A Historical Perspective of Web Engineering

Guangzhi Zheng
Georgia State University, USA

INTRODUCTION

As Web-based information systems and applications have been growing more and more complex, the concept of "Web engineering" begins to prevail. Generally, Web engineering refers to a field concerned about the employment of systematic and disciplined approaches to the development and management of Web sites and Web applications (Lowe, 1999; Murugesan, Deshpande, Hansen, & Ginige, 1999). It utilizes diverse principles from multiple disciplines including hypermedia, software engineering, human computer interaction, conceptual modeling, information retrieval, networking, etc.

The concept of Web engineering was developed in the middle 1990s when Web development practitioners and researchers eventually realized that Web development is beyond just Web page design and a systematical process is needed to develop large scale Web sites. White from Stanford organized a tutorial about Web document engineering (White, 1996) at the 5th World Wide Web Conference in 1996 and introduced approaches from hypertext engineering field. Since then, the Web has had such a tremendous advance that it has grown from a mere publishing medium to a platform for a full range of applications and systems. This adds all kinds of difficulties to the already problematic Web development.

It has been more than 10 years since White's tutorial at the 5th WWW conference. What has changed in all these years? What is the status of Web engineering at present? What are the challenges and opportunities in future? This article will introduce Web engineering with a historical perspective on the following three issues:

1. Evolvement of the Web information system concept
2. Web application development characteristics and major techniques/methodologies
3. Development of Web engineering as a discipline and as a community

BACKGROUND: WEB INFORMATION SYSTEM

The core artifacts of Web engineering are information and computer applications based on the Web. A number of terms have been used to label this artifact, including homepage, Web page, Web document (White, 1996), Web site (Lowe, 1999; Powell, Jones, & Cutts, 1998), Web(-based) application (Kappel, Pröll, Reich, & Retschitzegger, 2004; Lowe, 1999), Web(-based) system (Lowe, 2001; Murugesan et al., 1999), and Web(-based) information system (Isakowitz, Bieber, & Vitali, 1998). These terms reflect a diverse understanding of Web information system, especially at its different evolvement stage. Many people use these terms interchangeably and do not specifically define the one they choose to use. But the meanings of them are not exactly the same, especially viewed in a historical perspective. (To be consistent, the term "Web information system (WIS)" will be used in this article.)

Earlier names reflect the Web as a publishing medium. Before the middle 1990s, Web sites were mainly static and were used mainly for information or document presentation. People felt confident that "Web is a publication, not a piece of software" (Stross, 1996, p. 180). The term "Web document" given by White (1996) also reflects this document centric feature:

Web document (or hypertext document) is a collection of Web pages (hypertext pages). Each Web page is typically a single computer file... A Web site can be thought of as a single Web document if it represents a single, logical information space.

During the late 1990s, with the wide spread of server side technologies such as CGI and ASP, and database connectivity technologies such as ODBC, more functionalities were developed and provided on the Web. The scope of WIS had been expanded to include more complex applications, such as transaction, analysis, information management, entertainment, collaboration, etc. It also grew into a platform to support complex enterprise systems. The Web was eventually perceived

to be software applications besides just information and documents. Thus, terms such as Web application and Web information system were used to mean something different from Web page or Web document. The definition given by Holck (2003) indeed captures this evolving role of Web:

(WIS is) a computer-supported information system, utilizing the technology of the WWW, and accessed by the majority of its users via a browser.

Christodoulou and Papatheodorou (2005) go further to distinguish Web page, Web application, and WIS, thus giving a more complete understanding of these related terms:

WIS is an information system utilizing Web technologies to provide information (data) and functionality (services) to end-users through a hypermedia-based presentation/interaction user interface on Web-enabled devices.

Web applications are the different functionality-oriented components of a WIS. A Web application is actually a small-scale WIS, providing very specific information or functionality. Many developers use these terms as synonymous, especially for small WISs.

This is a fairly good definition of WIS for its development in the late 1990s and earlier 2000s. However, even this definition cannot catch up with the fast growth of WWW. For example, it does not reflect the latest development of Web services, which typically do not provide a user interface. To accommodate this new development, the definition above can be updated slightly as:

A Web information system is an information system that utilizes Web technologies to deliver information and services to users or other information systems/applications. A Web application is a specific functionality-oriented component that utilizes Web technologies to deliver information and services to users or other applications/information systems. A Web information system usually consists of one or more Web applications, together with information components and other non-Web components.

However, even this definition is likely to be revised in the near future as the Web is still growing at a fast pace. Any definition at present will soon be questionable because of new systems or new applications that expand the use of WWW (for example, the recent prevailing concept and implementation of Web 2.0). The diversity of WISs also creates problems for a good definition, because it is hard to have a simple but accurate and complete definition to cover all aspects. This diversity is also a difficulty for establishing a good classification of WISs.

Classification of Web Information Systems

Different types of WISs vary in their purpose, scope, and complexity. A good classification of WIS will help to understand WIS more in depth. It is also necessary because different types of WIS potentially need different techniques and approaches to development and maintenance. As yet, there is no complete or widely accepted way to categorize WIS (Murugesan & Ginige, 2005).

Earlier classifications were pretty simple and they tended to be based on the scope and complexity of WIS. Powell et al. (1998) simply categorized Web sites into Internet (public access), intranet (internal to an organization), and extranet (limited and controlled access to the public). Such classification was also used by Isakowitz et al. (1998), which further divided public access sites into Web-presence sites that are marketing tools designed to reach consumers and electronic commerce systems that support consumer interaction. Powell et al. (1998) also categorized Web sites by complexity levels into static (pure static Web pages), dynamic (database driven), interactive (dynamic data access), and fully functional sites (supporting processes and transactions).

Later, functionalities by Web sites were considered in more meaningful classifications. Ginige and Murugesan (2001), Deshpande et al. (2002), and Murugesan et al. (2005) provide similar taxonomies based on WIS functionality: informational (news, classifieds), interactive (information search, customized information presentation), transaction (shopping, banking, billing), workflow oriented (planning and scheduling, management), collaborative (group/team site, Wiki), online community (forum, BBS), and Web portals (which provide a mixture of information and functionality above).

The previous classifications correspond to the more traditional perception of WIS as Web sites. They do not cover new developments of WIS. Thus, the clas-