

## **Abstract**

Master's attestation work on:

“Multi-criteria evaluation tools for websites and fake websites detection”

Rublyov Dmytro

### **Relevance**

We live in a time of rapid development of information technology. They penetrate into all spheres of life of modern man. Increasing the value of information, rethinking its role in human life. With the rapid development of the Internet is becoming more informative among the “noise” to find the necessary information to us. Now, to create a website is not necessary to know HTML, and programming languages, because there are many software tools for creating Web pages, where all the work on the development page is reduced to dragging graphic elements and type. In many cases, the quality of sites created so better keep quiet - the absence of understanding the fundamentals of building web sites, web standards, and sometimes even the ordinary rules of grammar, what kind of quality can we speak? The existence of free webhosts can quickly fill a virtual web site are informative and useful of which tends to zero, and which complicate the search for useful information in global information network.

The existence of “garbage” sites now built a business - so-called “black” search engine optimization and “moneymaking”. Both events use that number in time develops into quality – creating many doorway pages and satellites can be for some time to increase attendance to some Internet resource. When a “garbage” sites are created, the question of their contents is often not the first place, and the site content is created by borrowing texts from other sites, rewriting with replacement of the words verb, and sometimes even generating source of the fragments of phrases,

selected at random. In all cases, Internet users suffer from the existences of such sites and in the first - the authors of borrowed texts.

But borrowing texts sometimes occurs not only among black search engine optimization, but also among large sites of major companies. There are cases borrowing texts very large corporations, which led to the discovery of cases and punish the perpetrators.

Unusual stands a category of sites that fully duplicate the appearance of any particular site. The main part of these - sites in the fraud. Having a replica authorization page in a popular online banking system, and sending spam message, asking to visit this page and log in there under any pretext. If this letter gets to the inattentive user bank, which copied the authorization page - it (user) himself actually give thieves their passwords.

In such a situation, the question of assessing the quality of web sites and identify fraudulent Web sites appears quite serious.

## **Objective**

The aim of this work is to study existing Web sites assessment tools by highlighting the basic criteria that together give an idea of quality sites.

## **Results**

The paper describes the main criteria to determine quality of the site, an overview of tools to evaluate these criteria and describes ways to improve their own websites.

Among the criteria that define websites quality was selected Top 4:

1. Content;
2. Link popularity;
3. Technical realization;

#### **4. Web standards matching;**

Among existing tools for assessing the quality of the website is considered as online services and applications. Using existing programs with reasonable accuracy can judge the quality of websites.

### **Scientific novelty**

For the first time were considered approaches:

1. to evaluate the information for search engines value;
2. to the method of determining the source of original texts;
3. for minimizing negative consequences from borrowing other people's lyrics sites.

### **Practical achievements**

Currently the system has only the basic functionality on which can developed into more powerful tools to identify borrowed information on the Internet.

### **Conclusion**

Due to impossibility of quality evaluation as a criterion that at present cannot be formulated and used to evaluate sites, was considered the approach to the evaluation of Internet sites by the criteria used by search engines for information ranking.

Results of these experiments, it was determined that the best way for the creation and further development of websites is to use CMS with testing situations access to pages with the incorrect address canonization URL, customer optimization and setting up a web server for caching and compression of downloaded information.

This approach allows us to keep a clear site structure, reduce time to open the page the user and reduce load on the equipment and bandwidth.

The work contains 98 pages, 21 figures, 1 table and 26 sources.

**Keywords:** CMS, W3C, PRIMACY OF INFORMATION, SEARCH SPAM, INTERNET SITES QUALITY, USER-SIDE OPTIMIZATION.