

HTML5 – A NEW HOPE AND A DREAM

Tăbușcă, Alexandru¹

ABSTRACT

The present paper aims to discuss the introduction, use and future trends for using HTML5. After going through an entire collection of different flavors of the good old HTML – DHTML, XHTML – and adding a lot of other technologies like JAVASCRIPT, or CSS, the developers of the today IT environment must learn the new and ultimate HTML5 ☺. This technology aims to provide not only a more complex and robust HTML syntax (with more strict XML implications) but also to be able to replace JAVA or FLASH in contexts regarding graphic approaches.

Keywords: html5, css, javascript, dhtml, xhtmls, svg, xml

1. INTRODUCTION

The world today is almost unthinkable if we try to set apart the computers from our lives. Everything today is more or less controlled by the use of computers. And we should also mind the next step – we don't only rely on computers but we take for granted the use of the global network, the “internet” that is now a noun, a concept used even by three year old children. I should also add that computers and internet are even legalized into our lives now – Several years ago Finland pioneered the implementation of a law regarding the guaranteed access to the internet of its citizens [1].

Throughout our “internet” history, even if actually humanity does not really have a long time experience regarding this issue from the historical point of view, the most important and influential technology is arguably the HTML language. Together with the classic HTML, grown up to 4.01 version, we must also mention two of its quite close “cousins”: the XHTML and DHTML.

HTML (Hyper Text Markup Language) is now clearly the most widely used, taught and accepted programming language for building websites. It provides the main structure of a website, it builds all the basic elements that a website should involve: tables, divisions, images, windows messages and text.

During its development history the HTML programming language (I call it this way even if among the programming community there is an issue regarding the “programming” particle being added to HTML description) has gone through several landmarks:

- HTML 2.0 – it actually was the first standard supported by the early browsers

¹ Lecturer, PhD, School of Computer Science for Business Management – Romanian-American University; e-mail: alextabusca@rau.ro

- HTML 3.0 – it is a standard that did not really see wide adoption due to the fact that it has been developed during a period of intense browser innovation and this fact did not help in establishing a clear and somehow rigid standard
- HTML 3.2 – this variant aimed to be the “real-world” 3.0 specification. It included a lot of browser-invented elements and also removed the 3.0 elements that did not see actual implementation within the browser of those days
- HTML 4.0 Transitional – this is the first iteration of the version we actually see today almost everywhere on the world wide web. This variant comprises frames, scripting elements, style sheet and presentational elements
- HTML 4.0 Strict – this variant of HTML removes almost every presentational element from the specifications, leaving this part to the CSS
- HTML 4.0 Frameset – this minor flavor of HTML defines the specifications for supporting the frameset, frames and inline frames
- HTML 4.01 – this variant which is the “de facto” standard of today is only a minor bug fix of the 4.0 version actually

At one moment, the environment of the web programmers’ community felt that HTML needed the rigor of the XML and as a result, at the end of the last millennia, in the year 2000, on January 26th a new programming language has seen the light: XHTML 1.0.

XHTML (eXtensive Hyper Text Markup Language) is almost the same as HTML, with the difference being existent in regard to a cleaner syntax. XHTML uses the same tags as HTML, and as a result programmers who know HTML will know actually a sort of “unclean” XHTML. There are new rules that have to be applied inside XHTML, such as tags that always need to be ended or tags that need to be "nested" properly.

- XHTML 1.1 – it is a more strict specification defined by the W3C² consortium and it represents another markup language with powerful functionality as regarding to structure but with reliance on the CSS for the presentation part. This variant is focused on the modularization of XHTML in order to make it easier to define and divide the features. We can find only several changes between versions 1.0 and 1.1 of XHTML
 - - the “lang” attribute has been replaced with “xml:lang”
 - - the “name” attribute has been replaced with “id” for the “a” and “map” tags
- XHTML Basic – it is a version of XHTML 1.1 aimed for the mobile applications development.
- XHTML 2 – this version of the language is quite a radical change. It aims to build a logical language without regard for backwards compatibility. The specification is still under development

DHTML (Dynamic Hyper Text Markup Language) is not exactly a version of the HTML programming language - it represents the way of using HTML, JavaScript and CSS together in order to create dynamic things, such as navigation menus.

² W3C – World Wide Web Consortium is an international community that develops open standards to ensure the long-term growth of the Web; www.w3c.org

DHTML represents a collation of web technologies used to move the programming universe beyond the static presentation of information, in order to be able to create more interactive websites. Technologies that play a key role in DHTML are HTML, JavaScript, Cascading Style Sheets (CSS) and the Document Object Model (DOM). CSS is actually used to control the presentation part of the structured content that is defined by the use of HTML. The DOM is the client-side programmable interface used in order to show the state of the HTML document that is published and its associated CSS presentation. The dynamic part of DHTML comes into the scene the moment that the JavaScript technology is being used to manipulate the state of the DOM in order to cause the appearance or behaviour of the published web page to change.

As an heir to all this HTML and XML mixtures of web programming languages, the HTML5 has a lot of things to cover. It represents a new hope for the standardization fans and a dream for most web developers. This new paradigm is trying to cover both static and dynamic approaches of the HTML, comprises several former independent languages such as SVG or MathML and deals with graphics in both JavaScript and SVG style – all these technologies still relying on CSS (and CSS 3 in the near future) for the presentation part.

2. FACTS

Nowadays HTML5 appearances, or at least articles about it, can be seen all over the place. Not only the enthusiast programming community but also the large and renowned companies of the web universe are starting to implement the new markup language. But, even with the entire hype made around this “revolutionary” HTML one must maintain an objective look and ponder if this is really the case – do we see a “revolution” or just an “evolution”? Can we see the new HTML5 as a universal solution for both classic and mobile web applications?

Going a little bit in depth in regard of the HTML characteristics and specifications, most web programming specialists agree on a list of five issues that one must take into consideration when offered a choice: design a classical HTML website or going with HTML5. I can not exactly rank these elements as I consider all of them to be very important, of course with different degrees of importance for different cases scenarios.

I. Actual implementation and interpretation

HTML5 is not exactly a native technology for any device. A real native technology is actually compiled separately for each device and utilizes its capabilities at maximum level. The new markup language does not differ in any way from the old variants, the HTML5 design must actually go through a browser that publishes the objects, synchronizes elements with the help of JavaScript and arrange them according to the CSS code; all these actions might as well be transparent for the users but in real life they actually incur a penalty regarding the raw performance of the software.

II. Client technology

It is obvious that HTML5 is a client technology – even though it would be an excellent opportunity for the programmers if they did not have to manage the requests and

synchronization issue, unfortunately, this does not happen (at least yet) in the HTML5. The new HTML flavor is easy to learn, even if it became a little bit more complex and complicated than its predecessors, but the programmer is aware that in fact it is only a front end (more evolved than before) and in the end everything has to be submitted to the server in a way or another. This fact brings into life the need for the divide of the application algorithm between the client and server, from the point of view of its logic and operations. This, of course, is only a normal case for any 3-tier application but for the today development techniques it seems rather outdated.

All the logical divisions between the server and the client and the communication/synchronization issues, as regarding the application, must be deal with manually by the web developer in the case of HTML5. There are a lot of other development platforms on the market today that take care of this problem in an automatic manner; this thing brings down the need for code writing and adds an easier way to address an entire list of issues.

III. State of the specifications

Even of HTML5 is already several year old is still pretty much under development. There are still a lot of problems with the implementation of the programming language tags in different browsers. Most of the time, the developer has to not only implement HTML5 but in the same time implement different workarounds for the cases when the main code is rendered differently, or not at all, on some different relevant browsers. In fact, the HTML5 developers can be considered for the moment as beta-testers ☺. This issue brings a lot of problems especially when taking into consideration the actual costs in time and manpower needed to solve an issue that should not have been there in the case of an established and tested set of specifications.

IV. Security

All markup languages are by design and way of use much more exposed from the point of view of the security issue. Even if the new web applications designed on the framework of the HTML5 can use the HTTPS protocol in order to try to secure the communications, there still are some problems common to all HTML designs: macro command, malware cases, fishing attacks [2], code injection, cross-site scripting etc.

In 2010 the famous Steve Jobs³ stated the HTML5 is the future and all Apple device will implement it, using it as a counter-weight for the unsupported Adobe Flash. The Apple switch on the HTML5 side made a bang and the specification, despite its under development status, gained wider and wider adoption; together with this also came an entire list of security issues that began to be noticed.

The ENISA⁴ chose to scrutiny in more depth these security issues and made a complete study on 13 HTML5 specifications, taken into account their public and official definition

³ Steve Jobs – founder and leader of Apple company, one of the largest and most innovative digital-age companies

⁴ ENISA – European Network and Information Security Agency; <http://www.enisa.europa.eu/>

given by the W3C. The result was quite important – no less than 51 security threats have been identified and documented.

Another institution with world-wide reputation and recognition, the security software company Sophos⁵, made an extensive analysis on the HTML5 and identified four issues

- The problem of cross-origin resource sharing
- The problem of click hijack
- The problem related to the geolocation and privacy issues (specific to HTML5)
- The WebSocket API

To remain absolutely true, all these issues are not drawing a bleak and gloomy image for HTML5, or at least not a worse image than of any other modern web development markup language. Actually, the HTML5 problems are continuously approached by the web community and solutions are implemented faster than in most other cases. The fact the HTML5 updates of implementation are done through the browsers and not by only one company, as for example in the case of Adobe Flash Player, also helps in speeding up the handling of different issues. Besides the browser vendors I must also mention that some of the largest digital-environment companies like Google, Microsoft, Facebook or PayPal are also actively supporting the HTML5 designs – a fact that helps even more the acceptance and dissemination of the new language.

V. Communications and synchronization

By design, all HTML flavors are built as asynchronous technologies. The synchronization part within HTML comes from the implementation and use of JavaScript. Unfortunately, JavaScript is not the most reliable technology and therefore it is not well suited for enterprise-grade applications. JavaScript can be quite slow sometimes and can be quite easily misled into feeding inaccurate or even plainly false data to an application. Even though virtually all the benchmarks focused on browser applications take into consideration for the speed-test only the speed of the JavaScript calculations, there are still cases when a JavaScript intensive web application might suffer from a visible slow-down in certain cases. Perhaps the widespread adoption of using the modern graphic cards capabilities for all types of calculations might bring these issue to an end but, for the time being this is just a desiderate. For example, the use of existing CUDA capabilities from NVIDIA graphic cards should bring a huge leap in calculation speed [3].

One of the most widely used case of HTML5 in real-life today is related to social-media. The web programming community considers that social-media is becoming a critical element in regard to user engagement and content distribution and that HTML5 can be the most convenient way to bring on the market richer applications, with more dynamic and more user-friendly application experiences. In fact, an important survey⁶ that was done in 2012, involving more than 2100 programmers focused on the development of mobile

⁵ SOPHOS – one of the oldest security companies of the IT environment, founded in 1985; it provides software and hardware security solutions

⁶ Survey published on June 20, 2012, by Appcelerator; <http://research.venturebeat.com/content19997>

applications, resulted in a figure of no less than 79% of the answers stating that the programmers were planning to use HTML5 for application development in the near future.

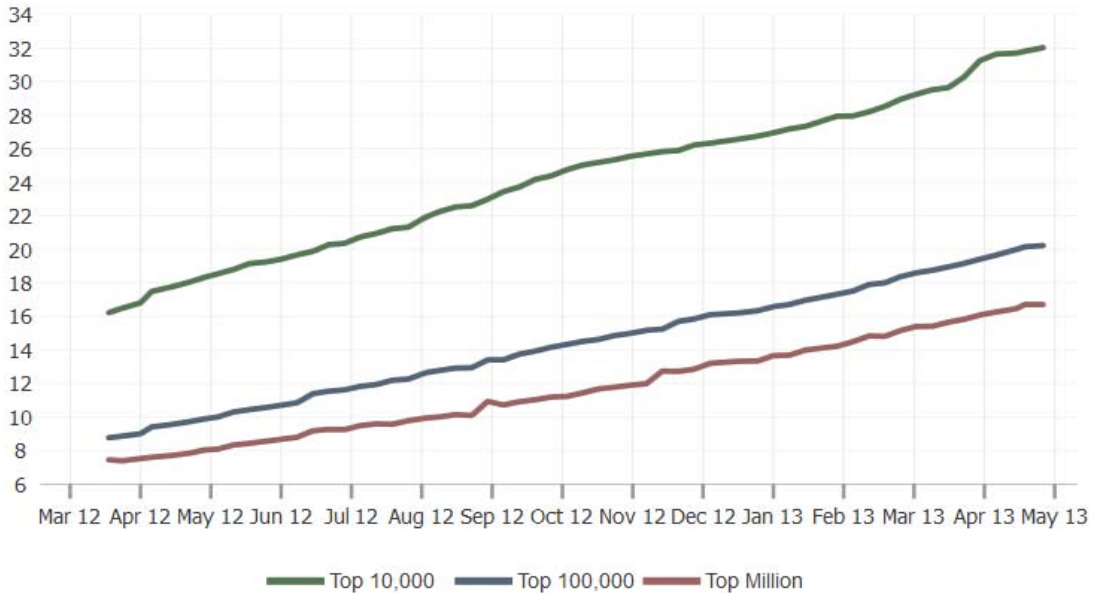


Figure 1. The usage trend for HTML5 elements on a specific set of websites⁷

A large automatic survey of 7604730 websites, conducted by the Australian company BuiltWith Trends, shows the obvious and constant positive trend of HTML5 use in all websites. The chart shows the HTML5 use trend for the top 10000, top 100000 and top 1 million websites, ranked by the number of visits.

The analysis is based on the actual existence of the HTML5 trademark tag: <DOCTYPE>. In fact there might be even more websites using actual HTML5 elements, such as form parts, graphic effects, local validations and so on, but they do not have the <DOCTYPE> tag available for the web robots to collect. These figures are very recent, the chart being based on the data compiled in the first part of May 2013.

The use of HTML web applications should also be increased if we look at the rate of support by the browsers. Each week brings new implementations of HTML5 elements to all the major browsers on the market today. From the statistics available for the last quarter of 2012, the market penetration of HTML5 capable browsers was excellent; in the United States and Canada around 75% of the users were deploying such browsers and in Europe no less than 83% of the users were in the same situation. These figures state a huge advancement when compared to the same analysis for the last quarter of 2011 – at that moment only 57% of browsers in the United States of America and Canada were HTML5 capable and only 62% were in the same situation in Europe⁸.

⁷ Source: <http://trends.builtwith.com/docinfo/HTML5-DocType>

⁸ Source: reports compiled by Forrester Research; <http://www.forrester.com>

One of the most acclaimed things about the HTML5 is the simplicity of the audio and video elements implementation. The use of the <AUDIO> and <VIDEO> tags, even if not all the capabilities of the <VIDEO> tag especially are not yet widely implemented, has an important role in the adoption process of the new language for the media-driven websites.

A lot of users prefer to access audio and video content directly through the browser, without the need to rely on a dedicated application. Taking into consideration the most recent data about this issue I can cumulate no less than 83% of the browser market of today as being able to play HTML5 video content.

Browser/Device	Market Share	HTML5 Video	Flash Video
Chrome	32%	✓	✓
Firefox	18%	✓	✓
Internet Explorer 9/10	16%	✓	✓
Internet Explorer 6/7/8	10%	✗	✓
Safari	4%	✓	✓
Opera	1%	✓	✓
iOS	7%	✓	✗
Android	5%	✓	✗
Other (feature phones)	7%	✗	✗

Figure 2. The usage of browsers able to render HTML5 video content⁹

As I mentioned before, this is the minimum number to be taken into consideration. The actual figures can be even higher, at least a part of the browsers that could not be identified and analyzed must be also HTML5 capable. There is also a number of specific devices, such as smart TV sets, set top boxes or gaming consoles that might also involve HTML5 capable browsers.

Unfortunately, I have to also come back to an above mention related to the “full implementation” of the HTML5 <VIDEO> tag. As I presented earlier, 83% of the browsers in use today can decode the <VIDEO> tag – but not all browsers can render the correct result of different other elements related to the <VIDEO> tag. For instance, perhaps one of the most important and innovative elements of the <VIDEO> tag environment is the <TRACK> tag. This command gives the programmer the possibility to insert captions, subtitles, descriptions, chapter markers and other metadata over the video content. Some

⁹ Source: <http://www.longtailvideo.com/html5/>; based on data from StatCounter

other mentionable elements related to <VIDEO> tag are WebVTT [4] and JavaScript capabilities.

Browser/Device	Basic Captions	WebVTT Styling	JavaScripting
Chrome			
Firefox			
Internet Explorer			
Safari			
iOS			
Android			
Opera			

Figure 3. The implementation of basic captions, WebVTT and JavaScript for HTML5 video content¹⁰

3. CONCLUSION

The HTML5 is the best bet at the moment for most web developers.

This statement is fully arguable but it must be commented a little bit. At this moment, the HTML5 is the best mélange of web technologies, standardized by a set of common specifications, is quite easily learnt and in the same time is complex enough in order to cover most of the situations met during a web application development cycle. But, it is a bet – the language looks excellent of W3C specifications but we should not forget that not everything is yet implemented or, in some cases, the implementation renders a different result when looking at the same website with different browsers.

I have started using HTML5 a couple of years ago, up to that moment I did only reading on the topic. The first experiences with HTML5 can be very enthusiastic but the downturn comes quite quickly – when you start developing the workarounds for the cases when the specifications are not yet implemented. I would recommend the use of the HTML5 approach especially for trendy projects, with targeted and predictable customers but I would also caution about the use of the technology for enterprise-grade applications. The implementation of the HTML5 elements is clearly not at this level yet and the cost involved in developing an enterprise-grade web application based on such an immature framework are very high [5].

Even thinking about these real issues that HTML5 still has, a diligent web applications developer must foresee the path and the trends, must prepare for tomorrow and must always

¹⁰ Source: <http://www.longtailvideo.com/html5/>; based on data from StatCounter

learn first-hand the new technologies and paradigms that constantly appear on the digital-era scene. Yes, sometimes there is no factual gain from this but in any case each thing we learn at one time might at least be considered as an experience, and a larger set of such experiences will always enrich our horizon and capabilities.

4. REFERENCES & BIBLIOGRAPHY

- Tabusca, Silvia Maria - "The Internet Access as a Fundamental Right"; published in "Journal of Information Systems and Operations Management", Vol.4.No.2, ISSN 1843-4711, pages 206-212
- Tabusca, Alexandru; Enaceanu, Alexandru – "Network security increase by using extended validation secure socket layer certificates for avoiding the phishing threats", published in Annals of DAAAM for 2009 & Proceedings of 20th DAAAM International Symposium "Intelligent Manufacturing & Automation: Theory, Practice & Education", Austria, Vienna, pages 0817-0818, 2009
- Pirjan, Alexandru - "Improving software performance in the Compute Unified Device Architecture"; published in "Informatica Economica Journal", 14(4), ISSN 1453-1305, pages 30-47
- Ian Devlin – "HTML5 Multimedia: Develop and Design"; published by Peachpit Press, 2011, ISBN: 978-0-321-79393-5
- Gabriel Eugen GARAIS – "Web Applications Readability"; published in "Journal of Information Systems and Operations Management", Vol.5.No.1, ISSN: 1843-4711, pages 114-120