6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, DSAI 2015

Exploring web attributes related to image accessibility and their impact on search engine indexing

Reinaldo Ferraz*

Ceweb.br – Web Technologies Study Center/ NIC.br - Brazilian Network Information Center-
Av. das Nações Unidas, 11541, 7o andar São Paulo – São Paulo – 04578-000 - Brazil

Abstract

The purpose of this study is to analyze how search engines index web content inserted in image attributes for alternative/complementary texts that favor web page accessibility. The study discussed the importance of optimizing websites to improve search tool indexing and explored how these engines index image attributes. We conducted empirical observations of tests carried out in a controlled environment, representing a typical image publishing website. The basis of the experiment lay in observing how content inserted in ALT attributes in elements of web page images was indexed. Data were collected first by observing the indexing of a page without any declared attributes. Then, we analyzed a page with all the properly declared attributes. The experiments were conducted using popular web search engines, such as Google, Yahoo, Bing and DuckDuckGo. We also used Search Engine Optimization (SEO) online verification tools for analyzing criteria relevant to search engine indexing. The hypothesis for this paper is that the ALT attribute is relevant to SEO, whereas the TITLE attribute is not. At the end of the study, we present details on how some search engines indexed content specific to alternative/complementary for images.

Key words: Accessibility, web accessibility, images, alt, title, SEO

* Corresponding author. Tel.: +55-11-5509-3532.
E-mail address: reinaldo@nic.br

1. Introduction

The web has evolved from simple HTML pages full of text and hyperlinks to a rich repository of connected information. Such information goes far beyond online texts. Pictures and multimedia archives are common on the web of today and are increasingly used as technology evolves. Simply varying text-only to rich media sites is no longer enough for today's web. The purpose of a web page is not merely to present information, but to enable participation in the learning community [1].

In this context, web accessibility is an extremely important theme. However, it tends to be placed in the background when developers, companies and institutions have tight deadlines or budgets. They frequently consider the accessibility of their page to be unimportant, unaware of the extent to which improving page coding can interfere with important issues such as performance and indexing in the main search engines.

One item that is of utmost important to accessibility is the proper description of website images. Such descriptions allow individuals who cannot see the image to receive alternative and/or complementary information about the figure or picture. As this type of activity requires the description and manual validation of such resources, texts tend to be set aside or misused within the code.

In order to demonstrate that the importance of appropriate image description transcends issues related to web accessibility, the main purpose of this paper is to present how ALT and TITLE attributes are indexed by search engines and how they can be used to improve accessibility to both published images and pages indexed by search engines.

2. Hypothesis and Methodology

The hypothesis for this work is that the ALT attribute is relevant to SEO, because its use refers to the alternative text of the image, while the TITLE attribute should not be relevant because it has an advisory role, as will be investigated in this paper.

For the theoretical part of this article, a literature review was conducted and discussions on the concept of web accessibility were analyzed. The methodology used for the empirical part of the study consisted of conducting observations of web content publishing, focusing on search engine content indexing and automatic SEO verification. The publishing process was conducted in two steps:

- Publishing images without the declared ALT attributes;
- Publishing images with the declared ALT attributes.

The results were gathered after indexing the images found in the two prior publishing steps, using search engines and automatic SEO checker tools:

- Collecting and analyzing results of indexing without declared ALT attributes;
- Collecting and analyzing results of indexing with declared ALT and TITLE attributes.

These steps were necessary for analyzing how search engines index content with properly declared ALT and TITLE attributes.

The test environment used for this experiment was Reinaldo Ferraz' personal blog (http://www.reinaldoferraz.com.br/) and an article on the World Wide Web Consortium (W3C) Brazil website (http://www.w3c.br/Noticias/AgenciaGeniaParticipaDeWorkshopNoW3cBrasil). The first is a public Linux-based web environment, with WordPress and the robots.txt file configured so that search engines can index the website's content. The second environment is a FosWiki under a Free BSD platform.

All the tools used in the test are available online and require no installation. We used Google (http://www.google.com/) as the main search engine for the test, as it is the most accessed website in the world, according to the Alexa ranking of March 2015 [2]. Other search engines used in this experiment were: Yahoo (http://www.yahoo.com/), Bing (http://www.bing.com/) and DuckDuckGo (http://www.duckduckgo.com/).