

Automatic Evaluation of Interfaces on the Internet

Thomas Mandl

University of Hildesheim, Germany

INTRODUCTION

Empirical methods in human-computer interaction (HCI) are very expensive, and the large number of information systems on the Internet requires great efforts for their evaluation. Automatic methods try to evaluate the quality of Web pages without human intervention in order to reduce the cost for evaluation. However, automatic evaluation of an interface cannot replace usability testing and other elaborated methods.

Many definitions for the quality of information products are discussed in the literature. The user interface and the content are inseparable on the Web, and as a consequence, their evaluation cannot always be separated easily. Thus, content and interface are usually considered as two aspects of quality and are assessed together. A helpful quality definition in this context is provided by Huang, Lee, and Wang (1999). It is shown in Table 1.

The general definition of quality above contains several aspects that deal with human-computer interaction. For example, the importance of accessibility is stressed. The user and context are important in human-computer interaction, and the information-quality definition also considers suitability for the context as a major dimension.

The automatic assessment of the quality of Internet pages has been an emerging field of research in the last few years. Several approaches have been proposed under various names. Simple approaches try to assess the quality of interfaces via the technological soundness of an implementation, or they measure the popularity of a Web page by link analysis. Another direction of research is also based on only one feature and considers the quality of free text. More advanced approaches combine evidence for assessing the quality of an interface on the Web. Table 2 shows the main approaches and the discipline from which they originated.

Table 1. Categories of information quality (IQ) (Huang et al., 1999)

IQ Category	IQ Dimensions
Intrinsic IQ	Accuracy, objectivity, believability, reputation
Contextual IQ	Relevancy, value-added, timeliness, completeness, amount of information
Representational IQ	Interpretability, ease of understanding, concise representation, consistent representation
Accessibility IQ	Access, security

Table 2. Disciplines and their approaches to automatic quality evaluation

Approach	Discipline
HTML Syntax checking	Web software engineering
Link analysis	Web information retrieval
Indicators for content quality	Library and information science
Interface evaluation	HCI
Text quality	Human language technology

These approaches are discussed in the main sections. The indicators for content quality have not resulted in many implementations and are presented together with the interface evaluation in the subsection “Page and Navigation Structure.”

BACKGROUND

In the past, mainly two directions of research have contributed to establish the automatic evaluation of Internet resources: bibliometrics and software testing.

Link analysis applies well-known measures from bibliometrics to the Web. The number of references to a scientific paper has been used as an indicator for its quality. For the Web, the number of links to a Web page is used as the main indicator for the quality of that page (Choo, Detlor, & Turnbull, 2000). Meanwhile, the availability of many papers online and some technical advancement have made bibliometric systems for scientific literature available on the Internet (Lawrence, Giles, & Bollacker, 1999). The availability of such measures will eventually lead to an even greater importance and impact of quantitative evaluation.

Software testing has become an important challenge since software gets more and more complex. In software engineering, automatic software testing has attracted considerable research. The success of the Internet has led to the creation of testing tools for standard Internet languages.

MAIN ISSUES IN THE AUTOMATIC EVALUATION OF INTERFACES ON THE INTERNET

HTML Syntax Checking

Syntax-checking programs have been developed for programming languages and markup languages. Syntax checkers for HTML and other Web standards analyze the quality of Web pages from the perspective of software engineering. However, some systems also consider aspects of human-computer interaction.

One of the first tools for the evaluation of HTML pages was Weblint (Bowers, 1996). It is a typical system for syntax checking and operates on the following levels.

- Syntax (Are all open tags closed? Are language elements used syntactically correct?)
- HTML use (Is the sequence of the headings consistent?)
- Structure of a site (Are there links that lead one hierarchy level up?)
- Portability (Can all expressions be interpreted correctly by all browsers?)
- Stylistic problems (Is alternative text provided for graphics? Do words like *here* appear in link text?)

The examples also illustrate how syntax-checking programs are related to human-computer interaction. Some rules cover only the syntactical correctness. Others address the user experience for a page. For example, missing alternative text for images poses no syntax problem, but it may annoy users of slow-loading pages. In their generality, these simple rules do not apply for each context. For instance, a link upward may not be useful for nonhierarchical sites.

A more comprehensive system than Weblint is available from the National Institute of Standards and Technology (NIST, <http://zing.nsl.nist.gov/WebTools/>). Its system WebSAT is part of the Suite Web Metrics. WebSAT is based on guidelines from the IEEE and checks whether tags for visually impaired users are present. It also tests whether forms are used correctly and whether the relation between links and text promises good readability.

An overview of systems for code checking is provided by Brajnik (2000). Such systems can obviously lead only to a limited definition of quality. However, they will certainly be part of more complex systems in the future.

Link Analysis

One of the main challenges for search engines is the heterogeneous quality of Internet pages concerning both content quality and interface design (Henzinger, Motwani, & Silverstein, 2002). In Web information

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/automatic-evaluation-interfaces-internet/13100

Related Content

Family Businesses and Their Transition to Industry 4.0: Human Resource Perspectives From Bangladesh
Zarjina Tarana Khaliland Farzana Nahid (2022). *International Journal of Technology and Human Interaction* (pp. 1-21).
www.irma-international.org/article/family-businesses-and-their-transition-to-industry-40/306229

Evaluating Psychological Aircraft Accident Reports for Differences in the Investigation of Human Factors
Marco Michael Nitzschnerand Michael Stein (2019). *Human Performance Technology: Concepts, Methodologies, Tools, and Applications* (pp. 1302-1318).
www.irma-international.org/chapter/evaluating-psychological-aircraft-accident-reports-for-differences-in-the-investigation-of-human-factors/226618

Regaining Our Place as a Species among Species
(2014). *Technology versus Ecology: Human Superiority and the Ongoing Conflict with Nature* (pp. 254-268).
www.irma-international.org/chapter/regaining-our-place-as-a-species-among-species/78809

Computer-Mediated Communication in Virtual Learning Communities
Lisa Linkand Daniela Wagner (2009). *Human Computer Interaction: Concepts, Methodologies, Tools, and Applications* (pp. 316-322).
www.irma-international.org/chapter/computer-mediated-communication-virtual-learning/22258

How It Started: Mobile Internet Devices of the Previous Millennium
Evan Koblentz (2009). *International Journal of Mobile Human Computer Interaction* (pp. 1-3).
www.irma-international.org/article/started-mobile-internet-devices-previous/37457