fourth version of a database of administrative units with associated population figures for Africa. The first version was compiled for UNEP's Global Recertification Atlas (IMED 1007).

Layer Package by [consbio](#) 1/27/2011

[Details](#) [Add](#)

USA Recent Population Change 2010




This map shows recent population change from 2000 to 2010 in the U.S., by state, county, tract and block group.

Map Service by [AtlasPublisher](#) 2/19/2012

[Details](#) [Add](#)

Global rural population distribution (2000) (D

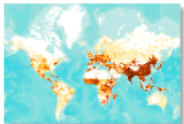


Distribution of rural population throughout the globe. Larger rural populations can be seen in developing nations whereas industrialized countries have higher urban populations. Global rural population fields were constructed using country-level demographic statistics from the World Resources Institute Earth Trends database. Rural dataset was developed by spatially distributing country level rural population data among DMSP-OLS nighttime stable lights imagery (Elvidge 1997a) and ESRI Digital Chart of the World populated places points (ESRI 1993). Rural population was spatially distributed equally among the DCW populated places points falling outside the DMSP-OLS city lights extents.

Layer Package by [consbio](#) 9/9/2010

[Details](#) [Add](#)

Global rural population distribution



Layer Package

Owner: [consbio](#)

Date modified: 9/9/2010

☆☆☆☆☆ 0 Ratings

Description

Properties

Comments (0)

Distribution of rural population throughout the globe. Larger rural populations can be seen in developing nations whereas industrialized countries have higher urban populations. Global rural population fields were constructed using country-level demographic statistics from the World Resources Institute Earth Trends database. Rural dataset was developed by spatially distributing country level rural population data among DMSP-OLS nighttime stable lights imagery (Elvidge 1997a) and ESRI Digital Chart of the World populated places points (ESRI 1993). Rural population was spatially distributed equally among the DCW populated places points falling outside the DMSP-OLS city lights extents.

Global urban population and total population datasets are also available from the same study. This dataset is in vector shapefile format, and each record corresponds to a 0.5 degree cell.

This layer package was loaded using [Data Basin](#). [Click here to go to the detail page for this layer package in Data Basin](#), where you can find out more information, such as full metadata, or use it to create a live web map.

1 2 3 4 5 Next

Authoritative Data Sources

- The website may not be the original creator of the data – especially in the case of **clearinghouses**
 - This means you **shouldn't cite clearinghouses!**
 - This also means that you shouldn't determine the usefulness of the data based on the clearinghouse.
- **Typical steps:**
 - Think about the **areal extent** and **resolution** of the data that you will need.
 - Select **attributes** and data that you will need.
 - Download and **unzip**.
 - **Read** the **metadata** and any read me files.
 - **Additional steps** may be needed to make your data ready for analysis – joining tables, geocoding, file conversion, formatting of attributes, etc.

Conclusion

- ◆ **Primary geographic data capture**
 - ◆ direct measurement of objects
- ◆ **Secondary geographic data capture**
 - ◆ the process of creating raster and vector files and databases from maps, photographs, and other hardcopy documents
- ◆ **Raster data capture** (Primary & Secondary)
- ◆ **Vector data capture** (Primary & Secondary)
- ◆ **New sources of spatial data**

Questions ?



<https://www.google.com/url?sa=i&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwhuvyghjz-AhU3DQIH2brj8QQjw6BqIEAU&url=http%3A%2F%2Fwww.cityofrockhill.com%2Fdepartments%2Finformation-technology-services%2Fmore%2Finformation-technology-services%2Fgeographic-information-systems-gis-%2Fgis-frequently-asked-questions&psig=AOvVaw2fELXAJbUy2Gw-bn50wY&ust=1531436220322311>

Upcoming

- **EXAM Week 6 (Check Syllabus)**
- **Lab 03 Due**