Communications of the Association for Information Systems

Volume 12 Article 14

August 2003

Web Engineering: An Assessment of Empirical Research

Bouchaib Bahli
Concordia University, bbahli@jmsb.concordia.ca

Dany Di Tullio
BOC Group Canada, dany.ditullio@ca.gases.boc.com

Communications of the Association for Information Systems

WEB ENGINEERING: AN ASSESSMENT OF EMPIRICAL RESEARCH

BOUCHAIB BAHLI
Concordia University
bbahli@jmsb.concordia.ca

DANY DI TULLIO
BOC Group Canada

ABSTRACT

Web engineering is the process used to create high-quality Web-based systems and applications that deliver a complex array of content and functionality to a broad population of end-users. As Web Engineering continues to grow in popularity with practitioners and academics alike, so far, there hasn't been any assessment of its accumulated body of knowledge in terms of academic research. Because Web engineering was established as a new discipline some five years ago, it is perhaps time to take stock of the efforts made in this field. Using the Web Engineering Process Model developed by Pressman [2000], this paper organizes and map progress made so far. The results suggest a significant need for theory-based research in Web Engineering. The paper discusses some of the managerial and research implications of the findings.

Keywords: web engineering, research agenda

I. INTRODUCTION

The Web is now used quite differently than it is the purpose for which it was originally conceived – sharing scientific information among a handful of academics. From simple static web pages, which are commonly associated to the humble beginnings of the Web, current Web applications vary both in scope and complexity [Cachero and Pastor, 2001]. Although users can still come across simple Web sites, large-scale enterprise applications distributed across the Internet, corporate intranets and extranets are commonplace [Ginige and Murugesan, 2001]. As Web-based systems and applications evolved rapidly, the complexity of designing, developing, managing, and maintaining these systems also expanded. Reflecting the significant growth of Web development, steady streams of research surfaced about various stages of the Web development process, from planning, analysis, to engineering, testing, and evaluation of Web-based systems**.** Furthermore, academics recently argued for the establishment of Web Engineering as a discipline in its own right to address the growing need for disciplined approaches and new methods and tools for the development, deployment, and evaluation of Web-based systems [Murugesan et al. 1999].

As Web Engineering continues to grow in popularity with practitioners and academics alike, its accumulated body of knowledge is not yet assessed in terms of academic research. It is perhaps time take stock of the efforts made in this fledgling field. Using the Web Engineering Process Model developed by Pressman [2000] as a framework, this paper sets out to organize and map the research progress made so far. We believed it is indeed important, to examine Web Engineering to determine what is already accomplished, what is not yet done, and how to fill the gap. This need for assessment is the main motivation of this paper.

The remainder of the paper is organized as follows. Section II presents the theoretical background on Web Engineering. Section III describes the Web Engineering Process Model [Pressman 2000] used to examine Web Engineering research in this area and the third section presents the findings of our assessment. Section IV discusses the findings and suggests some recommendations for future research.

II. THEORETICAL BACKGROUND

WEB ENGINEERING: THE EMERGENCE OF A NEW DISCIPLINE

In a relatively short period of time, the Internet, and more particularly the World Wide Web, became an integral part of every day life. Entire industries such as banking, travel, and government are Web-enabled to address customer needs and to enhance their operations and performance [Powel, 1998]. Countless organizations scramble to offer Web-enabled systems and services. The complexity of Web-based applications continues to grow significantly, moving , from simple static informational Web sites to complex, dynamic, interactive, and transactional Web-based systems [Spiliopoulos, 2000]. This frantic rush to be on the Web, coupled with the ever increasing complexity of Web applications created serious problems, which need to be addressed. Among the problems currently surrounding Web-based system development is the common use of ad hoc, hacker-type approaches, which lack rigor, systematic techniques, quality assurance, and sound methodologies [Murugesan et al. 1999]. Furthermore, a survey conducted in 2000 by the Cutter Consortium found that in the case of Web systems, projects exceeded budget 63% of the time, delivered systems did not offer the required functionality 53% of the time. Deliverables were of poor quality 52% of the time [Cutter 2000]. Faced with this situation in 1998, a group at the University of Western Sydney (Deshpande, Ginige, Hansen, and Murugasen), decided to propose Web Engineering as a new discipline [Ginige and Murugesan 2001]. They organized the first workshop on Web Engineering in 1998 in conjunction with the World Wide Web Conference (WWW7) in Brisbane, Australia,. Their aim was to address the issues, promises, and challenges facing the development of web-based applications and systems. Subsequent workshops confirm the increased interest by academics and practitioners alike in the emerging field of Web Engineering [Murugesan et al. 1999].

Web Engineering advocates a process and a systematic approach to the development of high-quality Web-based systems. While various but quite similar definitions of Web Engineering are proposed. we adopted the one provided by the founders of the field. It is quite thorough and precise:

> *Web Engineering is the establishment and use of sound scientific, engineering, and management principles and disciplined and systematic approaches to the successful development, deployment and maintenance of high quality Web-based systems and applications* [Murugesan et al. 1999].

Web Engineering is compelled to become a multidisciplinary field, encompassing diverse disciplines as human-computer interaction, user interface, systems analysis and design, software engineering, requirements engineering, hypermedia engineering, information structures, testing, modeling and simulation, project management, and arts and graphic design [Murugesan et al. 1999]. In short, Web development can be defined as a composite field;

> *“a mixture between print publishing and software development, between marketing and computing, between internal communications and external relations, and between art and technology” [Powell 1998].*

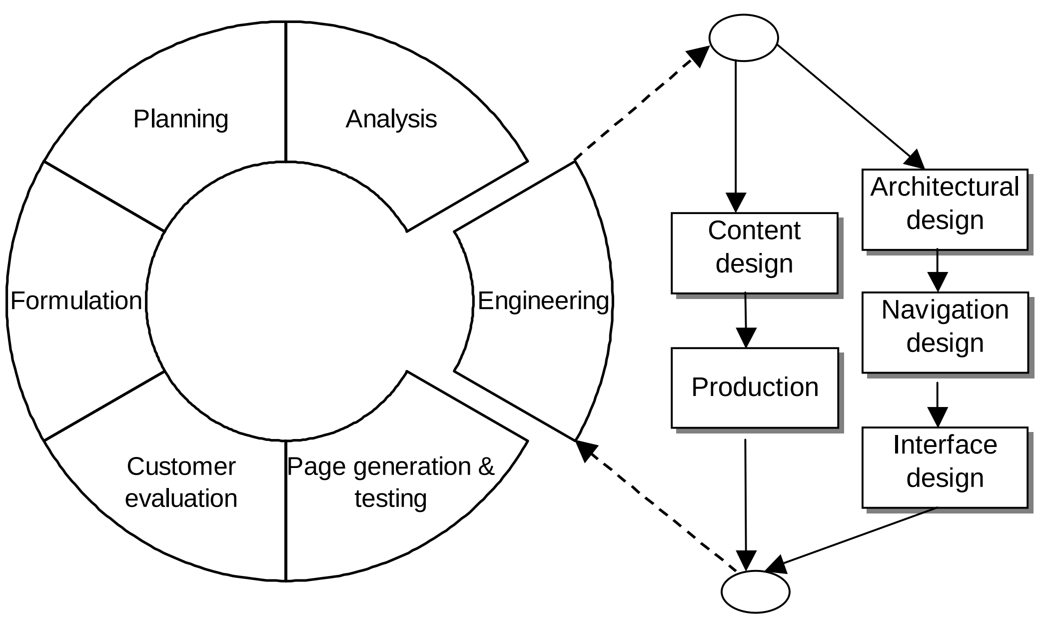
In addition, we note that Web Engineering affects and is affected by large scale deployment of hand-held devices with web and color access. Internet-enabled cell phones design emphasizes portability to accommodate web interaction [Schilit, Trevor, Helbert, Koh, 2002]. This portability brings new opportunities and challenges to the integration of a variety of small Web-capable wireless devices including availability, security properties, response time [Menascre, 2002], Web usability [Chi, 2002], and a stability of Wireless Web [Singh, 2001].

Faced with this amalgam of different facets of Web development and to attenuate the current ad hoc and sometimes chaotic approaches to the development of Web-based systems], Web Engineering strives to bring the entire Web development process under control, while minimizing risks, and enhancing quality and maintainability [Fewster and Mendes, 2001]. Web Engineering, therefore, encompasses all phases of the Web-based system development process including the analysis, design, implementation and, performance evaluation and continual maintenance stages [Deshpande and Hansen, 2001]. The following subsection describes the Web Engineering Process Model proposed by Pressman [2000]. This model is used in this paper as a comprehensive framework for mapping existing research in Web Engineering.

THE WEB ENGINEERING PROCESS MODEL

As Web-based systems evolve from static, content-driven applications to dynamic interactive and ever-changing ones, the need to apply sound methodologies throughout the development process and maintenance stages is called for by researchers [Murugesan et al., 1999; Pressman, 2000; Klapsing, Neumann and Wolfram, 2001].

To achieve this goal and to put discipline into the process of developing Web-based systems, Pressman [2000] proposed a Web Engineering framework that encompasses an effective process model, populated by framework activities and engineering tasks (Figure 1). To the best of our knowledge, this model is the only model that captures the main components of Web-based development process discussed in the literature.



Source: [Pressman, 2000]

Figure 1. The Web Engineering Process Model

As shown in Figure 1, the Web Engineering process begins with the:

Formulation phase. At this stage, the goals and objectives of the Web-based system or application are established in conjunction with the scope of the project for the first increment.

Planning is then conducted. The overall project costs are estimated and the risks associated with the development effort are estimated. A fairly detailed development schedule for the initial Web-based system increment is also created.

During *analysis*, the technical requirements for the system are specified and the content aspect of the system is addressed, including graphic design and content items to be incorporated into the web site.

The *engineering* activity involves both content design and production tasks, including acquiring all content that is to be included into the Web-based system (e.g., text, graphics, audio, video). *Architectural*, *navigation*, and *interface design* tasks are performed in parallel**.** These tasks involve the overall hypermedia structure of the Web-based system and the application of design patterns and templates to populate the structure.

Page generation follows engineering*.* It consists of merging the content defined in the engineering activity with the architectural, navigation, and interface designs to produce executable Web pages in HTML, XML, and other process-oriented languages. Integration with component middleware such as CORBA, DCOM, and JavaBeans is also accomplished during *page generation*.

The *testing* phase involves the intensive use of the Web-based system to reveal potential errors caused by applets, scripts, forms, or the use of various Web browsers.

The *customer evaluation* activity reviews each increment of the Web Engineering process. During this phase, customers can request changes to the system, which will then be integrated into the next incremental process.

We now elaborate on each of these activities.

FORMULATION OF WEB-BASED SYSTEMS AND APPLICATIONS

It is during formulation that the customer and the developers establish a common set of objectives for the development of the Web-based system. The development team determines broadly the scope of the development process and how to achieve the desired outcomes. Three questions should be considered at this stage [Powell 1998]:

1. What is the main motivation for the Web application?;
2. Why is the Web application needed?;
3. Who will use the web application?

In answering these questions, customers and developers will ultimately address two types of goals:

- informational goals which indicate an intention to provide specific content and/or information to the end user and
- applicative goals which indicate the ability to perform various tasks with the Web system.

Once these goals are identified, a user profile of the system can be developed followed by a statement of scope based on these same goals [Pressman, 2000].

Planning and Analysis of Web-based Systems and Applications

The planning phase includes estimates of the overall project costs and evaluation of risks associated with the development effort. A fairly detailed development schedule for the initial Web-based system increment and a broader one for subsequent increments are also determined

[Pressman 2000]. In the analysis phase of the Web Engineering process model, four types of analysis are conducted [Nanard and Kahn, 1998]:

1. *Content analysis* where graphic design and content items to be incorporated into the web site are defined;
2. *Interaction analysis*, that is, the way in which various users will interact with the Web application being considered. Use cases can be written to provide detailed descriptions of different kinds of interaction;
3. *Functional Analysis* in which the use cases developed at the interaction analysis phase define the operations that will be applied to the Web system's content and entail other processing functions.
4. *Configuration analysis* in which the infrastructure and the environment in which the Web-based system will run are described in detail.

Engineering Web-based Systems and Applications

As shown in Figure 1, five activities are identified including architectural design, navigation design, interface design, content design and production [Bernstein, 1998; Lynch and Horton, 1999].

- The architectural design involves design of the hypermedia structure of the system and the application of design patterns and constructive templates to populate the structure. Four different structures can be considered when designing a Web-based system [Powell 1998]:
    1. Linear structures which are employed when a predictable sequence of interactions is common;
    2. Grid structures used when the Web-based system's content can be organized in two or more dimensions;
    3. Hierarchical structures, a common Web application structure, allow navigation horizontally along vertical branches of the hierarchy;
    4. Networked structures where web pages are designed so that they may pass control (via links) to almost every other component in the system.

  It should also be noted that a combination of these structures could be used in designing and developing a Web-based system.

- Following the specifications of the system's architecture and the identification of its components (pages, scripts, applets, and other processing functions), designers create and design navigation pathways that will eventually enable a user to access Web application content and services [Pressman 2000]. As the design process moves forward, each navigation pathway is identified and links that are appropriate for the proposed content are determined [Gnaho and Larcher, 1999].
- The interface of a Web-based system is its "first impression", it must be carefully designed and laid out. Even an extremely efficient web site may not attract many users if the interface design is poor. A well-designed interface improves the user's overall perception of the web site and can only benefit the sites owner. [Nielsen and Wagner, 1996; Rosenfeld, 1998].
- Content design that involves the design, and/or acquiring of all content that is to be included into the Web-based system (e.g., text, graphics, audio, video). This activity also derives the overall structure and detailed layout of the information content that will be presented as part of the Web-based system [Lynch and Horton, 1999].
- The production phase entails acquiring the content to be included into the Web-based system [Laurent and Cerami, 1999].

Page Generation and Testing of Web-based Systems and Applications

Page generation consists of merging the content defined in the engineering activity with the architectural, navigation, and interface designs to produce executable Web pages in languages such as HTML and XML. Integration with component middleware such as CORBA, DCOM, and JavaBeans is also accomplished during *page* generation [Pardi, 1999].

Testing represents a considerable challenges in the development of Web-based systems and applications because these systems run on networks and interoperate with various operating systems, browsers, hardware platforms, and communications protocols. The following steps summarize the testing activity [Olsina 1999; Pressman 2000]:

- The content model for the Web system or application is reviewed to uncover errors;
- The design model for the Web system or application is reviewed to uncover navigational errors;
- Selected processing components and Web pages are unit tested;
- The architecture is constructed and integration tests are conducted;
- The assembled system is tested for overall functionality and content delivery; 6) The Web system or application is implemented in a variety of different environmental configurations and tested for compatibility with each configuration; 7) The Web system or application is tested by a controlled and monitored population of end-users.

Customer Evaluation of Web-based Systems and Applications

Each increment of the Web Engineering process is reviewed during the customer evaluation activity. During this phase, customers can request changes to the system, which will then be integrated into the next incremental process [Pressman 2000].

III. RESEARCH METHODOLOGY

To review Web Engineering research systematically, numerous studies discussing the topic were examined. Specifically, papers published in leading MIS journals and proceedings were considered. Among the journals and proceedings reviewed were Communications of the ACM, Information Systems Research, MIS Quarterly, Journal of Management Information Systems, IEEE Multimedia, Journal of AIS, Communications of AIS, International Conference on Information Systems Proceedings and, Hawaii International Conference on System Sciences Proceedings, a variety of conferences and workshops on Web Engineering. The journals were chosen notably because they are highly ranked by Mylonopoulos and Theoharakis [2001].

Journal and proceedings articles were selected for consideration only if they involved research pertaining to Web Engineering. The assessment was based on titles, abstracts, and key words emphasizing Web Engineering, the Web, Web-based information systems, Web-based methodologies, and Web sites development.

In our classification of the research papers relative to the activities described in Pressman's [2000] Web Engineering Process Model, both authors read the all the papers and independently categorized them. The inter-rate agreement was found to be 96%. This analysis and categorization method provides us with a comprehensive look at Web Engineering research to date and indicates the validity of our findings. Table 1 summarizes the findings of this study. The classification of Web Engineering literature based on the six phases of the Web Engineering Process Model [Pressman 2000] described in Section II.

Table 1. Empirical Studies on Web Engineering

FRAMEWORK ACTIVITIES	STUDY	SUBJECT MATTER	RESEARCH METHOD[S]	SAMPLE	MAJOR CONCLUSIONS
Planning	Mendes et al. [2001]	Metrics for estimating the design and authoring effort for Web applications.	Quantitative case study evaluation	43 students]	Several metrics can be used as indicators of effort such as code length, graphics complexity, and reused code length.
Analysis	Ricca and Tonella [2001]	Web sites Diagnosis	Design Research / Case Study	1 web site	Understanding and restructuring Web sites using the Re Web analysis tool.
	Catanio et al. [2003]	Relationship Analysis [RA]	Design Research	N/A	RA helps identifying and modeling relationships in a generic domain.
Engineering (Architectural Design)	Schwabe and Rossi [1995]	Object-Oriented Hypermedia Design Model [OOHDM]	Design Research	N/A	The four steps of the Design Model are presented: domain analysis, navigational design, abstract interface design, and implementation.
	Shibuya et al. [2001]	HTML generation.	Design Research / Case Study	N/A	Web-based applications can be automatically generated from the HMBS created with HySCHARTS computational tool.
	Garzotto et al. [2001]	HDM-Edit: a schema editor that supports the conceptual design of Web applications.	Design Research / Examples and simplified case study	N/A	HDM-Edit can be used as a design-support tool.
	Swabe et al. [2001]	Engineering Web applications for reuse	Design Research	N/A	Reuse design in Web applications

Engineering (Navigation Design)	Schwabe and Rossi [1995]	Object-Oriented Hypermedia Design Model [OOHDM.	Design Research	N/A	Description of the four steps of the Design Model.
	Gehrke and Turban [1999]	Web site implementation success..	Literature review and survey research	47 papers, 40 websites, 130 users	CSF are page loading speed, business content, navigation efficiency, security, and marketing/customer focus.
	Yoo and Bieber [2000]	Relationship Navigation Analysis [RNA].	Design Research / Example of use of RNA.	N/A	RNA provides a way to find a comprehensive set of linking and navigational opportunities.
	Cachero and Pastor [2001]	Conceptual modeling of device-independent Web applications.	Design Research	N/A	Object-orients Hypermedia model [OO-H] extends UML class diagrams with a navigation view.
	Schwabe et al. [2001]	Engineering Web applications for reuse	Design Research	N/A	Web design frameworks as a way to reuse design in Web applications.
Engineering (Interface Design)	Schwabe and Rossi [1995]	Object-Oriented Hypermedia Design Model [OOHDM.	Design Research	N/A	Description of the four steps of the Design Model.
	Rumpradit and Donnell [1999]	User interface design.	Experimental Design	50 students	User interface designs increase user confidence and satisfaction.
	Cachero and Pastor [2001]	Conceptual modeling of device-independent Web applications.	Design Research	N/A	The proposed Object-orients Hypermedia model [OO-H] extends UML class diagrams presentation view.
	Swabe et al. [2001]	Engineering Web applications for reuse	Design Research	N/A	The authors introduce Web design frameworks as a way to reuse design in Web applications. Underlying concepts include interface design.

Engineering (Content Design)	Gehrke and Turban [1999]	The determinants of a successful web site.	Literature review and survey research	47 papers, 40 websites, 130 users	The major categories of determinants of a successful web site are: page loading speed, business content, navigation efficiency, security, and marketing/customer focus.
	Kuhnke et al. [2000]	Document engineering: addresses all aspects that belong to the contents of pages, text and images.	Design Research	N/A	Present an approach to web engineering focusing on the document engineering part which comprises the activities document design, authoring, and document production.
Engineering (Production)	Gehrke and Turban [1999]	The determinants of a successful web site.	Literature review and survey research	47 papers, 40 websites, 130 users	The major categories of determinants of a successful web site are: page loading speed, business content, navigation efficiency, security, and marketing/customer focus.
	Kuhnke et al. [2000]	Document engineering: addresses all aspects that belong to the contents of pages, text and images.	Design Research	N/A	Present an approach to web engineering focusing on the document engineering part which comprises the activities document design, authoring, and document production.
Page Generation & Testing	Kallepalli and Tian [2001]	Measuring and modeling usage and reliability for statistical web testing.	web site survey	1 web site	Statistical testing and reliability analysis can be used to assure the quality of web applications.
Customer Evaluation	Katerattanakul and Siau [1999]	Information quality of web sites.	Theory building/ Questionnaire	64 students	Proposed an information quality framework.
	Gehrke and Turban [1999]	Web site critical success factors.	Literature review and survey research	47 papers, 40 websites, 130 users	CSF includes page loading speed, business content, navigation efficiency, security, and marketing/customer focus.
	Zhang et al. [1999]	Web user interface design and evaluation.	Literature review and theory building.	N/A	User's satisfaction/dissatisfaction with a Web interface depends on features in the web environment, the user's characteristics, and the information seeking task [strategies] involved.
	Spiliopoulos [2000]	Web usage mining for web site evaluation.	The WUM [Web Usage Miner] data miner was applied on several web sites.	N/A	Web usage mining for web sites allows for the extraction of knowledge from a Web server log, which can ultimately allow for improvements to the web site.

ASSESSMENT OF WEB ENGINEERING RESEARCH

In 1998, Ginige, Murugesan, Deshpande, and Hansen at the University of Western Sydney, Australia, established Web Engineering as a new discipline [Ginige and Murugesan 2001]. These same authors later presented a paper arguing for the emergence of Web Engineering as a new discipline which was conveniently titled "Web Engineering: A New Discipline for Development of Web-based Systems" [Murugesan et al. 1999]. The main objective of their paper consisted of giving an introductory overview of Web Engineering and to promote the new discipline among Web-based system developers, researchers, academics and students. Among other things, the authors present the principle roles of Web Engineering and identify its key activities, approaches and methods.

In similar papers, Ginige and Murugesan [2001a] again provide an overview of the emerging field in stressing the problems plaguing large Web-based projects and the differences between Web Engineering and software engineering. They also address what they refer to as being the essence of Web Engineering: the successful management of the diversity and complexity of Web application development and the avoidance of potential failures that can have serious implications [Ginige and Murugesan 2001b].

Deshpande and Hansen [2001] similarly discuss Web Engineering's classification, define its characteristics, and contrast its present issues with previous problems in information technology. Their main aim in writing the paper consisted of the positioning of Web Engineering as a distinct field among others such as computer science, software engineering, and engineering in itself. The essence of the paper is clearly captured in its title: "Web Engineering: Creating a Discipline among Disciplines" [Deshpande and Hansen 2001].

PLANNING OF WEB-BASED SYSTEMS AND APPLICATIONS

Among the papers classified and reviewed, none addressed the formulation as a separate activity. It was, implicitly included in the planning phase.

Mendes et al. [2001]. Although few studies address the planning phase for Web Engineering, this study focused on a set of size metrics and cost drivers for effort prediction in developing web sites. The metrics proposed characterize Web application size from two different perspectives: length and complexity. The prediction models derive from linear regression and stepwise multiple regression. A quantitative case study evaluation was used to measure several metrics of web applications and the effort involved in web design and authoring. The participants of the case study were computer science students who were asked to develop an educational web application with fewer than 50 pages. To collect the case study data, two questionnaires were handed out to the students: the first asked them to rate their web development experience on a scale of one to five, and the other to measure characteristics of the Web applications and the effort they put into developing them. The study ultimately identified key metrics statistically correlated with development effort, which they grouped as size metrics, complexity metrics, structure metrics, effort metrics, and reuse metrics. These metrics included total page allocation, total code length, total reused code length, graphics complexity, and connectivity.

ANALYSIS OF WEB-BASED SYSTEMS AND APPLICATIONS

At the analysis phase of the Web Engineering process model, four types of analysis are conducted: content analysis, interaction analysis, functional analysis, and configuration analysis.

Ricca and Tonella [2001]. While considering web site analysis, investigated automatic and semiautomatic web site analysis with a tool they developed named Re Web, which focuses on a site's architecture and evolution. Their main contribution consists of defining a set of analyses that extracts a description of a web site's main architectural features and its evolution through time. Using Re Web, they model a web site's structure through various graphs, which ultimately allow them to apply several known algorithms such as flow analysis, traversal algorithms, and pattern matching. Modeling and analysing web sites in this manner can provide developers with useful high-level structural view of the web site at any given time, thus given them a clearer picture of the web site's evolution and changing structure though time.

Catanio et al. [2003] Because neither structured nor object-oriented analysis techniques exist for identifying relationships/links explicitly during the analysis stage of Web development [Yoo and Bieber 2000], this paper proposes a systematic, domain-independent analysis technique. Called Relationship Analysis [RA], this technique focuses exclusively on a system's relationship structure. RA serves two major purposes.

- It helps analysts develop a deeper understanding of the application domain through making the relationships explicit.
- It results in a richer and fuller application analysis and design.

Furthermore, the paper suggest that while the analyst is working with the user in creating various use cases to understand the different functionalities required of the system, he or she could also be conducting Relationship Analysis and documenting it as part of the elicitation process. The ultimately aim of the technique is to fill the gap in other design artefacts such as ER diagrams and object-class diagrams. The diagrams capture an important, but often rather limited subset of relationships, leaving much of the system's structure out of the design and thus out of the system model. Consequently the method does not convey all the useful information that could be passed on to the system's designers and developers.

ENGINEERING OF WEB-BASED SYSTEMS AND APPLICATIONS

Architectural Design

The architectural design of a Web-based system or application involves the design of the overall hypermedia structure of the application or system and the application of design patterns and constructive templates to populate the structure and ultimately achieve reuse.

Shibuya et al. [2001]. In accordance with the Web Engineering architectural design phase, this paper proposes a Hypermedia Model Based on Statecharts [HMBS] as a model suitable for specifying highly structured hyperdocuments. The authors seek to bridge the gap between the design and implementation stages for developing Web applications. Essentially, they present and diuscuss an approach for automatic HTML code generation from formal HMBS hyperdocument specifications. HySCharts, an environment that supports the authoring of hyperdocuments based on the HMBS model is proposed as an environment to support hypermedia authoring using HMBS. In short, automatically deriving an HTML implementation from an HMBS hyperdocument model defined in HySCharts is presented.

Schwabe and Rossi [1995]. With the Object-Oriented Hypermedia Design Model [OOHDM], a hypermedia application is built in a four-step process supporting an incremental or prototype process model with each step focusing on a particular design concern with an object-oriented model ultimately built. These four steps consist of:

- domain analysis,
- navigational design,
- abstract interface design, and
- implementation.

Domain analysis involves building a conceptual model of the application domain using object-oriented principles. This phase results in a conceptual schema built out of subsystems, classes, and relationships. Navigational Design entails describing the navigational structure of the hypermedia application in terms of navigational context, which are induced from navigation classes such as nodes, links, indices, and guided tours. The abstract interface model is put together by defining perceptible objects in terms of interface classes. Implementation maps interface objects to implementation objects.

Garzotto et al. [2001]. Additional tools were developed and proposed as design-supporting tools for the architectural design of Web-based systems. HDM-Edit, a schema editor that supports conceptual design of Web or hypermedia applications as specified by the Hypermedia Design

Model [HDM] is one such tool. HDM-Edit is a smaller component of a larger environment called the JWeb system, which supports all phases of Web prototyping development. HDM-Edit promotes design reuse by support various navigation patterns as built-in modeling primitives and by allowing designers to derive the schema of a specific application as a set of hyperviews on the design of a family of applications [application framework]. HDM-Edit includes several features which are examined by Grozotto et al. It supports

- the latest version of the HDM model;
- design-by-reuse;
- developing families of application;
- producing high quality design documentation, as part of the design process;
- communication via XML files, with other design and implementation tools [Garzotto et al. 2001].

Schwabe et al. [2001]. This paper introduces Web design frameworks, a set of generic designs, as a way to reuse design in Web applications. The authors use OOHDM as a conceptual framework within which to discuss types of reuse. The most extensive types of reuse is achieved through complete application architectures specified using OOHDM-Frame [an extension of OOHDM]. The authors discuss the relation between design frameworks and Web application frameworks, arguing that although both must be integrated, the former can be specified independently of the latter. The underlying concepts addressed are reusing conceptual design, navigation design, abstract interface design, and implementation aspects of software development. The authors' examples are based on examples of conference paper review systems.

Navigation Design

In the navigation design phase, designers must create and design navigation pathways that will eventually enable a user to access Web application content and services. Designers must therefore identify the semantics of navigation for different potential users of the site and define the mechanics (syntax) of achieving the navigation [Pressman 2000].

Schwabe and Rossi [1995]. These authors also address this concern in their Object-oriented Hypermedia Design Model [OOHDM]. In particular, the navigational design step of their model describes the navigational structure of a hypermedia application in terms of navigational contexts, which are induced from navigation classes such as nodes, links, indices, and guided tours.

Yoo and Bieber [2000]. In addressing the need for designers to think of an application in terms of its relationships, links, and navigational structure, Yoo and Bieber proposed the Relationship Navigation Analysis [RNA] method. RNA provides system analysts with a systematic technique for determining the relationship structure of an application. RNA allows identifying of potentially useful relationships in application domains, which can be later, implemented as links. On top of these links, RNA allows for the identification of appropriate navigational structures. Ultimately, RNA aims for the enhancement of developers' implementations by including additional links, meta-information and navigation through the use of the Relationship Navigation Analysis method . A Relationship Navigation Analysis can be conducted either for describing a system or domain, or for analysing a system being designed. The RNA method consists of the following steps (excluding the final step in the case of the description of an existing system):

- stakeholder analysis,
- element analysis,
- relationship analysis,
- navigation analysis, and
- evaluation analysis.
-

It is at the navigation analysis phase that RNA fully addressed navigation design in identifying possible navigation structures for each stakeholder. In this step, analysts think of each element of interest in terms of how the stakeholder might usefully access it.

Gehrke and Turban [1999]. In addressing the determinants of a successful web site, the address the notion of navigation efficiency and its importance when it comes to users' interactions with the web sites they consult. The authors conducted a literature review of 47 papers and over 40 web sites that yielded over 100 issues relating to web site design. These issues were then classified into five major categories:

- page loading speed,
- business content,
- navigation efficiency,
- security, and
- marketing/customer focus.

To assess the relative importance of each category, the number of citations relating to each category was counted. Moreover, a survey among 130 e-commerce customers and potential customers was conducted to find the collective prioritization of the importance of the above categories . The paper makes several recommendations, including use well labelled accurate links, avoid the use of frames, keep navigation consistent, and use site maps/guides if the site contains many underlying pages.

Cachero and Pastor [2001]. The authors propose the Object-Oriented Hypermedia method [OO-H] as an object-oriented approach that captures relevant properties involved in modeling and implementing Web application interfaces. The OO-H method design process involves constructing two additional views which complement those captured in traditional conceptual modeling approaches that comply with the Unified Modeling Language [UML]: the navigational view and the presentation view. The navigational view extends a class diagram with hypermedia navigation features, a complementary diagram. The navigational access diagram [NAD] defines a navigation view and reflects the information, services, and required navigation paths for a user's navigation requirements fulfillment . They propose a CASE tool, which provides an operational environment that supports all the methodological aspects of the OO-H method. Once both views are refined, the CASE tool can generate a Web application front-end – either static or dynamic – for the desired environment such as HTML, JavaServer Pages [JSPs], and Active Server Pages [ASPs]

Schwabe et al. [2001]. As mentioned above, this paper also addresses navigational modeling in the context of the OOHDM-Frame model as OOHDM extends UML for defining node and link classes.

Interface Design

For many, the Web is often synonym with graphical interface. As the Web grows in terms of Web page content, presentation, and complexity, user interface design issues become an important concern of Web developers.

In their OOHDM paper, Schwabe and Rossi [1995] address this concern at the abstract interface design step of their framework. Interface design issues at this step involve the modeling of perceptible objects, implementing chosen metaphors, and describing interfaces for navigational objects.

Rumpradit and Donnell [1999] conducted an empirical study on the usability of user interfaces on the Web. They examined the differences between different users interacting with web sites and determined whether the user interface elements would have a significant impact on the performance (speed and accuracy), confidence, and satisfaction of web site users. User interface elements were evaluated in terms of four measures of effectiveness:

- speed [the amount of time required by a subject to complete a set of tasks],
- accuracy [the amount of correct answers on searching the information],
- confidence [the subject's confidence when searching for information]
- satisfaction between plain hypertext and four types of user interface techniques:
    1. an index

2. an imagemap
3. an index with a context path
4. an imagemap with a context path

This study involved 50 first-year students from The Washington University. The independent variable, learning style preference, was measured by using the Learning Style Inventory adapted from Suessmuth [1985]. Three categories of subjects were tested:

- a group of visual style preference subjects,
- a group of auditory style preference subjects, and
- a group of a combination of visual, auditory, and kinesthetic, learning style preference subjects.

The dependent variable data were collected using an experimental design with students randomly assigned to each type of user interface techniques [other independent variable] and to a set of tasks conducted during the second phase of the study. The subjects were asked to perform a series of tasks to evaluate the performance of the system design, in this case The Washington University web site. The study found

- a statistically significant relationship between user performance in completing given tasks and user interface designs and

- user interface designs with a context path feature increase user confidence and satisfaction more than other interface designs.

Cachero and Pastor's [2001] OO-H method discussed previously also addressed interface design, one of the additional views that use the different elements regarding interface appearance and behaviour to model a series of interconnected templates structures expressed in Extensible Markup Language [XML].

Content Design and Production

When it comes to the content design of Web-base systems and applications, Gehrke and Turban [1999], in addressing the determinants of a successful web site, discuss the business content and marketing/customer focus of web sites. Based on the literature review conducted in their study, the authors include the following recommendations about business content:

- provide contact information on each page,
- provide free services and useful information, and create a frequently asked question [FAQ] section.

For the marketing/customer focus aspect of web sites, the following advice is presented:

- write the web site in English or give the English choice because English is recognized is the worldwide universal business language,
- provide as many payment options as possible, and
- create a catalogue of products and/or services.

Kuhnke et al. [2000] present a two-fold approach to Web engineering:

1. the static part addresses all aspects that belong to the content of pages, text and images,
2. the dynamic part deals with interaction with databases and other application systems.

The static part, called document engineering, involves three major tasks:

1. document design which defines the quality, writing standards, and modular text structures of the desired document,
2. authoring which concerns text, tables, and images, and

3. document production which deals with the final publishing process, mapping the results of authoring and design onto a concrete layout .

Page Generation and Testing of Web-based Systems and Applications

While page generation consists in the merging the content defined in the engineering activity with the architectural, navigation, and interface designs to produce executable Web pages, testing requires some of the following activities:

- processing components of web pages are unit tested,
- the assembled system is tested for overall functionality and
- content delivery, and the Web system or application is implemented in a variety of different environmental configurations and tested for compatibility with each configuration [Pressman 2000].

Although no articles reviewed explicitly addressed the page generation phase in the context of Web engineering as defined by Pressman [2000], Pressman does mention that integration with component middleware such as CORBA, DCOM, and JavaBeans is also accomplished during page generation .

Kallepalli and Tian [2001]. For testing of Web-based systems and applications, use statistical testing and reliability analysis to assure quality of Web applications. The authors extract Web usage and failure information from existing Web logs. The usage information forms the basis for models for statistical Web testing. The related failure information is used to measure the reliability of Web-based systems and applications and the potential effectiveness of statistical Web testing. This approach was applied by the authors to the web pages of the School of Engineering and Applied Science at the Southern Methodist University. The models used are the Unified Markov Models [UMM]. By providing UMM construction guidelines and developing utility programs in Perl to analyze Web logs, the authors created various tools to support their reliability analysis for Web applications. Applying these tools to a university web site gave positive results and point to the viability of the authors testing strategy.

Customer Evaluation of Web-based Systems and Applications

Katerattanakul and Siau [1999]. Each increment of the Web Engineering process is reviewed during the customer evaluation activity. The authors address the evaluation of the content aspect of Web-based systems in proposing a framework and developing an instrument to measure the information quality of web sites. The framework consists of four major information quality categories:

- intrinsic information quality,
- contextual information quality,
- representational information quality, and
- accessibility information quality.

The main dimension of intrinsic information quality is the accuracy of the information. Contextual information quality highlights the requirements that information quality must be considered within the context of a given task. To add value to a specific task, the information must be relevant to that particular task. Representational information quality involves aspects related to the format of the information and its meaning. To obtain accessibility information quality, the information system must be accessible but secure. Based on these four information quality categories, the authors develop a questionnaire to test the importance of each of these newly developed categories and determine how Web users asses the quality of individual web sites. Research results showed that the framework needs further refinements in terms of for instance potential interpretation and the wording of each questions.

Gehrke and Turban [1999] also address the customer evaluation phase of Pressman's [2000] Web Engineering process model. As mentioned previously, they identify five major determinants of web site design: page loading speed, business content, navigation efficiency, security, and marketing/customer focus which can be addressed at each increment of the Web Engineering process.

Zhang et al. [1999] investigate the user satisfaction aspect of Web-based systems and applications. They provide a conceptual framework and foundation for systematically investigating features in the Web environment which contribute to user satisfaction with a Web interface. Their research is grounded in past theoretical models. They use Herzberg's [1966] motivation-hygiene theory to identify and distinguish the features that may be considered hygiene features (features that are necessary but not sufficient to ensure user satisfaction with a Web user interface) from those that could be considered motivators (features that contribute to user satisfaction with and a continued use of a web site) in Web environments. The authors propose that a user's satisfaction or dissatisfaction with a Web interface is the result of the interplay among three components: features in the Web environment, user characteristics, and the information seeking tasks. They relate Herzberg's motivation-hygiene theory to the features in the Web environment and user characteristics.

Spiliopoulos [2000] also emphasises the importance of web site evaluation and determining what the users want from web sites, what they like, and what distracts them. Spiliopoulos points to the evaluation of a web site based on the data automatically recorded on it though Web server logs; all traces left by the users of a web site being stored in these logs. Data mining is the methodology used to extract of knowledge from data. Data mining can be performed to optimize the performance of a web server, to discover which products are being purchased together, or to identify whether a web site or web page is being used as expected. WUM [Web Usage Miner] is presented as an example of a tool designed to meet demands of web site evaluation and navigation pattern discovery. The discovery of web site usage patterns or navigation patterns using tools such as WUM allow for web site redesigns based on user behaviours and preferences. Ultimately, navigation patterns reflect how a site is perceived by different groups of users and can eventually lead web site owners to personalize their web sites based on this knowledge (services for adaptive page generation).

IV. DISCUSSION AND CONCLUSIONS

Over the past five years, a large amount of research was published that addresses issues related to developing Web-based systems and applications. This paper reviews this work with a focus on mapping progress to date and suggesting future research directions. This paper provides a preliminary classification and review of Web Engineering research through the use of the Web Engineering Process Model proposed by Pressman [2000]. Table 2 summarized the number of articles found for each element of the Pressman model.

Table 2. Web Engineering – An Overview of Progress to Date

Web Engineering Framework Activities		**Number of Reviewed Papers Addressing this Activity**
Formulation		0
Planning		1
Analysis		2
Engineering	Architectural Design	4
	Navigation Design	5
	Interface Design	4
	Content Design	2
	Production	2
Page Generation & Testing		1
Customer Evaluation		4
Total Number of Articles		**25**

It is clearly evident, based on this preliminary review that much of the small body of Web Engineering research centers on the engineering activity. The architectural design, navigation design, and interface design tasks of the engineering activity and customer evaluation seem to draw more research. Almost 70% of the published research is concerned with the engineering part of Web-based systems development process. Two possible reasons for these findings are:

1. The most significant research disciplines for Web Engineering are Network Engineering, Software Engineering, Databases and Storage Systems, and Hypermedia [Whitehead, 2002]. Since it is an engineering discipline, one may expect research on Web Engineering to draw more attention among researchers from these disciplines.
2. Much attention focused on the engineering activity of Web-based systems development because curriculum programs in Web Engineering are developed by Software Engineering departments in most universities [Faulk, 2000].

The page generation & testing activity, which can reasonably be considered an important phase of any system development project, did not yet garner much academic attention. The formulation, planning and analysis activities also lag behind compared with other engineering activities. This result is not surprising. A decade ago, Barki et al. [1993] have conducted a large survey on 120 systems development projects. Based on five life cycle stages, Barki et al. found that risk of failure of a particular project increases from the preliminary analysis to detailed analysis but decreases during the remaining stages. The same authors later tested a research model with longitudinal data obtained from project leaders and key users of 75 software projects [Barki et al. 2001]. High-risk projects were found to call for high information processing capacity approaches in their management.

This paper's classification and review of Web Engineering research provides insight on the critical elements in addressing Web-based systems development projects. To avoid or attenuate risk of failure of these systems, more research is needed particularly in the formulation, planning and, analysis, phases. Furthermore, the framework used in this paper to examine research on Web Engineering could be extended bearing in mind that many crucial decisions are related to risk assessment and management of Web-based systems. It is possible to include and examine these critical parts within Pressman's framework of Web Engineering.

Clearly, more research is needed in the area of Web Engineering. Theory-based and empirical work needs to be conducted because Web-based systems development projects are growing rapidly [Schwabe, 2001]. For a long time, IS researchers examined issues related to systems development (including transaction systems, decision support systems) and built a cumulative body of research. It is, indeed, an opportunity for researchers interested in Web engineering to capitalize on prior research to examine Web Engineering critical issues.

Editor's Note: This article was received on May 23, 2003 and was published on August 27, 2003. It was with the authors two weeks for one revision.

REFERENCES

Barki, H., S. Rivard, and J. Talbot (1993) "Toward An Assessment of Software Development Risk", Journal of Management Information Systems, (10)2, pp. 203-227.

Barki, H., S. Rivard, and J. Talbot (2001) "An Integrative Contingency Model of Software Project Risk Management", Journal of Management Information Systems, (17)4, pp. 37-69.

Bernstein, M. "Patterns in Hypertext" Proc. 9th ACM Conf. Hypertext, ACM Press, 1998, pp. 21-29.

Cachero, C. and O. Pastor (April-June 2001) "Conceptual Modeling of Device-Independent Web Applications", *IEEE Multimedia*, (8)2, pp. 26-39.

Catanio, J. et al. "Relationship Analysis: A Research Plan for Enhancing Systems Analysis for Web Development", *Proceedings of the 36th Hawaii International Conference on System Sciences*, 2003.

Chi, E. (2002) “Imporving Web Usability Through Visualization”, *IEEE Internet Computing*, pp. 64-71.

Cutter Consortium Research Briefs (Nov. 7, 2000). http:\\www.cutter.com

Deshpande, Y. and S. Hansen. (April-June 2001) “Web Engineering: Creating a Discipline Among Disciplines”, *IEEE Multimedia*, (8)2, pp. 82-87.

Fewster, R. and E. Mendes. “Measurement, Prediction and Risk Analysis for Web Applications”, *Proceedings of IEEE Metrics Symposium - 7th International Software metrics Symposium, IEEE CS Press*, 2001.

Garzotto, F et al. “Supporting Reusable Web Design With HDM-Edit”, *Proceedings of the 34th Hawaii International Conference on System Sciences*, 2001.

Gehrke, D. and Turban, E. “Determinants of Successful Web Site Design: Relative Importance and Recommendations for Effectiveness”, *Proceedings of the 32nd Hawaii International Conference on System Sciences,* 1999.

Ginige, A. and Murugesan, S. (2001a) “Web Engineering: An Introduction”, *IEEE Multimedia*, (8)1, pp. 14-18.

Ginige, A. and Murugesan, S. (2001b) “The Essence of Web Engineering: Managing the Complexity and Diversity of Web Application Development”, *IEEE Multimedia*, (8)1, pp. 22-25.

Gnaho, C. and F. Larcher. “A User Centered Methodology for Complex and Customizable Web Applications Engineering” Proc. First ICSE Workshop on Web Engineering, ACM, Los Angeles, May 1999.

Herzberg, F. (1966) *Work and the Nature of Man*. World Publishing, NY, pp. 71-91.

Isakowitz, T., A. Stohr, and P. Balasubramanian (1995) “RMM: A Methodology for Structured Hypermedia Design”, *Communications of the ACM*, (38)8, pp. 34-44.

Kallepalli, C. and J. Tian (2001) “Measuring and Modeling Usage and Reliability for Statistical Web Testing”, *IEEE Transactions on Software Engineering*, (27)11, pp. 1023-1036

Katerattanakul, P. and K. Siau “Measuring Information Quality of Web Sites: Development of an Instrument”, *Proceedings of the 19th International Conference on Information Systems*, 1999.

Klapsing, R., G. Neumann, and C. Wolfram (2001) “Semantics in Web Engineering: Applying the Resource Description Framework”, *IEEE Multimedia*, (8)2, pp. 62-68.

Kuhnke, C., J. Schneeberger, T. Andreas, “A Schema-Based Approach to Web Engineering”, *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000.

Lynch, P.J. and S. Horton. Web Style Guide: Basic Design Principles for Creating Web Sites, Yale University Press, 1999.

Menasce, D. (2002) “QoS Issues in Web Services”, *IEEE Internet Computing*, pp. 72-75.

Mendes, E., N. Mosley, and C. Steve (2001) “Web Metrics- Estimating Design and Authoring Effort”, *IEEE Multimedia*, (8)1, pp. 50-57.

Murugesan et al., “Web Engineering: A New Discipline for Web-Based System Development”, *Proc. First Int’l Conf. Software Engineering [ICSE] Workshop on Web Engineering*, University of Western Sydney, Australia, 1999.

Mylonopoulosm, A. and V. Theoharakis (2001) “On Site: Global Perceptions of IS Journals, *Communications of the ACM*, (44)9, pp. 29-33.

Nanard, M. and P. Kahn. “Pushing Reuse in Hypermedia Design: Golden Rules, Design Patterns and Constructive Templates”, Proc. Ninth ACM Conf. On Hypertext and Hypermedia, ACM Press, 1998, pp. 11-20.

Nielsen, J. and A. Wagner. “User Interface Design for the WWW”, Proc. CHI’96 Conference on Human FACTORS IN computing Systems, ACM Press, 1996, pp. 330-331.

Olsina, L., et al. “Specifying Quality Characteristics and Attributes for Web Sites”, Proc. First ICSE Workshop on Web Engineering, ACM, Los Angeles, MAY 1999.

Powell, T.A. Web Site Engineering, Prentice-Hall, 1998.

Powell, T.A. Web Site Engineering: Beyond Web Page Design, Prentice Hall, Upper Saddle River, N.J., 1998.

Pressman, R.S. Software Engineering: A Practitioner’s Perspective, 5th edition, McGraw-Hill, New York, 2000.

Ricca, F. and P. Tonella (2001) “Understanding and Restructuring Web sites with Re Web”, *IEEE Multimedia*, (8)2, pp. 40-51.

Rumpradit, C. and L. Donnell, “Navigational Cues on User Interface Design to Produce Better Information Seeking on the World Wide Web”, *Proceedings of the 32nd Hawaii International Conference on System Sciences*, 1999.

Schwabe, D. et al. (2001) “Engineering Web Applications for Reuse”, *IEEE Multimedia*, (8)1, pp. 20-31.

Schwabe, D. and G. Rossi (1995) “The Object-Oriented Hypermedia Design Model”, *Communications of the ACM*, (38)8, pp. 45-46.

Schilit, B et al. (2002) “Web Interaction Using Very Small Internet Devices”, IEEE Computer, pp. 37-45.

Shibuya, R. et al. “Automatic HTML Generation from Formal Hypermedia Specifications”, *Proceedings of the 34th Hawaii International Conference on System Sciences,* 2001.

Singh, M. (2001) “The Invisible Wireless Web”, *IEEE Internet Computing*, pp. 6-7.

Spiliopoulos, M. (2000) “Web Usage Mining for Web Site Evaluation”, *Communications of the ACM*, (43)8, pp. 127-134.

Whitehead, E.J. (2002). A Proposed Curriculum For A MASTER in Web Engineering. *Journal of Web Engineering*, (1)1, pp. 18-22.

Whitley, E.A., “Method-ism in Practice: Investigating the Relationship between Method and Understanding in Web Page Design”, *Proceedings of the 19th International Conference on Information Systems*, 1999.

Yoo, J. and M. Bieber, “Towards a Relationship Navigation Analysis”, *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000.

Zhang, P. et al. “Websites that Satisfy Users: A Theoretical Framework for Web User Interface Design and Evaluation”, *Proceedings of the 32nd Hawaii International Conference on System Sciences*, 1999.

ABOUT THE AUTHORS

Bouchaib Bahli is Assistant Professor of Management Information Systems in the John Molson School of Business at Concordia University. His Research interests include Risk Management of Software Development Projects, Information Technology Outsourcing and, e-Business Security Systems. Dr. Bahli’s publications appear in the *International Conference on Information Systems, and the Hawaii International Conference on System Sciences*, among others. He is an experienced systems developer.

Dany Di Tullio is a system analyst at BOC Group Canada, Montreal. He specializes in Web-based systems development and holds a Master in Information Systems from John Molson School of Business at Concordia University.