XXII. Website Design

The Web

- Hypertext Data Independence
- Data Models for Hypertext Documents
- The Araneus Data Model (ADM)
- The Navigational Conceptual Model (NCM)
- The Araneus Methodology for Website Design

The Web

- The spread of World-Wide Web (hereafter “Web”) technology is one of the most remarkable phenomena of the last few years in all areas of computing and communication.
- The Web (e.g., Web browsers) is becoming a standard interface for the general public to access and exchange information:
  - The protocol is very simple and public;
  - The interface is uniform;
  - The content is extremely rich (both in breadth and in depth);
- Moreover, the Web is becoming a standard interface for accessing many specialized services, specifically information systems and databases of every type.
Web Features and Open Problems

- The Web is a simple and powerful data integration tool.
- Two basic approaches to Web-based data integration:
  - Coarse-grain: pages of hypertext;
  - Fine-grain: unified interface for accessing different (usually similar) information systems available on the Web.
- The Web is built out of semi-structured (HTML/XML) documents, databases contain structured (i.e., tuple/record) data.
- Databases can be queried in a flexible way; hypertext documents are easy to access, but cannot be “queried”.
- Web sites are often difficult to explore, use and monitor.
- Web sites are also difficult to design and maintain.

Problems with Large Websites

- Information is often poorly organized and difficult to access.
- It is often unclear what information is available on a given website.
- The access structure of many websites is casual and idiosyncratic, causing frequent dangling references.
- The style of presentation is often heterogeneous.
- Large websites are usually difficult to update, or change their structure.
- It is also difficult to change the presentation structure and/or details.
Data Independence for Hypertext Documents

- You might say that there are three facets to the Web:
  - **Data** -- what information is offered through the site and what are the conceptual details and the logical organization;
  - **Hypertext** -- how data is arranged in pages and what navigation links correlate them;
  - **Presentation** -- the appearance of each piece of information on each pages.

- As much as possible, we’d like to decouple the three, so that changes to one affect minimally the other two facets of the Web.

**An Example**

[Image of a web page with links and a search bar]

**Doctoral Programs**
- Doctoral Program in Machine Engineering
- Doctoral Program in Environmental Hydraulic Engineering
- Doctoral Program in Methods and Techniques for Environmental Monitoring
- Doctoral Program in Mineral Science and Technology
...Another...

Doctoral Program in Machine Engineering

Involved Universities:
- Ancona
- Basilicata
- Bologna
- Calabria
- Perugia

Contact: Giuseppe Grottoni

...and Another...

Doctoral Programs

Doctoral Program in Machine Engineering
Ancona, Basilicata, Bologna, Calabria, Perugia
Giuseppe Grottoni

Doctoral Program in Environmental Hydraulic Engineering
Calabria, Bari, Salerno, Basilicata
Luca Vesi

Doctoral Program in Methods and Techniques for Environmental Monitoring
Firenze, Pisa, Genova, Basilicata
Katia Colanelli
Design Issues

- **Data** -- choose the content
- **Hypertext** -- choose navigation paths
- **Presentation** -- define layout and graphics

### Maintenance Issues

- **Data** -- changing the content
- **Hypertext** -- changing navigation paths
- **Presentation** -- changing layout and graphics

Components and Models

<table>
<thead>
<tr>
<th></th>
<th>ER and Relational</th>
<th>What is missing is a model for hypertexts!</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hypertext</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Presentation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HTML</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Models for Hypertext Documents

- In **data-intensive Websites** (and often in general) there are (many) pages with a similar (or even the same) structure.
- Forty years ago people realized that in an application it is often the case that there are many records **with the same structure**; files and file technology were invented to exploit this fact.
- Likewise, the notion of a **schema** for a database was later introduced as an overall description of the content of a database.
- We need something similar for the Web!
A Web Page

Professors and Researchers

Gian Lorenzo VALENTI

Name: Gian Lorenzo VALENTI
Position: Full Professor
Address: G.I.F.A. - Universita degli Studi di Cagliari
Via delle Terrece, 3
09100, Cagliari - Italy
Telephone: +39 - 071 - 474610, 646
Fax: +39 - 071 - 64637
E-mail: valenti@unica.it
Research Areas of Interest: Chemistry and Applied Chemistry

A Page Schema: ProfessorPage

ProfessorPage

Text elements

List element
ADM (Araneus Data Model):
A Logical Model for Hypertext Documents

- Developed at the University of Rome III (Università di Roma Tre) by Paolo Atzeni, Paolo Merialdo, Giansalvatore Mecca and colleagues.
- Its features include:
  - Page schemas
  - Simple attributes
    - text, images, ...
    - link (anchor, URL)
  - Complex attributes such as lists, possibly nested.
  - A heterogeneous union operation.
  - Forms as virtual lists over form fields and links to a result.

Another Web Page -- Containing a List of Links

Professors and Researchers

- Franco AMATucci
- Salvatore AMURÁNO
- Gianfranco BONO
- Ada BONISLAVI
- Monica CARBONELLI
- Paolo CLAPES
- Vito De Rista COPETTRO
- Vincenzo CUOMO
- Bernardino DE RICCIARDI
- Paolo DI GIROLAMO
A Page Schema for ProfessorListPage

An ADM Schema
Heterogeneous Union and Forms

SearchProfPage

ProfessorListPage

ProfessorPage

Form element
Data Models, Again

There is considerable conceptual distance between the two!

A Simple ER Schema
The Navigational Conceptual Model (NCM)

NCM fills the gap between the two

Database Conceptual Schema (entities - relationships)
Hypertext Conceptual Schema (macroentities, directed relationships, aggregations)
Hypertext Logical Schema (page schemas, links)
Macroentities and Directed Relationships

![Diagram showing relationships between entities: Professor, Student, Teacher, Tutorship, Course, Lesson, Department, People, Activities, Student, Professor, Course, Seminar, Tutorship, Teacher. Diagram includes 1:1, 1:N, 1:1 relationships.]

Macroentity (was two entities)

Aggregation Nodes

![Diagram showing aggregation nodes: Department, People, Activities, Student, Professor, Course, Seminar, Tutorship, Teacher. Diagram includes 1:1, 1:N, 1:1 relationships.]

© 2001 Paolo Atzeni
An NCM Schema

The Araneus Methodology

Database conceptual design

Hypertext conceptual design

Database logical design

Hypertext Logical design

Presentation design

Page Generation

© 2001 Paolo Atzeni
The Araneus Methodology: Design from Scratch

Database conceptual design → Hypertext conceptual design

Database logical design → Hypertext Logical design

Presentation design

Page Generation

The Araneus Methodology: Design from an Existing Database

Database conceptual design → Hypertext conceptual design

Database logical design → Hypertext Logical design

Presentation design

Page Generation
Hypertext Conceptual Design: From an ER schema to a NCM Schema

Database conceptual design → ER Schema → Hypertext conceptual design → NCM Schema

Database logical design → Hypertext Logical design → Presentation design → Page Generation

Hypertext Conceptual Design: Step 1

Choose and describe macroentities; design views over the input ER schema.

ER

- Course (1:N)
  - Name
  - Description
  - Day
  - Hour

- Lesson (1:1)

NCM

- Course (1:N)
  - Name
  - Description
  - Day
  - Hour

- Lesson
**Hypertext Conceptual Design: Step 2**

Choose navigation paths

**ER**

- Professor \(1:1\) Paper \(1:N\)
- Research-Group \(1:N\) Paper \(1:N\)

**NCM**

- Professor \(1:N\) Paper \(1:N\)
- Research-Group \(1:N\) Paper \(1:N\)

---

**Hypertext Conceptual Design: Step 3**

Shape the hypertext access structure

**NCM**

- Professor \(1:1\) Seminar \(1:N\)
- Research-Group \(1:N\) Seminar \(1:N\)

- Professor \(1:1\) Research Activities
- Seminar \(1:N\) Research-Group \(1:N\)
The Input ER Schema

The Output NCM Schema
Hypertext Logical Design: From an NCM to an ADM Schema

Hypertext Logical Design: Step 1

Map each macroentity into either a page schema or a list element inside a page schema
Another Example for Step 1

Hypertext Logical Design: Step 2

Map each directed relationship into a (list of) link attribute(s)
**Hypertext Logical Design: Step 3**

Map each aggregation into a unique page schema with link attributes to the target page schemas

**Resulting ADM Schema …Sideways…**