Title: Understanding the role of Digital Commons in the Web; The making of HTML5

Authors: Raúl Tabarés Gutiérrez

Work Center: Fundación TECNALIA Research & Innovation

E- Mail: raul.tabares@tecnalia.com

Orcid: 0000-0002-8149-3534

Acknowledgments: Publication subsidized with funds from the program to support Technology Centers and their Corporations and Technological Alliances, Emaitek Program, Basque Government.

Summary: The last version of Web´s hypertext standard has been developed from 2004 to 2014. During this era, HTML5 has experienced different crossroads due to the variety of motivations and needs that the main stakeholders interested in its development had. On October 2014, the standard got official by the W3C and closed a period of uncertainty around the future of the Web but at the same time, this agreement also introduced a major change in the own conception of the hypertext´s standard. In this paper we review the current status of digital commons on the Web and the development of HTML5. We also confront this analysis with several semi-structured interviews carried out with different experts in web development that represent at the same time different players of Web´s value chain.

We argue that the development of HTML5 represents a new digital commons that prevented the proliferation of proprietary software that took place during the “Web 2.0” period. We claim that the World Wide Web promotes the development of new digital commons due to its own basis as a non-proprietary socio-technological platform. We also conclude that the development of standards and non-proprietary digital technologies is of utmost importance for the future of web business models that are fuelled by major digital players.

Keywords: HTML5; web 2.0; internet; metadata; web standards; web history
Bibliographical note

Dr. Raúl Tabarés comprises an area of expertise that combines Digital Skills, Social Media and Social Studies of Science & Technology. Raúl has published several journal articles, conference papers and book chapters and has been a speaker in different seminars, workshops and events. He has contributed to developing several strategies to foster the embrace of Web 2.0 technologies and skills across TECNALIA. He has taken part in several RTD projects in these topics, mainly at European level (especially H2020). Nowadays, his research is tightly related with the maker movement, DIY culture and user innovation.
1. Introduction & Methodology

Digital commons are a prominent element in the knowledge economy (Bauwens and Kostakis, 2015; De Rosnay and Le Crosnier, 2012; Vercellone et al., 2015). The predominance of open source software in different platforms and ecosystems (Castells, 2001; Himanen, 2002; Levy, 2010; McNamee et al., 2012; Mounier, 2002) has made possible new kinds of structures where several business models co-exist thanks to the prior development of digital commons structures (Fuster-Morell, 2010; Kostakis and Papachristou, 2014; Wittel, 2014). In this sense, we would like to pay attention to the hypertext standard called HTML5 that was finally concluded in 2014 (W3C, 2014). This major innovation has added significant technological capabilities to the World Wide Web and has laid the foundations for an evolution towards a much more interoperable, dynamic and automated platform (Tabarés-Gutiérrez, 2016). In this paper we review the state of the art of the digital commons and we describe at the same time the making of HTML5 Web standard. The aim of the text is to understand the role that play these structures in the Web for positioning different business models and capturing users value from different platform users. At the same time, we want to shed some light in the complex development of Internet standards which are affected by commercial interests, user motivations and participatory cultures in an increasing fragmented ecosystem.

To achieve that goal, we provide a critical analysis about the collective and cooperative innovation processes that have been carried out in the making of the new Web standard and how these processes were framed according to the personal, collective and institutional motivations for developing or deterring particular technologies. In order to contrast this analysis with some lead users of this technology we have carried out 17 semi-structured interviews with 21 HTML5 experts for obtaining qualitative outputs. These representatives are at the same time members of different stakeholders of Web’s value chain and they have been selected as key informants upon the involvement of them in the development of HTML5. That’s why participants have been chosen from major Web browsers (Opera, Firefox, Microsoft), search engines (Google, Yandex), standardization authorities (W3C), social media websites (Tuenti, Karmacracy), research organizations (TECNALIA, University of Deusto), new start up’s focused on the technology (Ludei, Otogami, Wimi5), mobile devices manufacturers (Blackberry), Web design studios (La personnalité) and technology user communities (HTML5 Spain).

An interview guide has been developed for enabling fruitful conversations with the stakeholders involved in the research. Selected questions were framed in the use of the technology by the participant, their attitudes and values, user adoption reasons, the role of multimedia contents, business models associated, digital commons views, etc. Content analysis have been carried out following a proper tutorial (Olabuénaga, 2012) and paying attention to the tipping points that were previously identified in the historical development of the Web and HTML5. The hypothesis that we wanted to be tested in the interviews were enclosed in the social construction of the technology approach (SCOT model) (Bijker et al., 1987; Bijker and Pinch, 1984) for understanding the different motivations, attitudes and values that major digital players had in the future development of this powerful technology.
The interviews have been deployed from July 2014 till the beginning of 2015 and all interviewees have taken part in a voluntary basis, they have been informed properly and they have received an Informed Consent Form (ICF) that guarantees their rights at any time in their involvement of this research. All interviews have been taped and transcripted for enabling content analysis and have happened physically, when possible, or throughout Skype when not (Iacono et al., 2016). Last, outcomes of this research have been shared with the different participants.

2. A growing interest in Digital Commons

Commons interest has risen in recent decades due to different causes. The lack of responses of capitalism (Brynjolfsson and McAfee, 2014; Mason, 2016; Rifkin, 2014) to the different and emerging challenges that society is facing has pushed the development of alternative approaches (Bauwens and Kostakis, 2015; Dafermos, 2015; Lafuente and Valverde, 1998; Wittel, 2014). Globalization, climate change, sustainability, unemployment, inequality and other great problems have become a major force that has shaken welfare capitalism (Prandini et al., 2016). The industrialism paradigm is also facing a significant crisis due to the transformation that has been experimenting from the 1980s towards a network–based form of organization (Castells, 2001). This new structure of work demands of flexible and agile structures that can create value through the personalization of goods and services (Beer and Burrows, 2010; Brake, 2014; Ritzer and Jurgenson, 2010). In this sense, the open nature of information technologies is an outcome of the social construction of the commons (Vercellone et al., 2015).

These kinds of “Social production processes” or “Commons-based peer production” (Benkler, 2006) that happens at the digital space have recently emerged as an alternative approach to the traditional state and market predominant organizational forms (Fuster-Morell, 2010). Commons refers to a variety of solutions, resources and knowledge that are shared by a community of commoners (Benkler, 2006; Fuster-Morell, 2010; Ostrom, 1990). These resources can be multiple kinds of things but commons always consists of three elements (Wittel, 2014):

- People that share the commons
- Resources that are being produced/shared/consumed
- A normative framework that establishes the rules of how commons are producing/sharing/maintaining and developing further

Elinor Ostrom’s detailed study of the commons demonstrates that this system operates in disparate areas such as the Spanish Levante (through the Water Tribunal), the Swiss region of Törbel (with communal pasture management) and the Turkish coast (in the Alanya fishing grounds) (Ostrom, 1990). In each of these case studies, Ostrom describes the resources and the regulatory framework governing their use, which is tailored to local needs and how can be modified by their users with the lurking of external authorities. Perhaps, Ostrom’s greatest contribution in this book is to empirically refute Hardin’s argument in his article “The Tragedy of the Commons” (Hardin, 1968) wherein he contends that any collective use of resources leads to their excessive use and depletion. This argument was
used to lay the ideological foundations for the claim that private or state ownership is by far the best way to manage a resource since rational individuals only act in the interest of maximizing profits. However, studies of the commons have identified examples of optimal collective management of resources all over the world which are efficient, careful and sustainable over time (Bauwens and Kostakis, 2015; Kostakis et al., 2015; Ostrom, 1990; Vercellone et al., 2015). Furthermore, anything that is private or public property can change hands, a risk that the commons avoid since they belong simultaneously to all and none of their users (Lafuente and Valverde, 1998).

In this sense and with the extraordinary development of digital technologies we have witnessed the emergence of different resources created throughout a commons basis. These digital commons does not refer to the material world but to the immaterial world and the realm of culture (Wittel, 2014). These are the products of social production (Benkler, 2006) processes that are necessary for social interaction and further development. Resources that can be framed as languages, codes, knowledge, etc (Hardt and Negri, 2009). A clear definition can be found in the dissertation thesis of Mayo Fuster Morell. She enounces digital commons as “information and knowledge resources that are collectively created and owned or shared between or among a community and that tend to be non-exclusively, that is, be (generally freely) available to third parties. Thus, they are oriented to favor use and reuse, rather than to exchange as a commodity. Additionally, the community of people building them can intervene in the governing of their interaction processes and of their shared resources” (p5 Fuster-Morell, 2010)

Mayo also pays attention to the growing importance of digital identities for the new generations. That is why she stress that creativity, remix and sharing cultures are really at the backbone of the young people that embrace these kinds of digital technologies (Fuster-Morell, 2013). This fact creates an enormous space for user generated contents in different digital platforms (OECD, 2007; Ritzer and Jurgenson, 2010) that have to be channeled through new digital business models (Dijck, 2013).

Sharing and collaboration have become significant values in the current society due to the new forms of production and creativity that have been fostered in the digital space (Beer and Burrows, 2010; Botsman and Rogers, 2010; Brake, 2014; Munar and Jacobsen, 2014; Ritzer and Jurgenson, 2010; Sundararajan, 2016). But these forms doesn`t offer novel ways for producing goods and services, they also offer alternative ways for reorganizing our society too (Lessig, 2004). This is one of the most special characteristics of the digital commons because digital resources cannot be exhausted if overused, but at the same time there is a great risk of underuse if no one curates, develops, edits and update these digital assets (De Rosnay and Le Crosnier, 2012). This is one of the reasons that have led to some authors to stress that the digital culture that is been developed throughout the World Wide Web and other digital platforms is spreading simultaneously a participatory culture in society (Beer and Burrows, 2010; Jenkins, 2006).

FLOSS (Free/Libre/Open Source Software) is one of the most popular examples of digital commons but other examples likewise IETF and W3C standards, Internet protocols, Wikipedia encyclopedia and Mozilla Firefox browser are also famous too. Most recently, open hardware (Anderson, 2012; Johar et al., 2015; Lindtner, 2014; Morozov, 2014; Smith
et al., 2013) is contributing to the popularization of digital commons in society too. Some examples of this emerging trend are Arduino, Raspberry Pi, Rep Rap and Makerbot (Dafermos, 2015; Kostakis and Papachristou, 2014). Open source projects that make digital technology available in the real world and allows to co-create products and goods near a zero marginal cost (Rifkin, 2014) throughout the emergence of Makerspaces (Niaros et al., 2017).

3. The making of HTML5

The birth of HTML5 constitutes a major turning point in the development of Web standards (Tabarés-Gutiérrez, 2016, 2015). The origins of this innovation lie in a group of professionals from private companies who were unsatisfied with the roadmap of the organization that supervises the development of the standards, the W3C.

In June 2004, at a workshop on Web Applications and Compound Documents hosted by the W3C, several developers from Opera Software and the Mozilla Foundation (and subsequently others from Apple) formed the Web Hypertext Application Technology Working Group (WHATWG) independently of the W3C (Franganillo, 2010). The workshop brought together browser vendors, Web development companies and other agents that were part of the W3C. At the event, Opera and Mozilla professionals presented their vision of the future of the Web which could be summed up as follows: “To transform HTML4 into a standard capable of including new features for modern Web application developers” (Pilgrim, 2010).

They also stressed the importance of the 7 principles that informed their proposal (The Mozzilla Foundation & Opera Software, 2004):

- **Backwards Compatibility, clear migration path**
  Web application technologies should be based on technologies authors are familiar with, including HTML, CSS, DOM, and JavaScript. Basic Web application features should be implementable using behaviors, scripting, and style sheets in IE6 today so that authors have a clear migration path. Any solution that cannot be used with the current high-market-share user agent without the need for binary plug-ins is highly unlikely to be successful.

- **Well-defined error handling**
  Error handling in Web applications must be defined to a level of detail where User Agents do not have to invent their own error handling mechanisms or reverse engineer other User Agents’.

- **Users should not be exposed to authoring errors**
  Specifications must specify exact error recovery behaviour for each possible error scenario. Error handling should for the most part be defined in terms of graceful error recovery (as in CSS), rather than obvious and catastrophic failure (as in XML).

- **Practical use**
  Every feature that goes into the Web Applications specifications must be justified by a practical use case. The reverse is not necessarily true: every use case does not necessarily warrant a new feature. Use cases
should preferably be based on real sites where the authors previously used a poor solution to work around the limitation.

- **Scripting is here to stay**
  But it should be avoided where more convenient declarative markup can be used. Scripting should be device and presentation neutral unless scoped in a device-specific way (e.g. unless included in XBL).

- **Device-specific profiling should be avoided**
  Authors should be able to depend on the same features being implemented in desktop and mobile versions of the same UA.

- **Open process**
  The Web has benefited from being developed in an open environment. Web Applications will be core to the Web, and its development should also take place in the open. Mailing lists, archives and draft specifications should continuously be visible to the public.

Besides this proposal, participants’ opinions were appealed through a poll on a question asked by Ian Hickson of Opera Software: “Should the W3C develop declarative extension to HTML and CSS and imperative extensions to DOM, to address medium level Web Application requirements, as opposed to sophisticated, fully-fledged OS-level APIs?” (The Mozilla Foundation & Opera Software, 2004).

The result was 8 votes in favor and 11 votes against. In addition to this rejection, the W3C made a formal declaration assuring that no support or resources would be offered to any proposal that did not come from the working groups previously established (Pilgrim, 2010). Because of this resounding denial from the W3C, the group of people who had made this pitch decided to continue their work out of the W3C. They registered the domain whatwg.org and began to work on their ideas founding the Web Hypertext Applications Technology Working Group (WHATWG).

This unofficial working group open to third parties consisted of browser vendors and different stakeholders who wanted to return to the origins of HTML rather than developing other languages such as XHTML. This approach was based on guaranteeing backwards compatibility of the new standard (Chau, 2009), a feature highly appreciated among web developers, and one asset that undoubtedly contribute to the coherence of future technological development.

However, the W3C wished to continue the development of XHTML2. This language did not offer backwards compatibility with HTML and required another “MIME type” (the document type specified at the beginning of web pages). This option entailed dumping all previous work achieved on HTML because browsers had always “overlooked” HTML syntax errors and until then, no one had thought about how to specify them. We can say that till this point, web development had relied largely on different browser developers making their products compatible with those of their competitors and ignoring specifications and standards. Most browsers were focused on “being able to present” the “tag soups” generated in the best way possible (Andersson, 2007). The WHAT Working
Group devoted 5 years of work to document how to analyze and dissect HTML properly in a manner consistent with existing web pages (Pilgrim, 2010).

One of the experts interviewed, who formed part of the original working group, explained the causes of this situation: "We were worried about XHTML2 and we didn’t believe XHTML was the future so, the features that we cared about was extending the Web without breaking backwards compatibility. Everything else all the other stuff like Web Workers, Web Sockets, all of these things are extras. But the primary goal was two things;

- Extend the Web without breaking backwards compatibility.
- Interoperability

So, one of the central features of the HTML5 that people don’t talk about so much because it is not sexy is what we called the parsing algorithm. The HTML5 parsing algorithm fixes that so every website will make the same DOM in every browser."

(Bruce Lawson, Opera)

Besides this staggering task, the group also worked on native support for audio and video (without plugins), the “canvas” element (which is used to draw graphics using a document structure) and other web application specifications. The reason for setting up the group was the disagreement with the W3C’s academic vision of standards and the aim to provide a much more practical focus for the new standard (Chau, 2009).

However, the W3C appeared to be seeking a replacement for HTML from among various technologies, particularly XHTML 2.0 (O’Mara, 2012). The truth is that two and a half years after the workshop XHTML2 was languishing while the new features of HTML raised high expectations. This fact combined with the criticism to the W3C’s slow progress and little concrete results (Castro, 2007) was one of the reasons that accelerated the announcement of Tim Berners-Lee and the W3C in October 2006. From this date they would be working together with the WHATWG to add extra features and push forward HTML (Berners-Lee, 2006). Subsequently, in October 2009, the W3C dismantled the XHTML2 working group (Le Hegaret, 2009) and abandoned activities in this language in order to concentrate its efforts on developing HTML5. In 2008, this consortium also published a working draft of the standard (Hickson and Hyatt, 2008) thanks to the joint efforts of the two organizations. In the following years browsers began to support HTML5 (Mozilla Firefox was the first to take the plunge) and a phase of raising awareness about this technology started.

However, it was Steve Jobs’ open letter, “Thoughts on Flash”, that achieved most impact in terms of social dissemination. His post highlighted the priority that should be given to mobile devices when developing standards and the problems generated by proprietary software such as Flash. Steve praised HTML5, CSS3 and JavaScript for their features and their nature of open standards (Jobs, 2010). Several of the experts interviewed during our research acknowledged the importance of these statements by the famous ex-CEO of Apple.

Other influential companies and platforms such as YouTube (Harding, 2010), Slideshare (Slideshare-Blog, 2011), and Scribd (Calore, 2010) also declared their public support for HTML5 as well as their intention to implement this technology on their applications as far as possible. Since then, development of HTML5 has continued progressively, first as a
W3C candidate recommendation (W3C, 2011) and later as an official recommendation on October 28, 2014 (W3C, 2014).

As noted by several of the experts we interviewed during our fieldwork, the appointed day did not represent a complete break, as most of the features of the standard were already in use and were implemented on most browsers, which was one of the main ideas behind this development (Hickson, 2011; WHATWG, 2005). The specification is available at the following link: http://www.w3.org/TR/2014/REC-html5-20141028/.

4. Changing a paradigm in Web standards

The rising of HTML5 has introduced different innovations into the hypertext structure but above of all it has created a new software platform that can be used by different agents to promote new applications and business.

The new hypertext standard has tackled most of the problems of standardization that introduced the Web 2.0 phenomenon which made possible that a new kind of user interested in creating and consuming multimedia contents entered the platform. That’s why Social Media platforms like YouTube, Facebook and Blogger delivered several innovations for promoting User Generated Contents (Dijck, 2013, 2009; OECD, 2007) that created conflicts with the previous hypertext technologies and were forced to use external elements because of the lack of native capabilities of the standard to deal with multimedia and dynamic contents (Harding, 2010). At the same time the irruption of mobile devices (Vogelstein, 2013) and native apps created a momentum of uncertainty around the future of the Web (Andersson, 2007). This “battle of decade” between Apps and Open Web (Mikkonen and Taivalsaari, 2011) has resulted in a new version of HTML that is not only an hypertext standard but also a software platform that has been benefited from the cooperative work of different companies that have created an infrastructure where new services, applications and platforms can be built. In this sense, we can observe how different actors have cooperated towards the creation of a new technology with different motivations, meanings and interests (Bijker et al., 1987; Bijker and Pinch, 1984; Latour, 1992).

HTML5 constitutes a major turning point in the history of web standards as it has extended the web environment to other kind of devices that were not present in the original design of the hypertext. It also has changed the approach of standardization processes as a more practical focus has been imposed in detriment of the academic basis that was the rule in previous episodes of the history of the Web. Last but not least, the introduction of several semantic elements has turn this platform into a giant database much more automatized that has paved the way for introducing big data tools (Boyd and Crawford, 2011; Kusnetzky, 2010) that explodes different companies raised during the Web 2.0 decade.

4.1. HTML5 as new digital commons

As previous hypertext standards HTML5 is also an open source platform. This type of software can be used freely to develop web pages and/or modify existing ones. It is based
on universal free availability and use of content, offering total freedom of use, modification, reuse and redistribution of the source code (Alonso and García, 2012). These rights are granted voluntarily by the creator and owner of the software copyright through a specific license which concedes part or the totality of the rights. Such licenses are commonly known as copyleft (Lessig, 2004).

Besides the standards involved within HTML5, such as CSS3 and JavaScript, other types of FLOSS software are commonly used on the Web in both collaborative and personal projects. The existence of content management systems (CMS) such as Wordpress, Drupal, and Joomla, the Apache HTTP web server, browsers like Mozilla Firefox and databases as MySQL indicates the prevalence of this type of software on the Web. In fact, this type of applications has formed the basis for many collaborative projects, leading for instance to the development of Linux and Wikipedia (Anderson, 2007). The open source model has risen at the same time that the role of the Web has gained importance in society (Alonso and García, 2012). For that reason, web values and processes have also been transferred to different areas of society and in this way, the university education sector is one of the most recent examples (Figuerola et al., 2007).

Consequently, some authors have seen the predominance of FLOSS on the Internet as an evidence of the Web as a driving force for the social production of content (Benkler, 2006). Specially through what is commonly known as “P2P” or “commons-based peer production” (Benkler, 2006). This is the collaborative production basis that has enabled the development of many open source projects and other collaborative initiatives worldwide. In this case, we argue that the rising of HTML5 mirrors a crisis concerning the future of the Web. The proliferation of proprietary software to support multimedia elements (especially Flash), the emergence of mobile devices and the excessively academic approach to XHTML2 all combined to call into question the viability of the Web as a platform. The development of HTML5 represents new digital commons that has emerged to strengthen the technological structure of the Web.

As we have explained before, the role of digital commons in Internet standards is especially important as there are infrastructures that have created by the cooperative efforts of several companies (De Rosnay and Le Crosnier, 2012; Vercellone et al., 2015), institutions and individuals committed with different values and attitudes. Openness is an ideal that attracts entrepreneurs, hackers and innovative companies (Russell, 2014) but it also demands awareness, participation and negotiation (Schrock, 2014) for promoting a shared consensus about technological trajectories. Nobody has the power to shut down the Web because it is a platform on which a vast number of companies, organizations and users collaborate and the development of digital commons such as HTML5 has made it possible to tailor web standards to the real needs of users. This has prompted a way to incorporate the views and voices of a range of stakeholders who were previously unable to participate in this process and an element of meritocracy in decision-making processes that were not present before, contributing to introduce a more participatory approach to deliver web standards.

“If you make the web much stricter like XHTML2 it will lead to a better-quality Web overall, cause when you really look around in the Web these days you still really see horrible and awful markup, and I
comprehend why people argue for more quality markup. But at the same time, should the Web really be about W3C wanted to be or should it be about what it actually is? I think that it should be what people actually are doing rather than W3C want people to do. HTML5 more accurately represents what people do in the real world. Because the people do want to write.” (Chris Mills, Mozilla Firefox)

But at the same time this cooperative effort of different stakeholders in the digital economy also has been pushed by different interests of major players that were pursuing new commercial opportunities in the transition towards a mobile devices paradigm (Schrock, 2014). This episode in the recent history of the Web provides an idea of how the participation of different agents in technological development can contribute to developing standards that are more aligned to the real needs of users but also on how big companies work towards the promotion of different values that can attract participation and involvement of different collectivities that can being converted to their interests. This can be also observed in the “Living Standard” (Tabarés-Gutiérrez, 2016, 2015). A new way of developing web standards more focused in the standardization of useful elements without having to become entangled in excessively long and tedious processes of deliberation and consensus. The “Living Standard” is a way to test functionalities demanded by users in experimental environments safeguarding at the same time the stability of the web specifications but it is also a working group guided by major interests of web browsers.

4.2. Has HTML5 put an end to the Browser wars or it’s just a ceasefire?

One of the big questions that stands up after the rising of HTML5 lies on questioning if this web paradigm also implies a major agreement on the key aspects that subsequent web standards will have to incorporate as the recurrent series of conflicts among browser vendors throughout the history of the Web tend us to be skeptical about it. According to Ian Hickson, the evolution of web standards for enabling digital platforms building blocks is of outmost importance;

“I think the main difference in how browsers competed back then vs how they compete now is mainly that the standards world is more competent now, so we can spec out what browsers do faster. But we still have lots of disagreements and tensions between browser vendors, certainly. For example, look at pNaCl vs asm.js, or how the Web Components work is developing, or the many APIs that Mozilla is inventing around Firefox OS, etc.” (Ian Hickson, Google)

The robust technological capabilities that are at the backbone of the standard makes a stronger hypertext markup language that is not at the mercy of the innovations promoted by the different browsers. HTML5 now comprises a large number of previously nonexistent technologies over which new “browser wars” are being conducted. Everything related to the mobile and multimedia technologies has created new battlefields with new players and new application ecosystems but they are all possible because of the Web standards. Specially, major digital players have extended their operational area and have gained a dominant position in media industries thanks to the development of new native Web standards capabilities (Schrock, 2014).
The most significant change in the history of the development of web standards has been the shift from prioritizing PCs in favor of mobile devices (Tabarés-Gutiérrez, 2015) which are increasingly common as a gateway to the Web. This multi-device approach has created different needs from the PC era and has stressed the importance of values such as ubiquity, mobility, and interoperability. Consequently, the standard must be more robust in terms of technological capabilities and features. It should not be forgotten that mobile devices have less computing power than desktop devices and therefore these capabilities must be provided by the standard.

“HTML5 has provided the biggest boost in history to web technology. It has prompted a hitherto unseen evolution of something that works in a browser. Thanks to HTML5, browsers have become computing platforms. Now, I can use my browser to watch video, generate audio, see 3D applications, and consume content with a presentation rich in animations and transformations. This is what HTML5 has done.” (Ibon Tolosana, LUDEI)

Furthermore, it should be noted that the development of HTML5 also represents a quest for establishing an open innovation ecosystem in Web standards development. In previous episodes in the history of the Web the innovation potential of browsers has overruled many of the core values of the hypertext standard due to commercial purposes and technological innovations that didn’t fit to a minimum set of standards requirements (Tabarés-Gutiérrez, 2012). These inventions were aimed at attracting a greater number of users to increase the market share and therefore providing bigger profits for the companies behind these applications. But nowadays, with a more fragmented market, major digital companies like Apple or Google are also affected by the proliferation of proprietary software as those elements disrupt their expansion to the media ecosystem. HTML5 is a way to promote a technological platform for entering the media industry for several Internet companies that want to disrupt this business (Schrock, 2014) and at the same time creating an open digital infrastructure that is maintained by these emerging actors.

This is the revolution implied in the development of HTML5 and most particularly the great contribution of the WHATWG group, on which the work of Mozilla Firefox, Opera Software and Apple was based, and which was subsequently supported the W3C. In the history of standards we can find another similar experiences as Internet standardization communities are famous also for pioneering organizational innovations (Russell, 2014, 2006). These procedures, behaviors or patterns are the result of “constantly dealt with problems that stemmed from the tension between centralized authority and grassroots initiatives, as well as the rising influence of commercial values” (Russell, 2006).

These commercial values that are at stake at this reorganization of the Web as a source of emerging business models for digital companies is clearly one of the drivers for promoting a new common technological base that is backed up by social production (Benkler, 2006) processes and the role of digital commons (De Rosnay and Le Crosnier, 2012). Moreover, we cannot forget that there is an increasing number of companies, industries and countries involved in the so called digital economy (Fuster-Morell, 2013). Some years ago, companies from USA and Europe were really present in these standardization committees but nowadays it is more common that companies from no western countries have their
presence too. Russia, China, Brazil, India and other countries where the digital economy is emerging have also their big players and they also bet on HTML5.

These emerging business and markets are characterized by a tremendous complexity due to the different economic inequalities in their societies that they are also present in the digital scene. Yandex, Baidu, Weibo, Tencent and many other companies in these non-western markets that are not aligned with the values of the “Californian Ideology” (Barbrook and Cameron, 1996) have started to gain importance at the international level and have contributed to a multi-cultural approach in the construction of HTML5 (McKenzie, 2013). The social construction of technology (Bijker et al., 1987) is getting global and there is a growing need to build robust web standards that meet all requirements from all multi-cultural perspectives.

“Also, we in the West tend to forget that there are other companies and markets than ours and that there are other “players”. In China for example, where the Internet economy and web services are exploding, there are many browsers that we’ve never even heard of but which are also based on HTML5. It would be crazy to start another browser war, with the vast range of options out there today.” (Martin Álvarez, W3C Office Spain)

The development of all these collective and cooperative innovations in recent years by the major browser vendors has generated technological resources for communal use. These web standards comprise free software that can be used, adapted and/or reformulated according to the interests of potential users. One particular example is the Chromium project that was used to develop Google’s Chrome browser and operating system, and that was used also by other browsers and stakeholders.

This type of open source projects has opened up alternative possibilities for collaboration between Internet companies for sharing resources. Although each browser offers a different interface and a diverse user experience, the social production of these technologies has enabled many of them to share various capabilities and features (Anthony, 2013). These collaborations between browser vendors have occurred because of the increasing fragmentation of the ecosystem of devices, operating systems, platforms and users. Now, it is really difficult for a single company to occupy a dominant position as it was in the past with Internet Explorer (Tabarés-Gutiérrez, 2016, 2012). Given this disparity, a robust standard and amply endowed with technological capabilities is required as computing needs of mobile devices and the doorways of accessing the Web are so fragmented. Nowadays, the paradigm of the Web standards industry relies on a similar user experience no matter what kind of device is used to access and in this sense, it seems that all browser companies are committed to provide a common ground.

“A few years ago, there was a clear difference between IE, Firefox, Opera, and the others, but now Safari and Opera use the Chrome engine and vice versa, since many of the innovations they produce are joint improvements. There is an open fight, not because we have moved from the desktop to mobile devices, but because the ecosystem was and is increasingly fragmented, which is very good news for standards. The less power a manufacturer has, the easier it will be to recognize the need for an official standard.” (Pablo Garaizar, Universidad de Deusto)
4.3. Metadata and its dark side

Blogs and other informative sites have played an important role in the development of HTML5 as the new tags that have been introduced have greatly influenced on their conceptualization by the irruption of these platforms during the Web 2.0 phenomenon (Schafer, 2010). The new hypertext standard presents a series of semantic elements that organize and improve document presentation. Some of these are article, header, hrgroup, nav, section, aside, and footer.

The dynamism of these platforms and the need for a timeline to organize the information presented on web pages together created a momentum for imposing new elements in hypertext structure that were not being met in a suitable way (Tabarés-Gutiérrez, 2015). Thus, several semantic elements have been deployed to address the weaknesses of the previous standard in this regard and they have been delivered after a phase of research that uncovered the reality of web developing at that time.

Ian Hickson, that was the specification editor, used several data mining tools owned by Google in order to extract information from around one billion web pages for identifying the most commonly used names for these generic div elements (Lawson and Mills, 2010) that were employed elsewhere. The element names used in HTML5 reflects the twenty most commonly used types in the pages analyzed in this study.

The Opera Software company also conducted a similar study with a sample of three and a half million URLs known as “MAMA”. Although the study was based on a smaller sample, it included a higher number of statistics and values contained in the source code of the websites. The empirical methodology employed by the two companies for these studies was highlighted by several participants in our fieldwork and acknowledging how important it was for the future standard to reflect the reality of development at that time.

“The way these elements decided to come, were driven by Google and Opera works. These companies looked to all of the most common text that was used in websites. So, if so many people are using a class of aside or article or post or whatever, it really make sense to have a tag that deals with this, in a proper semantic way. I think they did this in a very analytical and scientific way. HTML5 has heavy influenced upon very common sites archetypes such as blogs.” (Chris Mills, Mozilla Firefox)

Several companies were really interested in renaming different elements of the standard for allowing machines to understand the information they handled. Many Internet companies, such as search engines, wanted a more semantic approach that could facilitate and improve the type of search performed by users. This more automated style would enable such companies to offer better search results to their users, to enhance their competitiveness, to increase their market share and therefore to generate bigger profits.

“The use of these elements is necessary so that search engines can understand the document or web page content. That’s the reason for all these new elements, because previously with DIV’s, search engines didn’t understand anything they had encapsulated. For example, Google will be able to deliver much more accurate and intelligent search results with these new elements. Everything about HTML5 represents the answer to a whole heap of community requirements at all levels.” (Kiko Hernández, WIMA5)
These tags that facilitate information retrieval and indexing by search engines also represent the framework necessary to develop data capture or create digital traces in the form of metadata. Although this aspect has might receive less attention, it undeniably enables the establishment of a multitude of digital business models from different companies around the Internet throughout targeted advertising.

HTML5 presents significant innovations regarding the collection of information of Internet users thanks to various sub-specifications devoted to metadata within the general specification (W3C, 2014). These metadata are produced when electronic information channels such as the Web, Internet or mobile phones are used and they provide information such as the time or location of the communication. Despite the fact that the content of the communication is not included several platforms use these metadata to categorize and index all kinds of information (Schuster et al., 2014). Photos and videos are now more easily recognizable because while a user uploads a photo to a virtual social network, the standard provides tools to add a lot of information to the file like the date and time it was taken, the location, the user, size, title, etc.
Other features of the standard like the geolocation API allow to locate the user via the browser (Pilgrim, 2010). Cookies have been redesigned and improved through HTML5 storage proposal, known as “local storage”, is designed to accommodate much larger volumes of information and saving bandwidth (Pilgrim, 2010). These types of elements that provide all kinds of information about users and their browsing patterns make it possible for many of the companies created during the Web 2.0 phenomenon to monetize user-generated content (Dijck, 2013).

The lack of transparency in these data collection, management and reusing processes has given rise to various doubts in society since users lack informed knowledge about the subsequent uses made of their data. To this problem can be added the technological evolution of the most large companies in the digital economy which use various big data tools to enable much more advanced processing of large amounts of data while lowering at the same time the costs of collection and storage (Boyd and Crawford, 2011; Gray, 2014). Moreover, Snowden's revelations have push forward these concerns due to the collaboration of cyberespionage agencies with major Internet companies (Schuster et al., 2014).

5. A needed infrastructure for the Platform Economy

As we have claimed during the last section, the rising of HTML5 has created a new platform software that allow to generate new applications based on standardized and open source technologies. This has granted that many successful platforms during the Web 2.0 period have accommodated their business models to this new technological infrastructure, collecting data and metadata from their users (Helmond, 2015). Nowadays, more and more people has onboarded to the Web and to different social media services that were originated during the Web 2.0 period. According to O’Reilly, the companies that rose during this time were creating the “Web as platform” paradigm because they were developing new technologies and building new business models that created a “perpetual beta” in software and making the Web more important than ever (O’Reilly, 2005). But as we have observed this transition has not been completed till the standard has been capable to implement these new functionalities and creating a shared consensus about what it will be the future of the Web between different stakeholders.

Platforms like Facebook or YouTube emerged and became rapidly the hubs that allow users to create, edit and share multimedia contents that entice attention (Goldhaber, 1997) providing valuable spaces for targeted advertising and reconfiguring the nature of the Web. Nowadays we can see how a platform economy (Kenney and Zysman, 2016) or a platform capitalism (Srnicek, 2017) has been established around the role of data that is been scraped in platforms (Sundararajan, 2016), devices (Law et al., 2010) and the behaviors that can be inferred (Dijck, 2013). Despite the fact these companies are totally dependent on the contribution of human beings for creating new services and remaining competitive, the availability of large datasets that have been aggregating since the consolidation of UGC sites has made possible for them to obtain a dominant position in the Internet for disrupting different already established business thanks to the free labor (Terranova, 2000) and digital labor (Scholz, 2012) that have transformed in economic value.
This exploitation of commons by platform owners on the Internet (Fuchs, 2010; Fuster-Morell, 2010; Tufekci, 2010) has been possible due to the generation of narratives that speak about openness and neutral ecosystems (Barbrook and Cameron, 1996; Gillespie, 2010) but also for the promotion of a new technological path that is shared by the majority of users that are affected by this innovation.

This is one of the reasons that make us claim about the need of generating innovative structures for promoting digital commons in data generation, using and storing for preventing the commodification of it imposed by the dominant position of these companies in the market. New ways are required to protect users’ rights and empower them in digital environments because this ecosystem systematically exploit the entire activity of their users for a variety of commercial purposes.

6. Conclusions

During this text we have briefly reviewed the state of the art about the digital commons and we have highlighted the critical role that these infrastructures play in the Web and the Internet. Special attention has been paid to the development of HTML5 and how this “living standard” has been developed in a multi-stakeholder environment thanks to the cooperative work of different stakeholders. As we have argued, HTML5 has rising as new digital commons to stop the proliferation of proprietary software (especially Flash) that occurred during the explosion of the Web 2.0 period (Anderson, 2007). We claim that at the backbone of this push for hypermedia lies the different digital business models that major Internet companies were forging in the transition to the mobile paradigm and media industry but also by the need of creating a user-oriented platform for promoting openness ideal as a shared consensus between different stakeholders.

The combination of UGC websites, mobile devices and other technologies have boost new business models that rely on data (Srnicek, 2017) and need the technological basis that HTML5 has delivered. Of course the emergence of disruptive technologies like Internet of Things, Wearables, Smart Cities, etc. will favor new complexities (Echeverría & Tabarés, 2016) but it will also need standardization stages that will need a multi-stakeholder approach in open innovation ecosystems. In this sense, HTML5 development show some lessons that can be of interest to other collectivities that will be involved in these processes. Special attention to this matter have to be paid to the role of the major players in web browser industry and how they are able to create alternative paths in standardization committees. The accumulation of such power can defy traditional institutions, creating alternative paths and developing alternative narratives to entice outsiders while pursuing at the same time their primal objectives.

In this article we have been reflecting about the need of extending the digital commons basis to the data generated by these platforms as the treatment of it is the cornerstone of the new digital business models. As we have stated all of these companies create value throughout the maintenance of digital platforms that are filled by the activity and work of generic users in a prosumerism basis. This trend has been favored since the irruption of the Web 2.0 period and have been fueled by different investments of venture capital in order
to create common behaviors that are now embedded in personal routines. That’s why it’s important to extend the concept of digital commons to data as this is the next frontier for creating techno-social structures that can promote digital inclusivity.

Experiences like SOLID (Hasslberger, 2016), the last project of Tim.Berners-Lee, are aimed to develop decentralized social applications that can provide an alternative approach to the usual one that is deeply rooted in exploitation practices that promote inequality between platform users and platform owners. In this sense is of outmost importance to rethink the role of web standards as there are the major foundations for generating structures of cooperation and value generation in the Internet and for stopping the commodification of every data and the commodification of everyday living. Digital commons must be spread around data to create wealth and resources that can be shared and maintained by a community. This is the next step for preventing inequalities in the digital sphere and to foster a much more inclusive digital space and user-oriented.
Notes

1 Details of this workshop are available at the following link: http://www.w3.org/2004/04/webapps-cdf-ws/ (last accessed on 5/11/2017).

2 Their website and letter of intent are available at this link: http://www.whatwg.org/news/start

3 MIME (Multipurpose Internet Mail Extensions) types specify the type of document being transmitted, and are indicated at the beginning of web page codes. More information on the concept itself is available on the following Wikipedia page: https://en.wikipedia.org/wiki/MIME

4 This reference document is continually updated and is available on the following link: http://www.whatwg.org/specs/web-apps/current-work/multipage/parsing.html.

5 Yandex is a Russian search engine, Baiyu is a Chinese e-tailer and Weibo is a Chinese social network platform. All of them are really popular in their countries and they have hundreds of millions of active users.

6 More information about the project is available at its website: https://www.chromium.org/

7 Data mining is a field of computer science and refers to processes aimed at revealing patterns in large volumes of data sets.

8 The element <div> is used to define a content block or section of the page in order to apply different styles. It is a very common element in web development.

9 More information about this study is available at the following link: https://developers.google.com/webmasters/state-of-the-web/2005/classes?csw=1 [accessed 5/11/17].

10 More information about this study is available at the following link: https://dev.opera.com/articles/mama/ [accessed 5/11/17].

11 Files sent by websites that are stored in the user’s browser when the user requests a web page. The next time the user accesses the website these cookies enable the site to recognize the user and then to consult previous activity, to load site elements more rapidly, to access information previously provided by the user, to remember passwords, etc.

12 Big Data is a term encompassing the use of techniques to capture, process, analyze and visualize potentially large datasets in a reasonable timeframe, not accessible to standard IT technologies.
Bibliography


de la Frontera.


Highlights

- A review of the digital commons status and the emergence of HTML5 web standard are included.
- The role of metadata in digital business models is extensively explored.
- Future directions for a much more inclusive Web are provided.