Visions for the Digital Library

Michael F. Goodchild
University of California
Santa Barbara
A geographer's perspective

- Space as an organizing dimension
- Space as a basis for understanding
- Metaphors for working with space
- Digital Earth
Library search/sort dimensions

- Shelf order
- Card catalog
  - author, title, subject
  - all discrete, unidimensional
- Automation (1)
  - fast search
  - key word in context
New dimensions

- Location, time
  - continuous
  - multidimensional
    - begin and end
    - bounding rectangle

- How often do people need to search by location?

- What role do search mechanisms play in building collections?
The geolibrary

- What have you got about *there*?
  - all relevant information about a place
  - a library whose primary search mechanism is location

- Information footprints
  - beyond maps and images
  - photographs, guidebooks, reports, music, the Bible

- Emergency management
Find a Spot on Earth by  ▶ Place Name ▶ Map Search ▶ Geo Coordinates

Database Rows covering shaded area 149,757,510
Total size of Images 628.57 GB

Coverage graphic courtesy of the United States Geological Survey

The green shading on the map identifies the locations covered by images stored in the Microsoft® TerraServer database. Click on a country or state/province you are interested in viewing.
"Distributed Geolibraries: Spatial Information Resources"

www.nap.edu
Organizing information by location

- Information with a geographic footprint
- Organizational metaphors
  - the desktop, office, workbench
  - the surface of the Earth
Research challenges

- Defining footprints
  - fuzzy, vernacular
- Mapping between georeferencing methods
  - the gazetteer
- Search over a distributed archive
  - search engines
  - object-level metadata (OLM)
  - collection-level metadata (CLM)
CLM of the Alexandria Digital Library
Research challenges (2)

- Approaches to CLM
  - by data type
    - ortho.mit.edu
  - by area of the globe
    - SRI's Digital Earth
  - the one stop shop
    - www.fgdc.gov
  - a new generation of search engines
    - identifying footprints
Everything digital
- bags of bits
- media-independent

Information for education: ADEPT
- texts, articles
- data
- methods, models, results of analysis
- personal storehouses, shared spaces
- evaluations
A complex of geographic information technologies

- Remote sensing
- Point of sale systems
- Cellphones
- GPS
- WWW
- Census geographic services
- Geographic data clearinghouses
Definition: geographic information systems

- Technology for processing geographic information
- Processing: acquisition, compilation, sharing, communication, publication, analysis, modeling, visualization, storage, archiving
- Five general principles
General principles:
1. Integration

- Linking data through common location
  - the layer cake
- Linking processes across disciplines
  - spatially explicit processes
  - e.g. economic and social processes
    interact at common locations
  - Center for Spatially Integrated Social Science (www.ncgia.ucsb.edu)
2. Spatial analysis

- Social data collected in cross-section
  - longitudinal data are difficult to construct
- Cross-sectional perspectives are rich in context
  - can never confirm process
  - though they can perhaps falsify
  - useful source of hypotheses, insights
The Snow Map of Cholera Incidence in the Area of Broad Street, London, in 1854. The contaminated water pump is located at the center of the map, just to the right of the D in BROAD STREET.
3. Spatially explicit theory

- Theory that is not invariant under relocation
- Spatial concepts (location, distance, adjacency) appear explicitly
- Can spatial concepts ever explain, or are they always surrogates for something else?
\[ I_{ij} = \frac{E_i A_j f(d_{ij})}{\sum_k A_k f(d_{ik})} \]
4. Place-based analysis

- Nomothetic - search for general principles
- Idiographic - description of unique properties of places
- An old debate in Geography
The Earth's surface

- Uncontrolled variance
- There is no average place
- Results depend explicitly on bounds
- Places as samples
- Consider the model:
  \[ y = a + bx \]
5. Knowledge and policy

- Policy requires the projection of general knowledge in spatial context
  - the implications of this process in this location
  - alternative futures visualized under local circumstances

- GIS combines the general (processes, models, algorithms) with the specific (database of local details)
Metaphors of GIS

- Data analysis
  - the machine in the service of the user
  - database, processor, user interface

- Communication
  - technology mediating communication
  - the network is the computer
Fundamental parameters of the communication metaphor

- **Technical**
  - bandwidth, speed, access, reliability
  - interoperability, semantics, understanding

- **Media and formal structures**
  - visual, auditory, tactile
  - speech, text, imagery, maps, tables…
  - facilitating or imposing
Precise or vague?

- Precise: "the air temperature at 120W, 34N at local noon on January 14 was 24C"
- Vague: "it’s very pleasant today in Santa Barbara"
It's very pleasant today in Santa Barbara

Spoken word

Text

Picture

$x, y, T$
Data models

- An infinite number of atoms
- Compression, abstraction, generalization
Information lost to the representation

- All sub-polygon spatial variation
- All within-decade temporal variation
- All identities
  
  <R, number>
  
  <xy, xy, xy, xy, ..., R>
Challenges of GIS

- How to characterize what is missing
  - error, accuracy, uncertainty

- How to choose the best representation
  - confounding influences

- How to support many data models in a single software package
Weaknesses of GIS

- There are too many possible data models
  - special-purpose GIS
  - lack of interoperability
- Difficult to add data models retroactively
Moonshots

- Targets for 2010
  - practically motivated
- Multidisciplinary research issues in GIScience
The Digital Earth metaphor

- Learning about the Earth
  - significance to many disciplines
  - information
  - processes and scenarios

- The organizational metaphor
  - desktop, office, workbench
  - Earth itself
    - immersion
    - dynamics
Digital Earth background

- The Gore speech (Los Angeles, 1/98)
  - “I believe we need a 'Digital Earth'. A multi-resolution, three-dimensional representation of the planet, into which we can embed vast quantities of geo-referenced data.”
  - physical and social

- NASA's Triana
  - the sunlit Earth
Digital Earth

A very visual Earth explorer that lets Scientists – both young and old – examine information about the Earth to learn how the forces of biology and geology interact to shape our home planet.
Building the Digital Earth

- Access to data
  - what do you have about here

- Tools for visualization
  - 4-5 orders of magnitude of zoom
  - user-centered
  - beyond the visible
  - analysis, modeling, simulation
Two moonshots: by 2010

- Possible to assemble all relevant information about a place from the resources of the Internet
- Possible for a child to learn about Earth and its infinite variety through the medium of digital technology
Final comments

- Geography as an organizing dimension for information
- Geography as a basis for understanding
- Communicating about geography
- Digital Earth moonshot
- Very active research, development, policy community
Some WWW sites

- alexandria.ucsb.edu
- www.ncgia.org
- www.digitalearth.gov