Web-based Information Systems and their Engineering

Geert-Jan Houben
Contents

• Web Information System (WIS)

• Evolution in Web and Web data

• WIS and Web Engineering
Web Information System (WIS)
World Wide Web

- Tim Berners-Lee
  - “One of the things computers have not done for an organization is to be able to store random associations between disparate things, although this is something the brain has always done relatively well.”
  - First played with this idea in 1980 and later in 1989 at CERN he proposed that a global hypertext space be created in which any network-accessible information could be referred to by a single "Universal Document Identifier".
  - He wrote in 1990 a program called "WorldWideWeb", a point and click hypertext editor which ran on the "NeXT" machine. Together with the first Web server, this was distributed to CERN and the hypertext and NeXT communities in 1991. The specifications of UDIs (now URIs), HyperText Markup Language (HTML) and HyperText Transfer Protocol (HTTP) were published on the first server to promote wide adoption and discussion. With help from Robert Cailliau. Presentation at Hypertext'91. First server outside Europe at SLAC.
TBL’s Dream

• “The dream behind the Web is of a common information space in which we communicate by sharing information. Its universality is essential: the fact that a hypertext link can point to anything, be it personal, local or global, be it draft or highly polished. There was a second part of the dream, too, dependent on the Web being so generally used that it became a realistic mirror (or in fact the primary embodiment) of the ways in which we work and play and socialize. That was that once the state of our interactions was on line, we could then use computers to help us analyse it, make sense of what we are doing, where we individually fit in, and how we can better work together.”
W3C

• Between the summers of 1991 and 1994, the load on the first Web server ("info.cern.ch") rose steadily by a factor of 10 every year. In 1992 academia, and in 1993 industry, was taking notice.

• TBL felt pressure to define the future evolution. He formed the World Wide Web Consortium (W3C) in September 1994. The Consortium is [defined to be] a neutral open forum where companies and organizations to whom the future of the Web is important come to discuss and to agree on new common computer protocols. [decision by consensus]
Hypertext

- History started (long) before 1991.
- Vannevar Bush’s article in Atlantic Monthly about a photo-electrical-mechanical device, Memex (memory extension), which could make and follow links between documents on microfiche.
- Later:
  - Doug Engelbart makes NLS for hypertext browsing and editing (1962-1968).
  - Ted Nelson introduces the word Hypertext (1965).
  - Much more work on Hypertext, also academic.
Web: Hypertext Implementation

• Hypertext research led to many rich prototype systems. It is about very complex and dynamic systems of linking and cross-referencing (relevant for our focus).
• It also led to hypertext (fiction) as part of literature.
• But it didn’t catch on, until there was the Web: its most popular implementation, even when it is a “poor” one.
• Be aware that hypertext was about richer support for linking data (than the Web):
  – not just clicking, bi-directional and back links, typed links, text expansion, editing as well, etc.
WWW

• Internet (“hypertext on the Internet”)
  – only uni-directional links, with its pro’s and con’s
• Browsers and HTML
  – Mosaic
• Web servers and HTTP protocol
Information system

- Exchanges information with *object system* (= business process)
- Stores and manages information: *data-intensive*
- Requires careful *engineering* of information exchange
- Requires careful *engineering* and *modeling* of *object system*
- Traditionally *database*-oriented
Web Information System

- Information System based on *Web technology* (Web-based, Web-aware, Web-enabled etc.)
- Web technology can be used as *front-end*, e.g. application is available on the Web (or Intranet) via a browser
- Enables easy use and maintenance of (personalized) *end-user access*
- Web *metaphor* is appealing for end-users
- Requires different techniques for engineering the system’s *interfaces*
Web Information System

- Web technology can also be used in back-end of information system
- Organize (connect) the data inside the system using Web technology
- Use World Wide Web as provider of data (or Intranet)
- Typically highly volatile information (distributed and heterogeneous)
- Requires different techniques for engineering the implementation
Examples

• Real-estate sales
• Employee databases
• Museum databases
• Digital libraries
• Mail order catalogs
• Reservation systems
• Auctions, virtual marketplaces
• Electronic TV program guides

WebML View

- A Web-enabled software system whose main purpose is to publish and maintain large amounts of data
- Interfaces directed to general public
  - Exploratory
  - Browsing-oriented
  - Personalized (1 to 1)
- Data stored by means of DBMS technology
  - Possibly pre-existing the Web application
  - Normally volatile
  - With severe “freshness” requirements
  - May be distributed and heterogeneous
• Object-Oriented Hypermedia Design Method
• WWW brought new generation of IS:
  – *Navigation* through heterogeneous information space
  – *Operations* querying or affecting that information
• Introduced *hypertext* (hypermedia) paradigm
• Applications are constantly *modified*, enriched with new services, and new navigation and interface features are added
OOHDM View (2)

• Web-based application, *first* good hypermedia applications
• Traditional (SE) methodologies have no notion of linking: little is said about incorporating *hypertext* into interface
• Size and complexity imply *systematic* approach for evolution and reuse of design knowledge
RMM’s View

• History: *graphics designers + programmers*

• Experience: central information architecture and shared, common mechanisms/services helpful for coping with problems of *scalability* and “*information anarchy*”
Nielsen’s View

• “On the Web, the only constant is change. A site that works perfectly as long as its stays the same will quickly die.”

• Healthy navigation structure key to success

• Building interface is also complex, connecting interface objects to rest of application
Our Focus: Web-based IS

• A software system whose main purpose is to publish and maintain large amounts of data and does so using hypertext-based principles to offer end-user access to webs of data.
  – Web-browser as front-end
  – Data repository (database) as back-end
  – Heterogeneity, lack of typing and constant change

• Design (model-based):
  – Data (content) structures (ER-modeling)
  – Navigation
  – Presentation (layout)
  – Database/repository access
Recommendation

• If you have never before looked at the history of the Web and the work on hypertext, have a look at the topic.
• For anything on the Web, a good starting point is the site of W3C: w3c.org
• There have been Hypertext conferences (ACM) for the academic work on hypertext. The Wikipedia page on Hypertext gives a good starting point for different aspects of hypertext (research) and its role in the history of the Web.
Assignment

• Consider two publicly accessible WIS, assume you are one of their (main) designers, and determine what are the main aspects to consider in the design (and for which support appears useful).

• Name the two WIS and enumerate these aspects in a bulleted list on a single Powerpoint slide (standard layout, font less relevant), and email this with subject “WIS homework 1” to g.j.houben@tue.nl by September 8th, 1100 (am, strict). [6 hrs]