

# Five Years of Web Site Evolution

## Invited Paper

Scott Tilley

Department of Computer Sciences  
Florida Institute of Technology  
stilley@cs.fit.edu

### Abstract

*From its inception in 1999 to its fifth anniversary in 2003, the Web Site Evolution (WSE) series of workshops have provided a forum for researchers and practitioners to present original work on subjects related to the disciplined evolution of large-scale Web sites. This paper summarizes some of the key themes that have appeared in the first five years of WSE events. Emerging research challenges for the next few years are also outlined.*

**Keywords:** Web site evolution, themes, research

## 1. Introduction

Consider the following scenario. A new employee of a young and rapidly growing company is given the task of adding e-commerce capabilities to the company's Web site. The employee is a recent computer science graduate, but is unfamiliar with the structure of the Web site and has little experience with some of the technology involved. For example, the Web site uses a library of Perl scripts to implement CGI (Common Gateway Interface) applications, but the traffic on the site has increased sufficiently that the CGI interface is no longer meeting requirements due to performance problems. Upon further exploration of the site, the employee finds that the company was one of the early adopters of CORBA (Common Object Request Broker Architecture), a distributed object technology [5]. However, the vendor that the company selected to provide the CORBA services shipped a product (possibly a beta release) that was flawed. Further complicating matters, the vendor has gone out of business.

In many ways, this scenario is similar to a traditional maintenance scenario. As software ages, the task of maintaining it becomes more complex and more

expensive. Poor design, unstructured programming methods, and crisis-driven maintenance can contribute to poor code quality, which in turn affects understanding. Better understanding of a program aids in common activities such as correcting defects, reengineering systems, and keeping documentation current. To minimize the likelihood of introducing errors during the change process, the software engineer must understand the system sufficiently well so that changes made to the source code have predictable consequences. But such understanding is difficult to recover from a legacy system after several years of operation. In the prevailing atmosphere of accelerated schedules, years are reduced to months or even weeks.

The main difference in the scenario outlined above and a "traditional" maintenance task is the application undergoing maintenance: a large-scale Web site. It is a prime example of a modern system destined to be a legacy system of tomorrow. Indeed, many Web sites can already be classified as a legacy system, given their age, size, and constituent components. In the past, the subject system might have been a monolithic, mainframe-based payment processing system written in COBOL. Or it might have been a two-tier client/server system written in C/C++. For the Web application, there is a much wider range of implementation languages, heterogeneous networked applications, scripts on both the client (e.g., JavaScript) and on the server (e.g., Perl), and several other dynamic aspects of the system's architecture and implementation that make it more difficult to understand, and hence more difficult to evolve in a disciplined manner.

Since its inception in 1999, the objective of the Web Site Evolution (WSE) series of workshops has been to foster the development, validation, and dissemination of a disciplined approach to problems such as the one described above. The maintenance scenario of adding e-commerce capabilities to an existing Web site is certainly

representative, but it is by no means exhaustive. Each WSE workshop has focused on a different aspect of the very challenging problem of evolving Web sites to meet new business requirements in a systematic manner.

The rationale for the creation of WSE in 1999 was to serve the needs of the nascent community of researchers and practitioners who were focused on this sub-topic of system evolution. Established venues such as the International Conference on Software Engineering (ICSE), the International Conference on Software Maintenance (ICSM), the Working Conference on Reverse Engineering (WCRE), the International Workshop on Program Understanding (IWPC), and the Reengineering Forum (REF) were beginning to accept papers in the area of Web site evolution. However, there was not widespread recognition of the area as a true discipline in its own right. WSE was formed as a separate entity to completely focus on this emerging research area.

The following five sections of the paper summarize each of the WSE workshops that have occurred so far, from the first WSE in 1999 to the current WSE in 2003. The final section of the paper outlines some of the emerging research challenges for the next few years in the area of Web site evolution.

## 2. WSE 1999

The first WSE workshop was held on October 5, 1999 in Atlanta, GA. The workshop was co-located with the 6<sup>th</sup> Working Conference on Reverse Engineering (WCRE). There were 14 papers accepted for the inclusion in the WSE workshop proceedings. The proceedings themselves were informally published by the University of California, Riverside, and had a limited copies distributed (workshop participants only) [6].

While there was no official theme for WSE 1999, the Call for Papers stated the following:

*The goal of this one-day workshop is to bring together members of the software engineering, reverse engineering for program understanding, and information technology communities to focus on techniques for Web site evolution. Web sites have moved from supplementary mechanisms for communication to become a primary component of most organization's infrastructure. Rather than serving simply as a passive means of disseminating information, Web sites have become the integration hub for a wide variety of activities, including electronic commerce, streaming media, and online collaboration.*

The rationale for this inaugural WSE workshop was that with “older” Web sites having just passed five years on the Net, it seemed prudent to examine how they could evolve in a more disciplined manner. As Web sites age, they suffer from some of the same afflictions as any complex software system: their structure degrades, maintenance becomes increasingly problematic, and legacy applications and interfaces hinder evolution. However, Web sites are also unique in several aspects. For example, people who may lack a formal computer science or software engineering background often develop them.

WSE 1999 provided a forum for the presentation of new research and empirical results in areas such as:

- Managing an evolving 3-tier information system
- Using Web tools for system understanding
- Leveraging traditional program analysis tools
- Integrating dynamic and static content
- Investigating emerging technologies, such as Dynamic HTML, E-Commerce, and XML

There were 23 people who participated in the workshop discussions, which were structured according to the topics of “Software Engineering Issues,” “Technology Issues,” and “Reuse and Product Lines.”

## 3. WSE 2000

The second WSE workshop was held on March 1, 2000 in Zürich, Switzerland. The workshop was part of the 2<sup>nd</sup> Reengineering Forum Europe (EuroREF), which met jointly with the 4<sup>th</sup> European Conference on Software Maintenance and Reengineering (CSMR). There were 8 papers accepted for the inclusion in the workshop proceedings, which were again informally published by the University of California, Riverside. Workshop participants received a printed copy, with electronic versions also available [7].

The unstated theme for WSE 2000 was “Web-Enabling Legacy Systems.” This focus was captured in the Call for Papers:

*The goal of this one-day workshop is to bring together members of the software maintenance, reengineering, and information technology communities to discuss techniques for Web-enabling legacy systems. One way of extending the lifespan of existing applications is to make them available over the Web. While there are numerous tools available to aid in this process, significant challenges remain. In addition, once migrated to the Web, legacy systems are often augmented with new capabilities, such as*

*electronic commerce and streaming media. Such additions add to the complexity of the Web-enabled system, making its subsequent evolution even more difficult.*

The workshop topics included:

- Extracting and reusing components from legacy systems
- Using distributed component technology in migrating legacy systems
- Managing an evolving  $n$ -tier information system
- Case studies of large-scale evolution
- Business issues in migrating to the Web

The workshop opened with a keynote presentation by Harry Sneed on the subject of “Accessing Legacy Mainframe Applications via the Internet.” Later in the day, Hausi Müller gave a mini-tutorial on “Software Migration Strategies.” The 20 workshop participants spent the rest of the day in discussions that were clustered into three sessions of short paper summaries loosely related to the topics listed above.

#### 4. WSE 2001

The third WSE workshop was held on November 10, 2001 in Florence, Italy. This was the first time WSE was co-located with the International Conference on Software Maintenance (ICSM). This was also the first time that the WSE proceedings were formally published by the IEEE Computer Society Press [8].

The central theme of WSE 2001 was “Access for All.” The rationale for selecting this theme was the beginning of the new millennium, which provided a singular opportunity to view Web sites in a new perspective: as a vehicle for truly universal communication. Such an inclusive definition implied that Web sites should provide comparable experiences to diverse users, irrespective of their national language, physical abilities, or computing platform. In other words, access for all.

The 16 papers included in the WSE 2001 proceedings were clustered according to the topics of “Languages,” “Accessibility,” “Evolution,” and “Devices.” WSE 2001 was also a companion event to the ACM SIGDOC 2001 conference, which took place October 21-24 in Sante Fe, New Mexico [9]. The SIGDOC 2001 theme of “Going Global: Communicating in the New Millennium” dovetailed nicely with that of WSE.

The Call for Papers captured the goals of WSE 2001:

*The goal of this one-day workshop is to bring together members of the communication design,*

*software engineering, and information technology communities to discuss techniques for evolving Web sites from their current condition to a state that meets the criteria outlined above. Architectural styles and tool support for multilingual Web sites is currently quite limited. Expertise in constructing Web pages that are accessible to the disabled is available but not widely utilized. The explosion of non-traditional computing platforms for browsing the Web, such as PDAs, WAP-enabled phones, and Internet appliances, is forcing Web professionals to rethink the separation of form from content.*

The workshop began with a keynote presentation by Cornelia Boldyreff, who talked about “Web Accessibility.” The keynote set the tone for discussions during the remainder of the day. The 25 participants at WSE 2001 engaged in discussions touching on topics such as:

- Migrating to multilingual Web sites
- Enhancing Web sites to make them accessible to the disabled
- Making Web site content available in multiple formats for multiple platforms
- Case studies of large-scale evolution
- Fundamental design issues that transcend the medium

The more formal nature of WSE 2001 (in comparison with the first two workshops), coupled with the crosscutting theme of “Access for All,” attracted a more diverse audience than at WSE 1999 or WSE 2000. The discussions were extremely insightful and interesting.

#### 5. WSE 2002

The fourth WSE workshop was held on October 2, 2002 in Montréal, Canada. WSE 2002 was again co-located with ICSM. The IEEE Computer Society Press published the 11 papers in the WSE 2002 proceedings [1].

The theme of WSE 2002 was “Migrating to Web Services.” In late 2001 and early 2002, the potential of the Web as a vehicle for offering a variety of services was becoming more apparent. Web sites were evolving their content from static information to more dynamic and interactive forms. A number of standards and commercial products were emerging, offering a basis for developing Web-based services from dynamically-composable resources, extending both the range and scope of services that could be made available through the Web.

Papers on research investigations involving evolution with respect to Web services were specifically solicited, as were papers on assuring that Web services are universally accessible and usable. These goals were captured in the Call for Papers:

*The goal of this one-day workshop is to bring together members of the Web design, software engineering, and information technology communities to discuss techniques for migrating to Web services. Architectural styles and tool support for service-based Web sites are currently quite diverse. Expertise in constructing service-based Web pages is quite limited, and experience accounts are lacking. The explosion of non-traditional computing platforms for browsing the Web, such as PDAs, WAP-enabled phones, and Internet appliances, is forcing Web professionals to rethink the separation of form from content and the range and scope of services offered via the Web.*

Some of the topics discussed at WSE 2002 included:

- Migrating to Web services
- Web quality determination and improvement through process management and product control
- Enhancing Web sites to make them more accessible, reliable, and usable
- Analyzing and reverse-engineering Web site content and structure
- Making Web site content available in multiple formats for multiple platforms
- Applying traditional software engineering activities such as architecture, metrics, and testing to Web sites
- Case studies of large-scale Web site reuse, reengineering, and evolution

In addition to the paper presentations, WSE 2002 included an invited talk and a panel session. Jim Isaak delivered the invited talk on the subject of “Web Site Engineering Best Practice Standards (IEEE 2001).” The panel session centered on the question of “What are the characteristics of systems that make them amenable to Web services?” The lively discussion led by the panelists and involving nearly all of the 30 workshop participants was a fitting conclusion to the event.

After the workshop, the Program Chairs selected several papers for consideration of inclusion in a special issue of the *Journal of Software Maintenance and Evolution: Research and Practice* [2]. This was a welcome development that served to publicize the WSE event and the excellent work on the papers’ authors. The special

issue of the journal is scheduled for publication in Fall 2003 [2].

## 6. WSE 2003

This year is fifth WSE workshop. It is being held on September 22, 2003 in Amsterdam, The Netherlands. For the third time, the WSE workshop is co-located with ICSM. The IEEE Computer Society Press is again publishing the WSE proceedings, which for 2002 consists of 11 papers selected by the Program Committee (and this invited paper) [10].

The theme of WSE 2003 is “Architecture.” Software architecture is now established as a discipline of study in its own right. Research into software architecture deals with activities such as extracting architectures from existing systems, modeling architectural aspects of complex applications, and performing tradeoff analysis between quality attributes of proposed architectures. As Web sites evolve from disorganized constructions to carefully engineered systems, issues related to Web site architecture will become increasingly important.

As stated in the Call for Papers:

*The goal of this one-day workshop is to bring together members of the software architecture, system maintenance, and Web engineering communities to discuss architectural issues of large-scale Web sites. For existing Web sites, techniques from traditional software reverse engineering can be used to extract assets for use in the next version of the site. For new Web sites, results from architectural tradeoff analysis can be used to ensure a Web site that will evolve in a disciplined manner.*

The likely topics to be discussed at WSE 2003 include:

- Architectural tradeoff analysis for Web sites
- Case studies of Web site evolution
- Documenting, modeling, and reasoning about Web site architecture
- Techniques for reverse engineering and understanding Web site structure
- Web services, accessibility, and browser technologies

It is expected that approximately 32 people will attend WSE 2003. The 2003 edition of WSE promises to follow the example set by the previous workshops in providing a stimulating and enjoyable event.

## 7. The Future of WSE

Web site evolution is a rich research area. There are enough similarities between large-scale Web sites and traditional legacy systems that many of the proven tools and techniques can be redeployed in this new application area. However, there are also enough differences that the problem requires innovative solutions to the many new problems unique to the discipline.

Of the many interesting topics that warrant more exploration by the community in the context of Web site evolution, there are at least three that deserve focused attention in the near future. The first topic is testing. The field of software testing is well established [3]. It has a solid theoretical foundation, and enjoys considerable respect within the practitioner community. Many testing techniques that are used on non-Web systems are equally applicable to Web sites (although in practice relatively few are actually used). At the same time, Web sites exhibit enough distinct traits that a new body of research has begun to form around it [4]. WSE could help increase the visibility of this important area in the coming years.

The second topic is security, which has become a first-class quality attribute for almost all modern software systems. The popular press is replete with stories of Web sites under concerted attack, sometimes by “script kiddies” and sometimes by more nefarious hackers. Because of the nature of Web site development, designing security into the system is not always done properly. The engineering of security mechanisms is also often woefully inadequate. WSE could serve as a medium for the creation of a “best practices” guide to Web site security, particularly related to retrofitting security concerns into already-deployed (and hence vulnerable) systems.

The third topic is usability. Much has been written about the sorry state of most application software in terms of usability and human-computer interaction. There is an extensive body of research and empirical studies on the subject, but for some reason many Web site designers don’t seem to follow the guidelines. WSE could help improve Web site usability through the sponsorship of further research into the area, particularly in the emphasis for the need of evidence-based assessment techniques.

To paraphrase the teaser found at the end of James Bond films, “WSE will return.” In 2004, the venue for WSE will be Chicago. The workshop will again be co-located with ICSM. The proposed theme is “Testing.” Please join the Steering Committee in making the next five years of WSE even more successful than the first five years!

Further information on events related to WSE can be found online at <http://www.websiteevolution.org/>.

## References

- [1] Boldyreff, C. and Tonella, P. (editors). *Proceedings of the 4<sup>th</sup> International Workshop on Web Site Evolution* (WSE 2002: Oct. 2, 2002; Montréal, Canada). Los Alamitos, CA: IEEE Computer Society Press, 2002.
- [2] Boldyreff, C. and Tonella, P. (editors). *Journal of Software Maintenance and Evolution: Research and Practice – Special Issue on Web Site Evolution*, Fall 2003. John Wiley & Sons, 2003.
- [3] Kaner, C. *Testing Computer Software* (2<sup>nd</sup> Edition). John Wiley & Sons, 1999.
- [4] Nguyen, H. *Testing Applications on the Web: Test Planning for Internet-Based Systems*. John Wiley & Sons, 2000.
- [5] The Object Management Group. “The OMG’s CORBA Website.” Online at <http://www.omg.org/corba/>.
- [6] Tilley, S. (editor). *Proceedings of the 1<sup>st</sup> International Workshop on Web Site Evolution* (WSE 1999: Oct. 5, 1999; Atlanta, GA). Available online at <http://www.websiteevolution.org/1999/>.
- [7] Tilley, S. (editor). *Proceedings of the 2<sup>nd</sup> International Workshop on Web Site Evolution* (WSE 2000: March 1, 2000; Zürich, Switzerland). Available online at <http://www.websiteevolution.org/2000/>.
- [8] Tilley, S. (editor). *Proceedings of the 3<sup>rd</sup> International Workshop on Web Site Evolution* (WSE 2001: Nov. 10, 2001; Florence, Italy). Los Alamitos, CA: IEEE Computer Society Press, 2001.
- [9] Tilley, S. (editor). *Proceedings of the 19<sup>th</sup> Annual International Conference on Systems Documentation* (SIGDOC 2001: Santa Fe, NM; October 21-24, 2001). New York, NY: ACM Press, 2001.
- [10] Wong, K. (editor). *Proceedings of the 5<sup>th</sup> International Workshop on Web Site Evolution* (WSE 2003: Sept. 22, 2003; Amsterdam, The Netherlands). Los Alamitos, CA: IEEE Computer Society Press, 2003.