

# Automated Assessment of the Quality of Wikis' Content: a User Study

Eti Yaari

## Abstract

One of the major challenges in evaluating the quality of information available on open content systems (e.g., Wikipedia) is the inability of users to always apply traditional methods for evaluation, such as checking the identity of the author and his or her motives in order to estimate the degree of his or her authority. Other well-known methods have similarly been found to be ineffective. Comparison/corroboration techniques are not widely adopted by users, mainly because they are generally considered time consuming and demanding tasks, and as such do not correspond to the desire for instant gratification. In addition, the dynamic nature of the content makes it difficult, if not impossible, to implement peer review processes in order to determine the level of its accuracy.

These concerns emphasize the need to refine existing methods and to develop new ones that are appropriate for the unique characteristics of environments in which the content is created collaboratively.

One of the most popular present approaches involves the use of automated assessment tools. This approach incorporates a variety of means that seek to automatically evaluate the quality of the content, i.e. without receiving user input, and to present the results of this action to information consumers.

The present research aims to examine the automated approach from the user's perspective. Three secondary aims were derived from this overarching goal: First, to check how information consumers evaluate the quality of content in an environment – like Wikipedia – in which the writing process is collaborative; Second, to check the possibility of applying the users' insights in order to construct a user-centered automated tool for evaluating information quality; And third, to examine the users' perspectives on the issues involved in affixing an automated tool, based on this evaluation method, to this sort of environments.

The research was conducted in two sequential stages, using an exploratory strategy. In the first phase, which was a qualitative paradigm, 64 users were requested

to complete two evaluation tasks, to respond to a semi-structured interview and to fill out two background questionnaires. In the first evaluation task the participants were asked to assess the quality of five articles from the Hebrew Wikipedia and to indicate which two articles of the five were of the lowest and the highest quality, providing reasons for these choices. Each participant, it should be mentioned, conducted this task twice: once while looking at the article's history page. This aimed to include within the analysis aspects related both to the article's current content and to its development.

In the second evaluation task the participants received a list of 18 quantitative attributes identified in the literature as able to include quality, such as the entry's length, its currency and the degree of variety of its contributors. The participants were asked to select three attributes, which, in their opinion, were the most effective in evaluating the quality of the collaborative content, and to specify the significance they ascribe to these characteristics. During the interview, the participants were asked about their perception regarding the Wikipedia environment and about their attitudes toward the integration of an automated tool for assessing information quality in this environment.

In the second phase of the research, 2224 articles from the Hebrew Wikipedia were assessed based upon the insights learned from the participants in the evaluation tasks in order to gauge effectiveness of this evaluation method.

The research findings show that using the quantitative attributes of the information items presents a valid method of evaluation from both the human and practical perspectives. Examining the participants' means of dealing with the task of evaluating content quality revealed that many attributes that assisted them in assessing the Wikipedia articles are of a quantitative nature. This enabled the definition of these attributes as automatically measurable, and their integration in an automated tool for information quality assessment. In addition, inspecting these characteristics in the Wikipedia environment revealed their ability to indicate the quality of the articles, that is, to differentiate low-quality articles from high ones.

Furthermore, in the framework of this research the participants were shown an automated evaluation tool that enable them to choose the quantitative attributes they believe indicate the quality of the articles and in accordance with their preference to see the raw data on every characteristic and/or the weighted score for the article.

The findings show that participants' opinion concerning the integration of the tool within the Wikipedia environment is chiefly positive, and stems from their belief in its ability to help them in consolidating their view regarding the quality of the piece of information.

The findings also reveal that the participants' degree of willingness to take part in the assessment process was not only influenced by their cognitive abilities; motivation too plays an essential role.

These issues discussed in depth within the discussion, and several ideas were suggested for follow-up research in the field. In addition, the theoretical contribution of the research was presented together with practical recommendations for planning automated systems for assessing the quality of the online content.

One of the prominent conclusions emerging from the research is that we should look positively on merging automated evaluation tools as an integral part of information environments found in the virtual setting. Yet before these tools are integrated, it must be ascertained that the method at their foundation, the way the tools are used, and the way their products are presented are indeed comprehensible to the entire population that will utilize its serviced in the future.

System No.

**1149962**