# 1 Creating a Custom Alberta Spatial Database

**Charlene Nielsen**
GIS Technologist, Biological Sciences
University of Alberta
www.biology.ualberta.ca/facilities/gis

## 2 …From Free Digital Sources

Through live demonstrations emphasizing **Alberta** content, this workshop provides an overview on how to acquire and troubleshoot some of the common issues that may arise while working with publicly available sources of spatial data.

## 3 Prior Knowledge Assumptions

- It is assumed that you already know…
  - … what a Geographic Information System (GIS) is
  - … what data models (vector, raster, TIN) are used to represent geographic features (point, line, polygon, surface)
  - … what basic formats of data are used in a GIS (thematic layers, satellite imagery, tables, etc.)
  - … how to navigate the World Wide Web and operate window-based computer software

## 4 Spatial Databases

- Collections of **spatial** data and related descriptive data, organized for efficient storage and retrieval.
- Important component in any GIS!
- Specialized structures include the personal geodatabase and database engines (not discussed in this workshop).

## 5 5 Steps of GIS Project Development

❶ Identify your objectives
- Define your question and criteria
- Determine the geographic boundaries for the region of interest
- Select the appropriate map scale (resolution, precision, accuracy)

## 6 5 Steps of GIS Project Development

❷ Assemble all required data
- Acquire existing data
- Create new data (digitize/field work)
- Check for errors and omissions
- Understand metadata

## 7 5 Steps of GIS Project Development

❸ Process data
- Import/convert
- Define coordinate systems and match projections
- Extract/modify existing features
- Update attribute information
- Generate symbology
- Document metadata

## 8 5 Steps of GIS Project Development

❹ Perform the analysis
- Select by location or attributes, create buffers, overlay layers

- Do calculations, apply complex modelling

❺ Present the results
- Interpret the output
- Design map layout, graphs, tables, reports

## 9 Develop the Spatial Database

- It is at **steps 2 and 3** that can take more time than necessary if you don't have the appropriate skills to
  - Import files
  - Work with various projections
  - Aggregate files, and basically get all your data "ducks" lined up in a row.
- Use standard conversion methods and procedures to
  - Create a well-organized spatial database containing data in a consistent format for a specifically defined study area.

## 10 Document the Spatial Database!

- Proper **documentation** of data
  - Is key to understanding the history of, determining the suitability of, and updating the database.
- **Metadata** = "data about the data"
  - Information about the content, quality, condition, and other characteristics of data.
  - Subject matter; How, when, where, and by whom the data was collected; Accuracy of the data; Availability and distribution information; Projection, scale, resolution, and accuracy; Data dictionary; Reliability with regard to some standard).

## 11 Workshop Content & Methods

- IPR and ©
- Useful websites and downloading hints
- File formats and importing
- Error and quality assessment
- Spatial referencing
- Aggregating
- Raster techniques
- Metadata documentation

## 12 IPR and ©

- Carefully research the ownership and restrictions of the data you wish to use.
  - Are the data free and cooperative or expensive and confidential?
  - What are the restrictive uses?
- Good to always assume that data will be subject to intellectual property rights and copyright.

## 13 IPR and ©

- At minimum, acknowledge the sources of the data!
- Suggested data citation format.

  *Author. Title from series [format]. Edition. Scale. Series, sheet number. Place of publication: publisher, date. <URL> (date accessed).*
- For more info see http://library.owu.edu/citing222.html.

## 14 Useful Websites

- Types of geographic information available for free on the internet include:
  - Topography and base layers
  - Imagery and elevation models

- Boundaries
- Census and statistics
- Natural Resources
- Etc.

15 ## Canada-Wide Datasets

1 • The Geomatics for Sustainable Development of Natural Resources (GSDNR) website provides links to:
- The Atlas of Canada
- GeoBase
- GeoGratis
- Toporama

2 • Other nation-wide sources include:
- Agriculture Canada
- Canadian Climate Impacts and Scenarios
- Natural Resources Canada
- Statistics Canada

16 ## Geomatics for Sustainable Development of Natural Resources (GSDNR)

**http://gsdnr.nrcan.gc.ca/data_e.cfm**
The Atlas of Canada
GeoBase
GeoGratis
Toporama

17 ## The Atlas of Canada

**http://geogratis.cgdi.gc.ca/nationalatlas/e_basemaps.html**

18 ## GeoBase

**http://www.geobase.ca**
Administrative Boundaries
Canadian Digital Elevation Data (CDED)
Landsat 7 Orthoimages

19 ## GeoGratis

**http://geogratis.cgdi.gc.ca**
Landsat, RADARSAT, Earth Observation Data
Atlas Vectors, National Scale Frameworks
CGIS (Canadian Geographic Information Systems)
Vector/Tabular Data

20 ## GeoGratis - Landsat

Landsat imagery are available through the GeoBase and GeoGratis (and searchable through the GeoConnections) web pages. Search for your desired image, but use the GeoGratis download directory to access the files.

21 ## GeoGratis - Landsat

22 ## GeoGratis - CLI

Search for CLI layers using the index map…

23 ## GeoGratis - CLI

… but download using the GeoGratis download directory (also obtain the handy legend files available for use in the GIS)!

## 24 Agriculture and Agri-Food Canada

**http://www.agr.gc.ca/pfra**
Prairie Farm Rehabilitation Administration (PFRA)

## 25 Agriculture and Agri-Food Canada

**http://sis.agr.gc.ca/cansis**
Canadian Soil Information System (CanSIS)
National Ecological Framework

## 26 Natural Resources Canada

**http://lsd.nrcan.gc.ca/english/clab-latc_e.asp**
Legal Surveys Division

## 27 Natural Resources Canada

**http://sts.gsc.nrcan.gc.ca/clf/digitalrelease.asp**
Geological Survey of Canada

## 28 Natural Resources Canada

**http://www.pfc.cfs.nrcan.gc.ca/monitoring/inventory/canfi/cnfi-overview_e.html**
National Forest Inventory
**http://www.glfc.cfs.nrcan.gc.ca/landscape/climate_models_e.html**
Regional, National and International Climate Modeling

## 29 Canadian Climate Impacts and Scenarios

**http://www.cics.uvic.ca/scenarios**
Gridded Observed Climatologies for 1961-1990

## 30 Statistics Canada - Census of Canada

**http://www12.statcan.ca**
**http://geodepot.statcan.ca/Diss/Data/Data_e.cfm**
Census Boundaries
Census Tables

## 31 Alberta-Specific Datasets

1. • Alberta government sites containing downloadable GIS data include:
   - Alberta Geological Survey
   - Agriculture
   - Community Development
   - Sustainable Resource Development
2. • Other provincial sites include:
   - Alberta Land Information Systems (AltaLIS)
   - Alberta Basement Transect Lithoprobe

## 32 Alberta Geological Survey

**http://www.ags.gov.ab.ca/GIS/gis_and_mapping.shtml**

## 33 AGRASID

**http://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/sag3249?opendocument**
Agricultural Region of Alberta Soil Inventory Database

## 34 Community Development

**http://www.cd.gov.ab.ca/preserving/parks/anhic/esa.asp**

ANHIC Environmentally Significant Areas

**http://www.cd.gov.ab.ca/preserving/parks/lrm/index.asp**

Parks and Protected Areas

## 35 Sustainable Resource Development

**http://www3.gov.ab.ca/srd/forests/health/surveydata.html**

Forest Health (Insect/Disease Surveys)

**http://envweb.env.gov.ab.ca/env/forests/fpd/fstgis.html**

Forest Protection (Historical Wildfires)

## 36 AltaLIS

**http://www.altalis.com/productsandsamples/index.html**

1:2 Million Base Dataset

## 37 ABT Lithoprobe

**http://www.lithoprobe.ca/transectsWebSites/ab/**

Alberta Basement Transect (ABT) Transect Lines

Township and Range

## 38 Other Useful Websites

- Various GIS data **"goodies"** – especially index map grids – can be downloaded from:
  - Centre for Topographic Information (NTS grid)
  - Environmental Systems Research Institute (UTM grid)
  - Landsat 7 (path/row index)
  - Alberta Basement Transect Lithoprobe (township/range)
- Excellent free global resources:
  - Land Processes Distributed Active Archive Center
    **http://edcdaac.usgs.gov/dataproducts.asp**
  - Global Landcover Facility (GLCF)
    **http://glcf.umiacs.umd.edu/data/**

## 39 NRCan CTI - NTS Map Sheet Index

**http://www.cits.rncan.gc.ca/cit/servlet/CIT/site_id=01&page_id=1-005-002-001-007-001.html**

National Topographic System (NTS) Map Sheet Index Grid

## 40 NRCan CTI - NTS Map Sheet Index

**ftp://ftp.ctis.nrcan.gc.ca/pub/ntdb/client/canada_nts/nts_canada/**

## 41 Geography Network From ESRI

**http://www.geographynetwork.com**

World Universal Transverse Mercator (UTM) Zones Index

## 42 Landsat 7 WRS-2

**http://landsat7.usgs.gov/wrsprshape.php**

(Worldwide Reference System) WRS-2 Path/Row Index

## 43 Downloading Tips

- Record the website URL!
- RIGHT CLICK and
  - Save Target As… (Windows 2000)
  - Copy to Folder… (Windows XP)
- Always download "readme" metadata files!
- If possible
  - Select files requiring the least amount of importing into your GIS software (i.e. no third-party software).
  - Select files in geographic coordinate systems for easier projecting later.

- File Transfer Protocol (FTP) sites may provide easier access for downloading.
- Set Internet browser options (e.g. passive FTP).

## 44 File Formats

- **Shapefile**
  - A **vector** data storage format that stores the shape and location (*.shp), attributes (*.dbf), and lookup index (*.shx) of geographic features in a set of related files having the same prefix that must be kept together in the same directory
- **Coverage**
  - A **vector OR raster** "folder-based" data storage format that stores geographic features (as arcs, nodes, polygons, and label points OR grid cells) in a self-named folder and the associated attribute tables in an info folder

## 45 File Formats

- **E00**
  - An interchange file format that is used to transfer **vector OR raster** coverages between computers (*. E00)
  - Requires importing using a specific application or utility with your GIS software
- **Image**
  - A **raster** representation or description of a scene usually produced by an optical or electronic device; e.g. satellite data, digital aerial photographs, and scanned maps
  - Many file formats – TIFF, JPEG, BIL, DEM, IMG, SID, etc. – some require importing using ArcToolbox or installation of a special extension

## 46 File Formats and Importing

- Refer to your GIS software help files for information on supported and imported data formats
- A very common file extension you may encounter is ***.zip** and will require *unzipping* prior to actual use:

**http://www.winzip.com**

- An excellent reference on "**File Formats in GIS, 3D & Imagery and What to Treat Them With…**" can be found at:

**http://www.grime.net/gistools/a-z.htm**

## 47 Error and Quality Assessment

- Not all data are created equal!
- Sources of error may involve:

  **Age – Consistency – Content – Generalization**

  **Measurement (Accuracy & Precision) – Scale – Etc**
- Explore your data layers by symbolizing and labeling features – getting to know the attributes associated with your layers should help point out any obvious errors (i.e. do the data display as expected?)
- Inspect for problems in
  - Edge-matching of adjacent map sheets
  - Sensor error or noise (e.g. satellite)
  - Missing data or blatantly wrong values

## 48 Spatial Referencing

- It is always a good idea to **project** all your layers to the same coordinate system and make sure to **define** the spatial reference (i.e. *.prj)
  - Today's GIS software has the ability to reproject on-the-fly allowing different layers to visually line up.
  - However, some tools and operations will balk or yield strange results when you use them

## 49 Spatial Referencing

- **Unprojected** versus **Undefined** data
  - Use data that are in an **unprojected** geographic coordinate system (GCS) when you don't need a high level of locational accuracy, or won't be performing queries based on location, area, and distance
  - Provide the definition to **undefined** data that do not have the auxiliary files (e.g. *.prj) that enable ArcMap to "read" and reference them with other data layers
- **Georeferencing**
  - Align the raster to the desired map coordinate system of existing spatial data (e.g. target layer or paper map) by identifying a series of ground control points – of known x,y coordinates – that link locations on the source raster with locations in the target data

## 50 Aggregating

- Some datasets contain a wealth of spatial information in multiple map sheets (e.g. CLI, AGRASID) or layers (e.g. Atlas, Wildfire).
- Various vector procedures allow you to dissolve, merge, clip, intersect, and union to combine them into single layers based on similar geographic features.
- Aggregating makes them easier to work with, especially when symbolizing and reprojecting, and is often necessary for large area analyses (but may also be subject to available computer resources).

## 51 Raster Techniques

- **ESRI Spatial Analyst's Functional Reference** has help on how to merge, mosaic, and "clip" (along with many other functions) – useful for CDED data

- **ESRI Support Center** has technical articles for working with multiband raster images

http://support.esri.com/index.cfm?fa=knowledgebase.techarticles.articleShow&d=22526
http://support.esri.com/index.cfm?fa=knowledgebase.techArticles.articleShow&d=22005

## 52 Metadata Documentation

- Without data documentation, the spatial database will be difficult to understand, use, and update.
- Always keep a notebook with you to jot down notes on sources (e.g. website) of your data, informal metadata, details on any processing you perform, information on useful help files, step-by-step instructions on the analyses you execute to answer your research question(s), etc.
- You may take advantage of built-in GIS software metadata tools or implement your own system using commonly available software, such as Notepad, MS Word, Excel or Access.

## 53 Metadata Documentation

- **Federal Geographic Data Committee (FGDC)**

**http://www.fgdc.gov/metadata/metadata.html**

  - Identification_Information
  - Data_Quality_Information
  - Spatial_Data_Organization_Information
  - Spatial_Reference_Information
  - Entity_and_Attribute_Information
  - Distribution_Information
  - Metadata_Reference_Information

- **ESRI Profile of the Content Standard for Digital Geospatial Metadata**

**http://www.esri.com/metadata/esriprof80.html**

## 54 Text References


- Burrough, Peter A. and McDonnell, Rachael A. 1998. **Principles of Geographic Information Systems**. Oxford University Press Inc. New York, NY. 333 pp.
- Davis, David E. 2000. **GIS for Everyone**. Environmental Systems Research Institute, Inc. Redlands, CA. 184 pp.
- Longley, Paul A., Goodchild, Michael F., Maguire, David J., and Rhind, David W. 2001. **Geographic Information Systems and Science**. John Wiley & Sons, Ltd. Chichester UK.
- Minami, Michael. 2000. **Using ArcMap**. Environmental Systems Research Institute, Inc. Redlands, CA. 528 pp.
- Tucker, Corey. 2000. **Using ArcToolbox**. Environmental Systems Research Institute, Inc. Redlands, CA. 105 pp.
- Vienneau, Aleta. 2001. **Using ArcCatalog**. Environmental Systems Research Institute, Inc. Redlands, CA. 286 pp.
- Wadsworth, Richard and Treweek, Jo. 1999. **Geographical Information Systems for Ecology: An Introduction**. Addison Wesley Longman Limited. 184 pp.


## 55 More Resources

- Access the companion website to this workshop by visiting:

**http://www.biology.ualberta.ca/facilities/gis/index.php?Page=2980**

- Additional GIS software instructions are available at:
  **http://www.biology.ualberta.ca/facilities/gis/index.php?Page=485**

56 ## Questions?

57 ## Demonstration

- Develop a spatial database containing recreation-related themes in the Pembina river watershed
  - Compare base layers for the province of Alberta
  - Include watersheds, CLI recreation capability, Alberta parks and environmental areas, geology, elevation, and imagery
  - Skills demonstrated can be applied to a wide range of datasets