Web development awareness about W3C standards: Croatia survey

Tomislav Jakopec, Eleonora Đekić, Tanja Slijepčević

Department of Information Sciences
Faculty of Philosophy

L. Jägera 9, 31000 Osijek, Croatia

`{tjakopec, edjekic, tslijepcevic}@ffos.hr`

Abstract. *W3C (World Wide Web Consortium) standards serve as guidelines and rules for web developers, so that web pages would be of the highest possible quality and made according to the rules. Based on surveys, it has been noted that W3C standards are seldom used when developing web pages. That's why research was decided to be conducted into how much familiarity web developers actually have with W3C standards, and how often they are used. The research was conducted in April 2011, when randomly chosen web developers from across Croatia were taken as a representative sample. The goal of the research is to assert, not just the approximate number of developers complying with W3C standards, but also their reasons for not doing so.*

Keywords. Web development, Web developers, W3C

1 Introduction

The World Wide Web Consortium (W3C) is the main international organization which determines World Wide Web standards. Its importance for web developers lies in the rules and guidelines it provides them. This paper focuses on W3C standards and their use in practice. By conducting a survey among web developers in Croatia, the goal was to find out their opinions and attitudes toward W3C standards, as well as the importance those standards have in regards to the validity of websites. In this research the working hypotheses were:

H1: web developers are not well-acquainted with W3C standards,

H2: web developers don't comply to W3C standards because websites do not have to comply with W3C standards to look good.

2 W3C standards

The term web standards refers to a set of suggestions and guidelines for the proper creation of websites. Their goal is securing unique and unobstructed access to web technologies for everyone. W3C is an international organization that handles the standardization of technologies used for creating and developing the Web. It was founded in 1994. in association with the Massachusetts Institute of Technology and the European Organization for Nuclear Research (CERN), under the initiative of Tim Bernes Lee. [1]

The W3C fulfills its mission by creating web standards and recommendations, but it is also involved in software development, education, and also serves as a transparent forum for Web-related discussions. Individual HTML document should be assigned to the corresponding DOCTYPE.

W3C standards provide for the following document types [2]:

- HTML 5*
- XHTML 1.1
- XHTML 1.0 Strict
- XHTML 1.0 Transitional

- XHTML 1.0 Frameset
- HTML 4.01 Strict
- HTML 4.01 Transitional
- HTML 4.01 Frameset
- other (MathML, Compound, Optional and Historical)

*Not standard, working draft

3 Previous researches

The realistic usage of W3C standards in the development of web spaces is not completely satisfactory. Karl Dubost in his paper from 2002 states [3]:

„Most of the Web sites on the Web are not valid. We may assume that this is the case for 99% of the Web pages, but there are no statistics to support this. It would be interesting to run a survey to prove that this case is indeed true.“

Brian Wilson in key findings [4] of Opera's project MAMA in 2008 states that only 4.13% of 3,509,180 URLs from Open Directory Project (http://www.dmoz.org) passed validation. Same Author states

„Even though this ratio shows great improvement over the results of previous validation studies (Parnas 2001[5] and Saarsoo 2006[6]), this is a very worrying figure, which shows that there is a lot of Web standards education still to be done to increase these levels.“

Group of author's in 2008 [7] and 2010 [8] revealed that about 20% of the most popular sites don't even use a DOCTYPE declaration which is element of W3C standards and level of compliance to the HTML standard of web pages is very poor.

The most recent research in Croatia was conducted 2010 [9], results revealed that only 14% of web pages, in the sample, inside the Croatian national top-level domain were valid according to W3C standards.

Reason of such poor compliance lay in the web developers (web designers, web programmers, etc.), they create valid/not valid web pages. A List Apart Magazine is conducting a survey among web developers from 2007 [10]. Survey results tell us who are the web developers but unfortunately not their awareness about W3C standards.

Importance of using W3C standard and benefit for both users and web designers is emphasized in paper from Constantine and others[11]. Web designers are able to design a website with minimal worries regarding different non-compliant browsers. That way, the site becomes accessible from a wider variety of devices by following simple W3C guidelines and recommendations.

Motivated by these findings and facts, the goal of this paper is to find out the reasons for such a small usage of W3C standards with Croatian web developers.

4 The course of the survey

4.1 Sample

Potential participants in the survey were found in the following ways:

1. By searching the Internet for companies (Companies) that offer web making services. With this criteria, 160 unique email addresses were found, which were then sent an invitation for participating in the survey.

2. By finding the contact email addresses of the 50 most popular networking spaces (Top 50), according to the Alexa [12] evaluation system. With this criteria, 35 unique email addresses were found, which were then sent an invitation for participating in the survey.

3. By developing a custom website analyzer (Crawler), it became possible to sift through all email addresses that are publicly available as a single string of characters. By collecting all email addresses from the first two levels, restricted to the .hr domain, 45.563 email addresses were obtained. Seeing how not all of those email addresses were directly connected to website creators, a filtering with the following conditions was conducted:

* only email addresses from the .hr domain
* the string of characters on either side of the @ sign was equally long (without the .hr)
* the email address contains one of the following words: web, izrad, kontakt, admin

8.267 unique email addresses were obtained, of which 2000 were randomly picked and sent an invitation for participating in the survey.

The total number of participants invited was 2.169. The survey was technically realized with the help of the LimeSurvey3 [13] system, which was hosted on the following address: http://web.ffos.hr/anketaW3C/.

The survey questionnaire also familiarized the participants with the goals of the survey and the results of previous surveys. Since the Croatian

language contains a wide range of names for jobs that are involved in the creation of webpages, all derivatives of said jobs were listed: webmaster, web developer, web designer, web programmer, web architect, web application developer, PHP developer, HTML programmer and any other derivatives found in literature or practice.

4.2 Survey

The survey consisted of 10 questions divided into 3 groups:

The first group contained general questions and its goal was to describe the examinees according to their gender, age, workplace, professional qualification and interest in web development.

Gender
Years
Place of residence
Qualifications (regardless of the current job)

The second group refered to web development. The goal of this group of questions was to find out how the examinees learned to develop web pages and which tools they used for their creation.

How did you learn create web pages?
What integrated development environment you use when developing web pages?

The third group of questions examined the understanding and application of W3C standards when developing web pages.

Are you familiar with W3C standards?
Whether your pages are compliant to W3C standards?
Which document type you use when creating web pages?

The survey was conducted from April 1st to April 15th 2011.

4.3 Survey results

155 people participated in the survey. Table 1. shows the distribution of survey participants according to the groups the email addresses were gathered by.

Table 1.

Group	Send mail	Total start	Total end
Crawler	1974	174	99
Companies	160	62	51
Top 50	35	7	5
Total	*2169*	*243*	*155*

Small response rate of 7% shows web developers inertness towards the polls. The results of the first group of questions (Figure 1.) show that the majority of participants were male (89,0%). Unfortunately, the total number of female participants was 17, therefore no statistically significant results can be gathered based on that. Half of the participants were between ages 25 and 35 (51,6%), while the others were equally spread among the other age groups (Figure 2.). In comparison to previous results [10] result are similar for gender and ages. Results differ in part: Qualifications. A List Apart result show that the most common group (56,3 %) have College diploma, associate's, bachelor's, or equivalent degree while we have the most common group (42,6 %) with high school diploma.

The goal was for the survey to cover the entirety of Croatia, which was successful, judging from what participants noted as their place of residence (Figure 3.). The participants came from 45 different towns, which, compared with the official number of 123 towns, makes for a 37% representation. Most of the participants have high school or higher education as there qualification (Figure 4.).
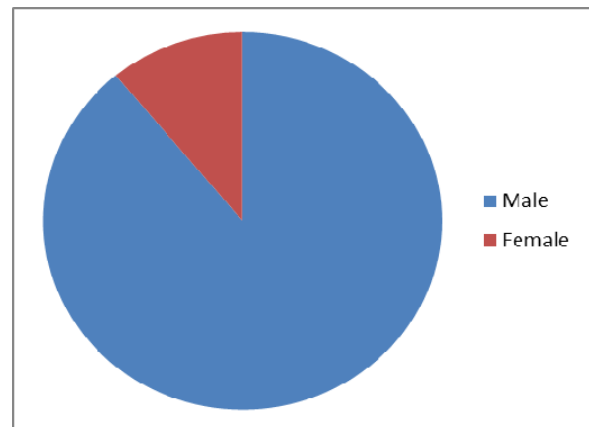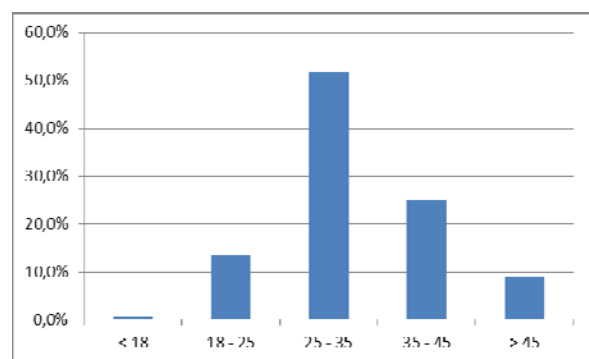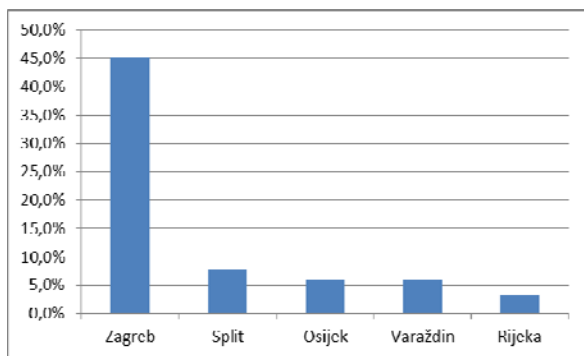

Figure 1. Gender

Figure 2. Age

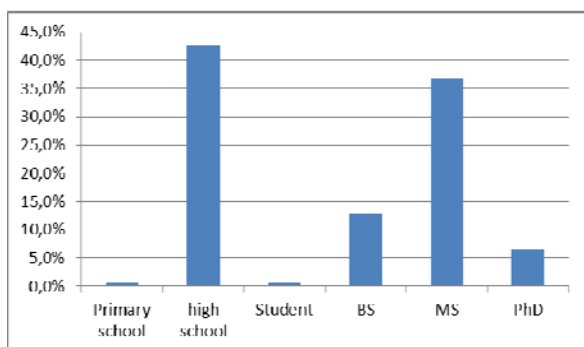Figure 3. The most common Place of residence

Figure 4. Qualifications

The results of the second group of questions show that the participants mostly learned how to create websites in an informal way (Figure 5.). For their work, they choose from a large pool of development tools (29), the first five of which we have shown the distribution of on figure 6. Dreamweaver domination coincide with BuiltWith [14] results. According to BuiltWith 4,228,144 websites using Adobe Dreamweaver while 1,206,026 websites using Microsoft Frontpage. The participants usually use one (63,2%) or two (24,5%) development tools for their work (Figure 7.).
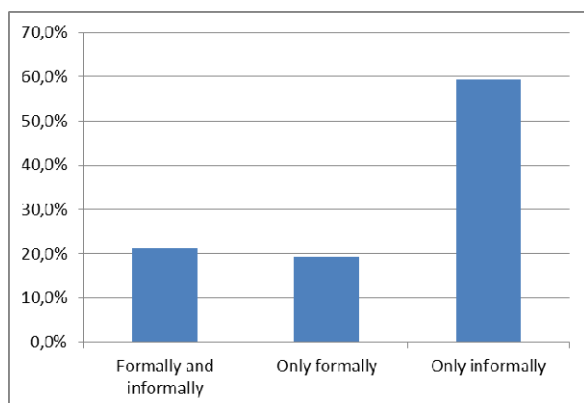
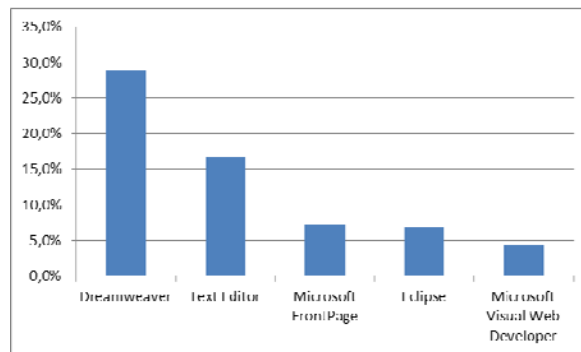Figure 5. Method's of learning web pages creation

Figure 6. Developing tools

The results of the third group of questions show that 78,1% of survey participants are familiar with W3C standards . For those of them that were not, the survey ended at this stage. Participants who answered that they were familiar with W3C standards check or partially check the validity of their websites, according to the distribution on diagram 7. For those of them that did not, the survey ended. Regarding the type of documents the participants use when creating websites, they answered mostly with Transitional or Strict. As a rule, they use one (76,1%) or two (22,1%) document types (Figure 8.) The literature search on the topic of Web developer's attitude toward W3C standards did not provide data directly comparable with data obtained in this study. However, there are articles [15] about how knowledge of W3Cstandards is one of the competencies required by every web developer.
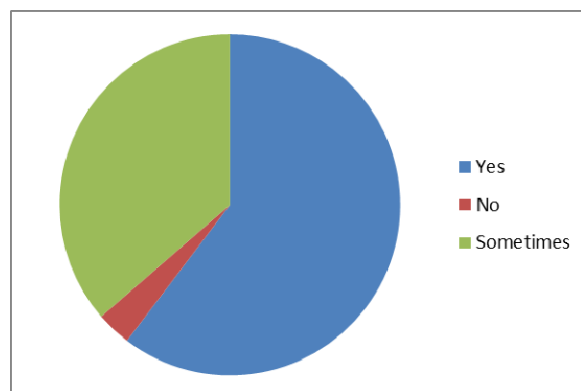
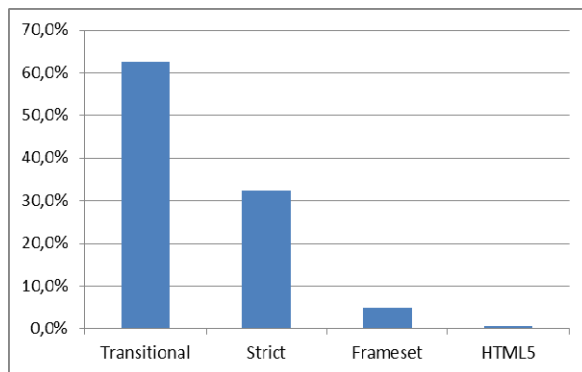Figure 7. Compliance to W3C standards

Figure 8. Used document type

In response to the last question, what they thought about W3C standards, the participants said that they were not familiar enough with them. Still, one point that everyone agreed upon was the problem with web browsers. Namely, web browsers are not synchronized, therefore web pages appear differently on different browsers. "W3C standards are an excellent idea, but the execution is sloppy; excessive fragmentation and Microsoft's disregard of any standards have ensured that these standards remain just guidelines, instead of concrete rules. The development of pages compatible with these standards would be practically impossible without an official validator." Also, the majority believes that obeying these standards would be quite demanding because of their restrictive nature, but they agree that W3C standards are necessary to make web pages of the best possible quality.

5 Conclusion

The survey successfully covered Croatia geographically. The results of the first and second group of questions coincided with the author's assumptions, but what was surprising was the fact that 16,8% of survey participants use some form of a common text editor. In the third group of questions, there was an incongruity between the web developers' statements and the results of previous surveys. According to the survey, 21,9% of participants are not familiar with W3C standards, while the others replied that in 96,7% of cases, they always or usually adjust their web pages according to W3C standards. This percentage isn't consistent with previous results, which have shown that only 12% of homepages in the .hr domain are valid. These results lead us to reject hypothesis H1. While Croatian web developers have the

willingness to use W3C standards and they make statements of doing so, actual "field results" show the opposite. The hypothesis H2 was partially confirmed by comments of participants. Further Webometrics, web mining studies and surveys are necessary to understand why such an inconsistency between results exists.

References

[1] **A Little History of the World Wide Web**, available at `http://www.w3.org/History.html`, Accessed: 20th June 2011.

[2] **W3C QA - Recommended list of Doctype declarations you can use in your Web document**, available at `http://www.w3.org/QA/2002/04/valid-dtd-list.html`, Accessed: 20th June 2011.

[3] Dubost K. W3C QA - **How to achieve Web standards and quality on your Web site?** available at `http://www.w3.org/QA/2002/04/Web-Quality`, Accessed: 29th June 2011.

[4] Wilson B. **MAMA: Key findings** - Dev.Opera, available at `http://dev.opera.com/articles/view/mama-key-findings/`, Accessed: 29th June 2011.

[5] Parnas, D., **How to cope with incorrect HTML** (Master thesis), available at `https://bora.uib.no/bitstream/1956/1731/1/Masteroppgave-parnas.pdf`, Accessed: 30th June 2011.

[6] Saarso, R.,**Veebilehtede kodeerimispraktikad** (Bakalaureusetöö), available at `http://triin.net/2006/05/10/veebipraktikad.pdf`, Accessed: 30th June 2011.

[7] Beatty P, Dick S, Miller J. **A Large-Scale Survey and Analysis of Conformance to W3C Standards: Is HTML in a Race to the Bottom?** IEEE Computer Society 2008; 76-80.

[8] Ofuonye E, Beatty P, Dick S, Miller J. **Prevalence and classification of web page defects**. Online Information Review 2010; 34(1): 160-174.

[9] Jakopec T., Papić A. and Selthofer J., **Inside Croatian National Top-Level Domain: Analysis of Technical Quality according to W3C Standards**, Proceedings of the ITI 2011 33nd International Conference on Information Technology Interfaces 2011. 471-476.

[10] A List Apart, **Findings from survey for people who make websites**, available at http://aneventapart.com/alasurvey2010/, Accessed: 29th June 2011.

[11] Constantine A. Murenin, Mohammad H. N. Tabrizi, **Development of Usable and Accessible Web-Portals Using W3C Standards**, itcc, vol. 2, pp.829-831, International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II, 2005

[12] **Alexa the Web Information Company**, available at `http://www.alexa.com/`, Accessed: 04th April 2011.

[13] **LimeSurvey** - the free & open source survey software tool, available at `http://www.limesurvey.org/`, Accessed: 04th April 2011.

[14] **BuiltWith Web Technology Usage Statistics**, available at `http://trends.builtwith.com/framework /Adobe-Dreamweaver` and `http://trends.builtwith.com/framework /Microsoft-Frontpage`, Accessed: 04th April 2011.

[15] Wayde C. **10 questions to ask when looking for a professional web designer or developer**, available at `http://newism.com.au/blog/post/54/10- questions-to-ask-your-web-developer/`, Accessed: 14th April 2011.