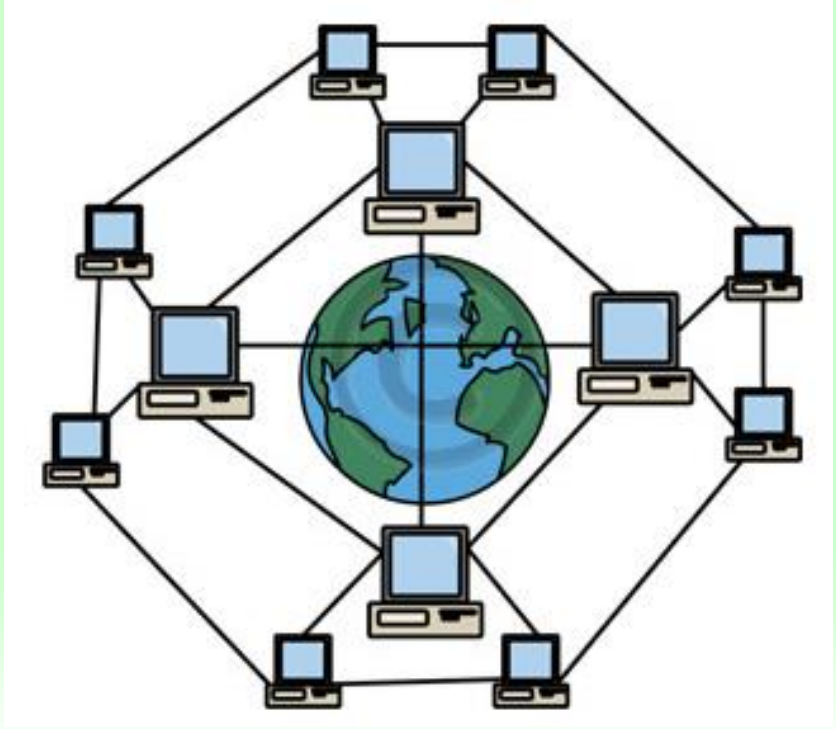

http://classroomclipart.com/cgi-bin/kids/imageFolio.cgi?action=view&link=Clipart/Animated_Clipart/Computers&image=Technology_12.jpg&img=&tt=

GEOG 482/582: GIS Data Management

Lesson 1: GIS Data Management Technology and Users

Overview – lesson in three parts

Learning Objective Questions:

Part A. Course Logistics

1. What are the course logistics?

Part B. Scoping GIS Data Management Technology

2. What kinds of information technology are important to GIS?
3. What are approaches to spatial data management?
4. Why employ user needs assessment for database systems?

Part C. Concepts and Architecture of Database Systems

5. What are some of the basic concepts of Spatial DBMS as per Y&H Figure 2-1?
6. What are fundamental approaches to data processing?

Lesson Preview

Learning objective questions act as the lesson outline.

Questions beg answers

Assignment 1: Forming the geodatabase schema

Part A. Course Logistics

1. What are the course logistics?

Course in General

Course Schedule

Student Assessment
- Lab Assignments
- Exams

Key Terms
Logistics

Course in general

GIS Data Management – three main thrusts

- Databases – a data perspective
- Data management technology – a software/hardware perspective
- Enterprise approach to data management -
  a people/organizational perspective combined with above

This is NOT the database (python) programming course.
That is Geog 465.

Key Terms
Course thrust

Course Schedule

Lectures/Lessons by session throughout quarter.
M&W lessons. Friday are integration sessions.

Readings throughout quarter – two required texts, multiple .pdfs

Y&H - Albert K. W. Yeung and G. Brent Hall (Y&H) 2007. *Spatial Database Systems*, Springer, Dordrecht, Netherlands. Referred to as Y&H in course schedule.

A&Z - David Arctur and Michael Zeiler (A&Z) 2004. Designing Geodatabases: Case Studies in GIS Data Modeling, ESRI Press
Referred to as A&Z in course schedule

Key Terms
Y&H
A&Z

Student Assessment

Exams – two exams
- Exam 1 – 100 points
- Exam 2 – 100 points

Lab assignments – Six lab assignments
- 200 points total

Key Terms
Exams
Lab assignments

Part B. Scoping GIS Data Management Technology

2. What kinds of information technology are important to GIS?

- Several technologies have always been important to compose GIS
- Data Collection technology
- DBMS is foundation for storing large amounts of data
- Spatial analysis – explore relationships among phenomena
- Map Visualization / Geovisualization – for depicting those relationships

Key Terms
Three technologies:
DBMS
Spatial analysis
Map visualization

What about other, more recent, technologies

- Data Communications
- Decision Science
- Mobile collection and displays
- High performance computation

Key Terms
Other technologies

Mainstream and GIS Technology coming together

- Technology streams merge
- Databases recognized as important asset to organizations
- Map-centric databases promoted fast retrieval
- Phenomenon-centric databases promoted better analysis
- 1st version of ArcInfo used the Info datafile manager, 1982

Key Terms
Technology streams

Databases existed before GIS

- When GIS adopted DBMS approach, data integration was easier
- Y & H emphasize distinction between GIS and “robust database system”
- Y & H argue database management system (DBMS) not as general
- Y & H Figure 1-2 – new working relationship between DBMS and GIS.

Key Terms

Three dimensions of spatial database systems are data-based and user-centric

- Stewardship – custodian of information mandated by law
- Sharing – spread the cost across units with organization
- Commodification – ownership in data and information, sales of value

Key Terms
Data-based
User-centric
Stewardship

3. What are approaches to spatial data management?

Four user-focused approaches …

Single-user

Personal database
single user database management system (DBMS)

Workgroup

Department/Application database
workgroup sever DBMS

Enterprise database

multiple workgroups large server; several servers DBMS

Consortium database

multiple enterprises; server archiving DBMS

Key Terms
User-focus

Migration Strategy from single user GIS to SDBMS

Review Y&H Figure 1-4

Two major contexts:

Computing & Interface

a) Local data
b) Enterrprise data server
c) Interoperable databases

Database & Application

a) Desktop computers
b) Enterprise GIS
c) Global, national, and spatial data infrastructure

Key Terms
Migration strategy

Migration Strategy from GIS to SDBMS

- Somewhat of misnomer
- Robust core of GIS being DBMS is certainly a trend that continues
- Technologies mature and grow together
- Project management becomes increasingly important
- Decision support apps become increasingly important

Key Terms

4. Why employ user needs assessment for database systems?

User Needs Assessment (UNA) is a term similar to:

- Systems analysis – understand workings of system by decomposing the work activity
- Functional requirements study – what capabilities are needed
- Business functions – what business functions are to be supported
- Requirements engineering – specification of functional capabilities needed

Key Terms
User needs assessment

User Needs

User need defines what a typical user requires of a database

Challenges vary according to

- Problem comprehension – what is understood about the world
- Communication – people speak to one another to gain meaning
- Problem and system structure match under conditions of continuous change
- Interdependency of needs among applications

Key Terms

Objectives of User Needs Assessment

- Facilitate knowledge transfer and integration between users and designers
- Identify business problems and associated business function activities
- Provide coherent framework for business information objects and activities
- Evaluate the functional feasibility of proposed database
- Structured and systematic approach for identifying desired system functions
- Reconcile different user needs associated with business functions and activities
- Outline framework for sharable information services and interoperability

Key Terms

User Needs Assessment Methods and Techniques

- **Document analysis** – scour documents for insights about what is to be done
- **Job observations** – participant observation of information work activity
- **Questionnaires** – ask questions about activities
- **Interview**s – face-to-face discussion about information activity
- **Focus group discussion** – brainstorming sessions get people to share insights

Key Terms
Document analysis

Steps and Workflow of User Needs Assessment

Y&H Figure 8-3 portrays user needs assessment in context of systems development and database development.

Two parts:

Needs development (next slide more detail)

- Elicitation
- Analysis
- Specification
- Verification

Needs management

- Rapid Prototyping

Plus the need for documentation of the outcomes from above

Key Terms

Workflow

Needs Development and Management

Development

- Elicitation – collecting information needs from users that imply data and software capabilities
- Analysis – decomposing the high-level needs statements into details that describe the features of the system
- Specification – document the findings
- Verification – checking the elicitation, analysis, and specification to make sure expressions are representative of the needs

Management

- Organize the documents so one can update them
- Maintain the documents over time for currency

Key Terms

Part C. Concepts and Architecture of Database Systems

- Problems being recognized as more complex
- Being able to accomplish more with same resources
- Data sharing continues to be important for workgroups, organizations, and consortium of organizations.
- Furthermore, gaining a better understanding of the changes in both space and time requires a better means of storing data about complex problems

Key Terms
Workgroup units
consortia

5. What are some of the basic concepts of Spatial DBMS as per Y&H Figure 2-1?

- Problem space – the realm of the complex problem
- Data model – the constructs, operations, and constraints framing the DBMS
- Database schema – data description language
- Database engine – software implementation of the schema
- Database - instantiations of data
- User interface – software through which people work
- Middleware – a layer of software between two others (DBMS engine and applications)
- Development tools and Application programs – task focused software

Key Terms
Problem space
Data model
Database schema
Middleware

Classification of Database Systems

Y&H Table 2-1

Data models

- Relational, object-oriented systems, object-relational systems

Primary database functions

- Data storage; transaction systems; decision support system

Nature of data

- Spatial information systems; non-spatial information system

Objectives of information

- Custodial systems and data warehouse
- Project-oriented systems

Hardware platforms and system configurations

- Distributed systems
- Desktop Systems

Key Terms

6. What are fundamental approaches to data processing?

File-based approach…
- data are stored separately in physical files (Y&H Figure 2-2a)

- Applications read the files directly

Key Terms

Data-based approach…
- one (or more) physical files contain logical separators like “tables” (Y&H Figure 2-2b)

- Application reads a schema, and fetches data.

Hardware and software architecture perspectives

- Centralized and Distributed Database Architecture
- Client-server computing
- Database Software
- Web-based Database Architecture

Key Terms
Centralized
Distributed
Client-server

1970's - Centralized Database Architecture

Y&H Figure 2-6 Centralized Database System Configuration

- Processing is performed on a single computer with primary memory (random access memory – RAM) and secondary memory (disk)
- Access is via "dumb terminal", that is terminal has no onboard memory
- All input/output is performed on single host computer

Key Terms
Centralized database

1980-90's - Distributed Database Architecture

Y&H Figure 2-7 Distributed Database System Configuration

- Nodes – computer, operating system and database system software
- Network cards – support protocols for data exchange
- Communication Protocol –Transmission Control Protocol / Internet Protocol (TCP/IP)
- Data Processor/Manager – Data engine for storing and retrieving data
- Transaction Processor – Controls database transactions

Key Terms
Distributed database

2000's – Client-Server Computing

- Software runs on several client computers together with one or more servers
- Client – software used for interaction with database
- Application server – applications programs are stored and executed
- Database server – programs for storage and retrieval of data

Key Terms
Client-server

Client-Server Options

Client-Server functions are split between clients and servers

Y&H Figure 2-8a-e

- Two tier client-server architecture (columns a - d)
- Three tier client-server architecture (column e)
- Thin client / Fat server architecture (columns a & b)
- Fat clients / Thin server architecture (columns c & d)

Key Terms
Client-server

Database Software – different modules in database system

Y&H Figure 2-9 – software layers

Database engine – operating system

SQL – database analysis routines

Network Middleware – client-side applications

Compiled database application interfaces – dynamic data analysis

Key Terms
Database engine
SQL
Network middleware

Web-based Database Architecture

Y&H Figure 2-10 (Numbers indicate data flow)

• Client submits TCP/IP “packet” request (1)

• using hypertext markup language (HTML) embedded in hypertext transfer protocol (HTTP) which is interpreted on Web server (2)

• as a script page submitted to a server-side extension, also called web-to-database middleware (3)

• processes database queries using Common Gateway Interface (CGI) or API protocols (4)

• that are interpreted by database server engine (5)

• then dynamically generates (wrapped back up) an HTML-formatted page that includes the data retrieved from database (6)

• returning to Web Server (7) for HTTP wrapping which

• forwards the HTTP/HTML packet wrapped in TCP/IP protocol with data (8)

• to the client (9) that can interpret all the packaging and present the data

Key Terms

HTML
HTTP
Web server

Summary

We addressed six learning objectives and learned about:

1. course logistics
2. kinds of technology that are important
3. approaches to spatial data management
4. character of user needs assessment
5. basic concepts of Spatial DBMS
6. fundamental approaches to database processing in terms of several system architectures

Contact Prof. Nyerges at nyerges@uw.edu if you have questions or comments about this lesson.

GEOG 482/582: GIS Data Management

Lesson 1: GIS Data Management Technology and Users