
 **Test Your Tech**

Blogging is:

- A. Someone's online journal.
- B. A Celtic dance with wooden shoes.
- C. How the Celtics keep the Knicks away from the ball.


1

 **Test Your Tech**

Blogging is:


- A. Someone's online journal.
- B. A Celtic dance with wooden shoes.
- C. How the Celtics keep the Knicks away from the ball.

2

 **QuickWrite 9**


- I sent an email announcement on Friday afternoon
- QuickWrite 9 is open book
- Turn in to Catalyst Collect It by 10pm Tuesday
- Write two paragraphs on this topic:
 - Define opt-in and opt-out and describe the differences. Which is used in the U.S. and which in Europe?

3

 **Announcements**

- Thanksgiving Holiday
 - Thursday and Friday
 - UW classes are canceled
 - Wednesday
 - TA holiday
 - No lecture or labs
 - CLUE tutoring Tuesday night


4

 **Announcements**

- Free copy of Access, Vista, etc., for educational/academic use:
 - Links on Computing page on Course Web site
 - Search for CSE or INFO to find the link
 - Username is your full UW email address
 - Password is different!
 - Click on "send a reminder"
 - Check wherever your email forwards to
 - If you are in **INFO 100**, send me an email.

5

FIT 100– Fluency with Information Technology

 **A Table with a View**

Data Storage and Transfer with XML and Databases

D.A. Clements

6

Differences Between Tables and Databases

- When we think of databases, we often think of tables of information
- Comparing Tables
 - Database tables
 - Metadata tag identifying each of the data fields
 - Spreadsheet tables
 - Rely on position to keep the integrity of their data
 - HTML tables
 - Data as table entries with no unique identity at all
 - Concerned only with how to display the data, not with its meaning

16-7 D.A. Clements, UW Information School

The Database Advantage

- Metadata is key advantage of databases over other systems recording data as tables
- Two of the most important roles in defining metadata
 - Identify the type of data with a unique tag
 - Define the relationships of the data

16-8 D.A. Clements, UW Information School

XML: A Language for Metadata Tags

- Extensible Markup Language
 - Tagging scheme similar to XHTML
 - No standard tags to learn
 - Self-describing, think up the tags you need
 - Works well with browsers and Web-based applications
 - Use a simple text editor
 - XML tag names cannot contain spaces

16-9 D.A. Clements, UW Information School

Extensible Markup Language

XML

16-10 D.A. Clements, UW Information School

An Example from Tahiti

- Area in km² for Tahiti & neighboring islands

```
<?xml version = "1.0" encoding="ISO-8859-1" ?>
<archipelago>
<island><iName>Tahiti</iName> <area>1048</area></island>
<island><iName>Moorea</iName> <area>130</area></island>
<island><iName>Maiao</iName> <area>9.5</area></island>
<island><iName>Mehetia</iName> <area>2.3</area></island>
<island><iName>Tetiaroa</iName> <area>12.8</area></island>
</archipelago>
```

Figure 16.1 XML file encoding data for the Windward Islands database. The first line states that the file contains XML tags.

16-11 D.A. Clements, UW Information School

An Example from Tahiti (cont'd)

- First line
 - `<?xml version="1.0" encoding="ISO-8859-1" ?>`
- File should be ASCII text
- File extension should be `.xml`

16-12 D.A. Clements, UW Information School

Table 16.1 Rules for writing XML.

| | |
|---------------------|--|
| Required first line | <code><?xml version="1.0" encoding="ISO-8859-1" ?></code> must appear on the first line, starting in the first position. |
| First tag | The first tag encountered is the root element, and it must enclose all of the file's content; it appears on the second or possibly third line. |
| Closing tags | All tags must be closed. |
| Element naming | Observe these rules: <ul style="list-style-type: none"> Names can contain letters, numbers, and underscore characters. Names must not start with a number or punctuation character. Names must not start with the letters xml (or XML, or Xml, etc.). Names cannot contain spaces. |
| Case sensitivity | Tags and attributes are case sensitive. |
| Proper nesting | All tags must be well-nested. |
| Attribute quoting | All attribute values must be quoted; paired single quotes (apostrophes) or paired double quotes are okay; use "dumb" quotes only; choose 'opposite' quotes to enclose quoted values. |
| White space | White space is preserved and converted to a single space. |
| Comments | XML comments are written as <code>!-- This is a comment. --></code> |

16-1

D.A. Clements, UW Information School

Expanding Use of XML

- Combine encodings of two archipelagos – the Windward and the Galapagos Islands
- Root element is the tag that encloses all of the content of the XML file
 - `<archipelago>` in Fig. 16.1
 - `<geo_feature>` in Fig. 16.2
- Indenting for readability and structure

16-14

D.A. Clements, UW Information School

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
<geo_feature>
  <archipelago>
    <a_name>Windward Islands</a_name>
    <island>
      <iName>Tahiti</iName>
      <area>1040</area>
    </island>
    <island>
      <iName>Moorea</iName>
      <area>130</area>
    </island>
    <island>
      <iName>Maui</iName>
      <area>9.5</area>
    </island>
    <island>
      <iName>Hawaii</iName>
      <area>2.3</area>
    </island>
    <island>
      <iName>Tetiaroa</iName>
      <area>12.8</area>
    </island>
  </archipelago>
  <archipelago>
    <a_name>Galapagos Islands</a_name>
    <island>
      <iName>Isabella</iName>
      <area>4588</area>
      <elevation>1787</elevation>
    </island>
    <island>
      <iName>Fernandina</iName>
      <area>642</area>
      <elevation>1494</elevation>
    </island>
    <island>
      <iName>Tower</iName>
      <area>14</area>
      <elevation>76</elevation>
    </island>
    <island>
      <iName>Santa Cruz</iName>
      <area>986</area>
      <elevation>944</elevation>
    </island>
  </archipelago>
</geo_feature>
  
```

16-15

Figure 16.2 XML ignores white space, so the text in the file has been indented for easier reading.

Attributes in XML

- Use attributes for additional metadata, not for additional content
 - Not good, name is content:

```
<archipelago name="Galapagos">
```
 - Better to give alternate form of the data

```
<a_name accents="Gal&aacute;pagos">Galapagos</a_name>
```

16-16

D.A. Clements, UW Information School

Effective Design with XML Tags

- Identification Rule: Label Data with Tags Consistently
 - You can choose whatever tag names you wish to name data, but once you've decided on a tag for a particular kind of data, you must always surround it with that tag.

16-17

D.A. Clements, UW Information School

Effective Design with XML Tags (cont'd)

- Affinity Rule: Group Related Data
 - Enclose in a pair of tags all tagged data referring to the same entity. Grouping it keeps it all together, but the idea is much more fundamental: Grouping makes an association of the tagged data items as being related to each other, properties of the same thing.
 - Groups together data for a single thing – an island
 - Association is among properties of an object

16-18

D.A. Clements, UW Information School

Effective Design with XML Tags (cont'd)

- Collection Rule: Group Related Instances
 - When you have several instances of the same kind of data, enclose them in tags; again, it keeps them together and implies that they are related by being instances of the same type.
 - Groups together data of several instance of the same thing – islands
 - Association is among the objects themselves (entities)

16-19

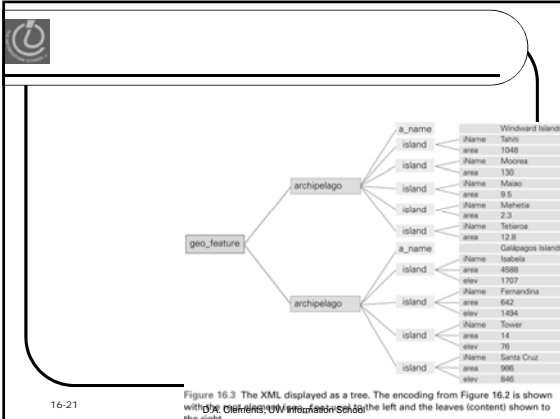
D.A. Clements, UW Information School

The XML Tree

- XML encodings of information produce hierarchical descriptions that can be thought of as trees
 - Hierarchy a consequence of how tags enclose one another and the data

16-20

D.A. Clements, UW Information School



16-21

Figure 16.3 The XML displayed as a tree. The encoding from Figure 16.2 is shown with the elements, attributes, and content on the left and the leaves (content) shown to the right.

DATABASES

D.A. Clements, UW Information School

22

What is a Database

- Any organized collection of data
- A collection of similar data
- Examples of databases:
 - Telephone book white pages
 - T.V. Guide
 - Airline reservation system
 - Motor vehicle registration records
 - Papers in your filing cabinet
 - Files on your computer hard drive

D.A. Clements, UW Information School

23

Data | Information | Knowledge

- Data
 - Can be defined in many ways
 - IS defines data as unprocessed information
- Information
 - Data that have been organized and communicated in a coherent and meaningful manner
- Knowledge
 - Knowledge—information evaluated and organized so that it can be used purposefully

Data is converted into information, and information is converted into knowledge

D.A. Clements, UW Information School

24

Data vs. Information

- We collect data
- Information is harvested from data
- Many companies are good at collecting data
- Fewer are good at harvesting information

D.A. Clements, UW Information School 25

Ultimate Purpose of a Database Management System (DBMS)

To transform

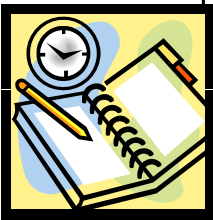
```

    graph LR
      Data --> Information
      Information --> Knowledge
      Knowledge --> Action
  
```

D.A. Clements, UW Information School 26

Why do we need a database?

- Keep records of our:
 - Clients
 - Staff
 - Volunteers
- To keep a record of activities and interventions
- Keep sales records
- Develop reports
- Perform research



D.A. Clements, UW Information School 27

Database Terminology

Phone book:

| | Fields (columns) | | | |
|----------------|------------------|--------------|-----|--|
| Records (rows) | Anderson Benson | Thomas Karen | A C | 123 Marine View Dr. 237-1234 1300 California Ave 237-1098 |
| | Cassery Drummond | Rick Lynn | W M | 12492 Rd 19 342-0502 12059 30th Ave W 931-1105 |

Table

| | | |
|--|--|--|
| Field (the columns in a table) | • Smallest unit of information in a table • Sometime called "attributes" | • First name • Last name • Middle initial • Street address • Phone number(s) |
| Record (the rows in a table) | • All related fields are collectively called a record • A collection of records is a data table | • All fields for one person are record • Collection of everyone's records |
| Database Management System (DBMS) | • All the related tables, queries, data and VBA modules constitute a database. | |

D.A. Clements, UW Information School 28


Database Management System (DBMS)

- Software tools for working with data
- Designed to:
 - Store (tables)
 - Organize (sort)
 - Add, modify or delete
 - Ask questions (queries)
 - Produce forms and reports
 - Summarizing
 - Displaying details
- Toolbox is a good analogy

D.A. Clements, UW Information School 29

DBMS Examples

- Microsoft Access
- Structured Query Language (SQL)
 - Microsoft SQL Server
 - Oracle
 - MySQL
- FileMaker Pro
- Lotus Notes
- Open Office Base



D.A. Clements, UW Information School 30

Flat-File vs. Relational Database

- Flat-File Database
 - All relevant data in a single table, or series of unrelated tables
 - Work best for small quantities of data; where viewing and sorting the data in a single list does not create a time-consuming task
 - Typically a person's first databases
 - Example: Excel spreadsheet or Word data list file
- Relational Database
 - Provide a solution to data entry redundancy problems
 - Linked through common fields (columns) with exactly the same data
 - Tables linked together can be queried as if one table

D.A. Clements, UW Information School 31

Flat-File Example

Staff Telephone List

| Last Name | First Name | Emp ID | Dept | Location | Work Phone | M/S | Supervisor Name | Supr Phone |
|-----------|------------|--------|-------|----------|----------------|-------|-----------------|----------------|
| Adams | Wes | 19929 | PROD | Seattle | (206) 221-1999 | GR-07 | Susan Buckle | (206) 221-2241 |
| Alberts | George | 21633 | PROD | Seattle | (206) 221-2153 | GR-36 | Marsha Mosley | (206) 221-1076 |
| Allen | Susan | 20256 | PROD | Renton | (206) 393-2025 | FB-18 | Frank Sullivan | (206) 393-1000 |
| Albert | Mara | 10544 | PROD | Seattle | (206) 221-1054 | GR-27 | Lynn Jarret | (206) 221-1336 |
| Andrews | Mike | 22113 | PROD | Seattle | (206) 221-2211 | GR-12 | Harry Hillis | (206) 221-2179 |
| Appery | Ward | 12244 | PROD | Renton | (206) 393-1224 | FB-14 | Molly Goldberg | (206) 393-1513 |
| Arthur | Diane | 12370 | MKT G | Bellevue | (206) 882-1237 | RL-27 | Wes Adams | (206) 221-1999 |
| Ashier | Jane | 11222 | ACCT | Seattle | (206) 221-1122 | EX-45 | Val Johnson | (206) 221-1999 |
| Astor | Lawrence | 20286 | PROD | Seattle | (206) 221-2028 | GR-10 | Peggy Farmer | (206) 221-2023 |
| Aynes | William | 22263 | PROD | Seattle | (206) 221-2226 | GR-10 | P. Kistner | (206) 221-2053 |
| Baker | Gerald | 19042 | ACCT | Seattle | (206) 221-1904 | EX-45 | Valerie Johnson | (206) 221-1999 |

- Weaknesses common to flat-file systems
 - Duplicate information in the table
 - Inconsistencies in the way Supervisor Names are entered

D.A. Clements, UW Information School 32

Relational Database Example

D.A. Clements, UW Information School 33

Database Tables

D.A. Clements, UW Information School 34

Query from Two Tables

| Emp ID | Lname | Fname | Hiredate | Dept | Grade | Salary |
|--------|---------|----------|-----------|------|-------|------------|
| 19929 | Adams | Wes | 0/23/1997 | PROD | 22 | \$3,812.55 |
| 21533 | Alberts | George | 2/18/2003 | PROD | 4 | \$1,611.05 |
| 20256 | Allen | Susan | 8/21/1999 | PROD | 14 | \$2,761.80 |
| 10544 | Albert | Mara | 1/17/1973 | PROD | 2 | \$1,280.90 |
| 22184 | Aly | Kim | 1/30/2004 | PROD | 3 | \$1,495.98 |
| 22113 | Andrews | Mike | 9/20/2004 | PROD | 13 | \$2,646.72 |
| 12244 | Appery | Ward | 9/13/1977 | PROD | 15 | \$2,876.88 |
| 12370 | Arthur | Diane | 1/17/1978 | MKTG | 20 | \$3,452.25 |
| 11222 | Ashier | Jane | 1/29/1974 | ACCT | 18 | \$3,262.16 |
| 20286 | Astor | Lawrence | 9/20/1999 | PROD | 5 | \$1,726.12 |
| 22263 | Aynes | William | 2/17/2005 | PROD | 1 | \$1,265.83 |
| 19042 | Baker | Gerald | 4/24/1996 | ACCT | 6 | \$1,841.20 |
| 17996 | Barker | William | 5/13/1993 | PROD | 9 | \$2,186.42 |

D.A. Clements, UW Information School 35

Forms

D.A. Clements, UW Information School 36

Reports

D.A. Clements, UW Information School 37

RELATIONAL DATABASES

D.A. Clements, UW Information School 38

Video

- Relational databases and tables

D.A. Clements, UW Information School 39

Relational Database Model

- Introduced by E. F. Codd in 1970
- A Logical View of Data
 - Enables developer to view data logically rather than physically
 - Greater logical simplicity tends to yield simpler and more effective database design methodologies

D.A. Clements, UW Information School 40

Tables

- Cornerstone of Relational DBMS
- Advantages – structural and data independence
- Conceptually Resembles a file
 - Note a file is actually a physical structure
- Easier to understand than its hierarchical and network database predecessors

D.A. Clements, UW Information School 41

Table Characteristics

- 1 A table is perceived as a two-dimensional structure composed of rows and columns.
- 2 Each table row (**tuple**) represents a single entity occurrence within the entity set.
- 3 Each table column represents an attribute, and each column has a distinct name.
- 4 Each row/column intersection represents a single data value.
- 5 All values in a column must conform to the same data format. For example, if the attribute is assigned an integer data format, all values in the column representing that attribute must be integers.
- 6 Each column has a specific range of values known as the **attribute domain**.
- 7 The order of the rows and columns is immaterial to the DBMS.
- 8 Each table must have an attribute or a combination of attributes that uniquely identifies each row.

D.A. Clements, UW Information School 42