

# CS 171 Group Project Proposal

*submitted by*

**Jennifer Le Hégarret, Zona Kostić, and John Greeley**

## **W3C – Test the Web Forward**

For our project, we propose to assist the W3C in its work to develop the standards for the web platform by helping it visualize its progress developing and testing its various specifications for web standards.

## **Background and Motivation**

Jennifer has personal connections with W3C and so offered that organization the chance to take advantage of some pro-bono labor via the CS 171 project. Together, they decided that the best data to work with would be the data related to progress made on the Specification drafts and Test Suite, as measured by both W3C GitHub activity and CanIUse.com statuses. This work will be very useful for W3C in terms of understanding what its working groups are actually working on and number of current to-do items which need to be addressed. The work is also very well-suited for this project due to the public availability of the data and the numerous data dimensions that are ripe for exploration and presentation.

Zona was happy to join Jennifer on this project as she has great respect for W3C and its work, and she wants to dedicate her efforts to a project with real-world applicability and relevance.

John was drawn to their team as he had previously worked with a government contractor to research and implement web services standards for cross-agency integration projects. He is therefore familiar with W3C standards and also willing to help support W3C's work through his CS 171 project.

## **Project Objectives**

We aim to help the W3C monitor what has been done and what needs to be done across all of the GitHub repositories that it uses to develop and to test its specs.

We aim to provide different views of the status of various pull requests and raised issues by working group, by spec, and by who is contributing. We would also like to incorporate available CanIUse.com data as a barometer measuring a given spec's implementation level.

In this way, we hope to help W3C and other interested parties are able to “Test the Web Forward” as effectively as possible. If we succeed in creating something truly useful to them, they will use it on their website: [www.testthewebforward.org/dashboard](http://www.testthewebforward.org/dashboard).

## **Data**

W3C has over 1000 Specifications that, combined, set the standards for web technologies. These standards are worked on by a collection of about 50 Working Groups. Each specification is an HTML document that moves through various stages – from Working Draft to Recommendation. Much of this HTML code is hosted on GitHub. For example, the latest HTML Spec is worked on here:

<https://github.com/w3c/html>.

Many of these specifications also have an associated Test Suite of JavaScript code, also hosted on GitHub. This Test Suite aims to provide a point-by-point test for every single detail within a given specification – like that, W3C, the browser developers, and other interested parties can analyze the current degree of real-world implementation of that specification. Basically, it is CanIUse.com to the *n*th degree. For example, the tests related to the latest HTML Spec are here: <https://github.com/w3c/web-platform-tests/tree/master/html>

In addition, [www.CanIUse.com](http://www.CanIUse.com) reports how well-implemented various particular features within a web technology have been implemented across browsers. We are able to tie their reporting back to a given W3C spec thanks to their Resources tab for a given feature. However, sometimes they reference the WHATWG version of the spec instead.

Our data will be obtained via the GitHub API, the CanIUse API, and our friendly connection at the W3C, who will list out for us each Working Group, the Specs that Group works on, and the status of each of those Specs. Our friendly connection will also remap CanIUse.com WHATWG Spec references to W3C Spec references.

Because we are focusing only on what data is available via GitHub for W3C work, our population of specs is reduced to about 70 specs, which are worked on by fewer than 20 Working Groups. While this is a significant reduction in scope (for example, CSS is not represented, as the CSS Working Group does not use GitHub), W3C is still very motivated to view the status of even these 70 GitHub-hosted Specs on our dashboard, and we find this still a very worthy effort.

## Data Processing

We are very fortunate that W3C is quite motivated itself to help make this happen. Our W3C connection dealt with working with the various APIs and assembling the W3C-specific data into JSON formats to our own specifications. They are hosted here: <https://github.com/plehegar/cs171-data>

We currently have:

<https://github.com/plehegar/cs171-data/blob/master/caniuse.json> - the mapping of CanIUse statuses to various elements across W3C Specs, judging implementation by latest browser versions only. Each browser's record provides its expected platform (desktop, mobile) and rates the known implementation on a scale of 4 (“yes”, “partial”, “no”, “unknown”). The element overall also has a name, description, a link to the caniuse page, and a reference to the W3C spec within its record in addition to the browser results for that element's implementation.

<https://github.com/plehegar/cs171-data/blob/master/groups.json> - a list of the Working Groups, with sublists of the specs each group works on, each spec's current status, and each spec's last publication date.

<https://raw.githubusercontent.com/plehegar/cs171-data/master/tests.json> - a list of all pull requests and issues for the test suite, with lots and lots of data about each one: the author, the state, the date created, how many lines were added, how many lines were deleted, and many more fields.

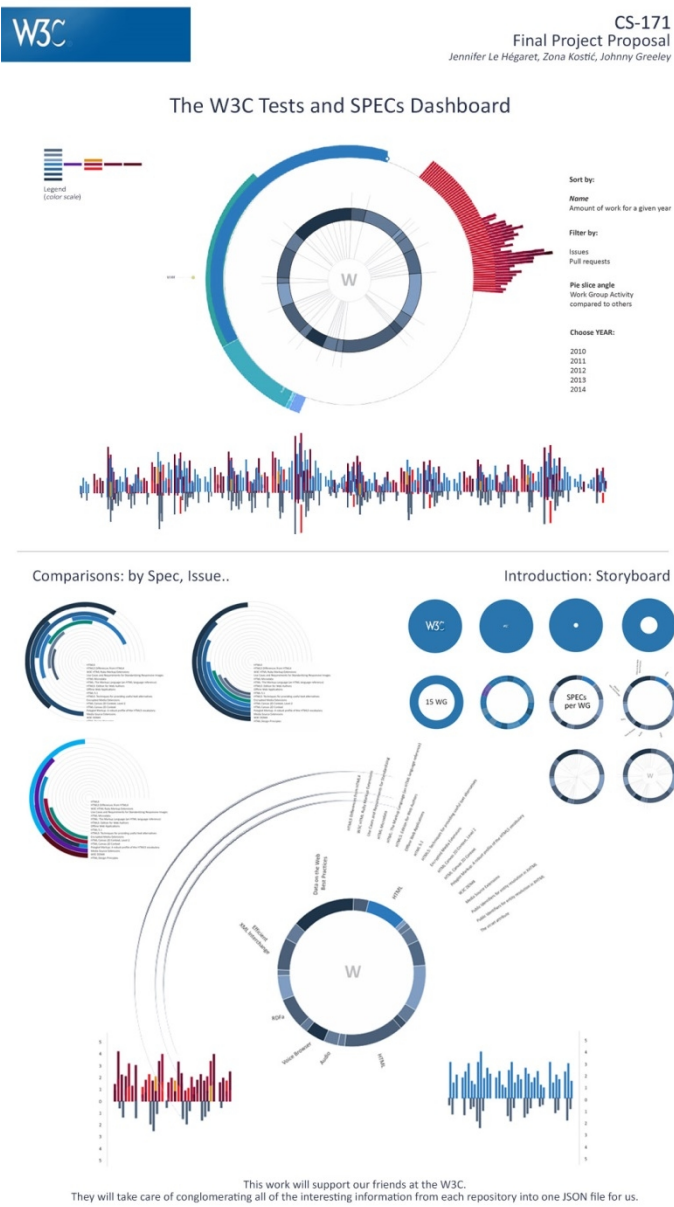
<https://raw.githubusercontent.com/plehegar/cs171-data/master/specs.json> - this last is still a work in progress, but will provide the same GitHub information for all specs as we have for the test suite.

Data processing will thus be very straightforward for us, as we will read these JSON files directly into the JavaScript environment via our D3 library's commands.

We expect to derive additional data as well, such as “Number of Days Elapsed” for timelines and a CanIUse “score” to summarize the overall implementation level of a Spec. We will keep an open mind as we work with the data and will probably add more ideas to this list as we go.

Visualization

We have a lot of ideas about how to present this data. Our favorite overall look and ambition is this amazing layout, as developed by Zona:



For other, rougher and less sophisticated ideas, you may look at the appendix of this document.

We plan to meet this weekend to discuss this visualization, exactly what it represents, and how it will work. What you see here is still a work in progress.

## **Must-Have Features**

At a minimum, we need to present the same information that is currently presented at the [www.testthewebforward.org/dashboard](http://www.testthewebforward.org/dashboard) website, but including Specification information and a sense of the CanIUse data as well as the Test Suite information which is already there. W3C wants to be able to judge progress over time primarily, and immediate action items secondarily.

Also, easier navigation between views of different levels is necessary: overall, by working group, and by spec within working group. Right now, there is just a very long, esoteric and jumbled pull-down menu to switch between different specs and groups.

## **Optional Features**

We very much want to provide a lot of interaction, so that the navigation between views is a matter of clicking and zooming, all the way to specific GitHub issues or CanIUse pages if desired, even while we provide a sleek overview of the current amount of work to be done.

We would also like to provide a sense of overall timelines about which issues have been outstanding for how long, or how long it has been since a spec's last published date as compared to its last GitHub update, etc.

Also, we hope to offer a view about “who” is doing what over time, given that this data is also available and interesting.

## Schedule

*(subject to change and refinement)*

Timing	Jennifer	Zona	John
3/30 – 4/3	<ul style="list-style-type: none"> <li>• Set up Repo</li> <li>• Obtain data</li> <li>• Proposal document</li> <li>• Team form</li> </ul>	<ul style="list-style-type: none"> <li>• Design work</li> </ul>	<ul style="list-style-type: none"> <li>• Come on board</li> </ul>
4/4 – 4/5	Review designs and allocate specific graphs to specific people		
4/6 – 4/16	<ul style="list-style-type: none"> <li>• Begin process document</li> <li>• Create website skeleton</li> <li>• Hopefully also help with creating specific graphs</li> </ul>	<ul style="list-style-type: none"> <li>• Refine design if needed</li> <li>• Create specific graphs to fit into website skeleton</li> </ul>	<ul style="list-style-type: none"> <li>• Create specific graphs to fit into website skeleton</li> </ul>
4/17	Submit working prototype		
4/17 – 4/21	Solicit feedback on working prototype from W3C and from CS171 TF <i>(We hope very much for a Tuesday face-to-face appointment during class time!)</i> Catch up on process document		
4/22 – 4/29	<ul style="list-style-type: none"> <li>• Incorporate feedback on specific graphs</li> <li>• Work on overall presentation website and documentation</li> <li>• Perhaps continue to add views of data</li> <li>• Perhaps add levels of interactivity across graphs</li> </ul>	<ul style="list-style-type: none"> <li>• Incorporate feedback on specific graphs</li> <li>• Design the story-telling aspect of our presentation website</li> <li>• Implement story-telling aspect of website</li> </ul>	<ul style="list-style-type: none"> <li>• Incorporate feedback on specific graphs</li> <li>• Continue to expand possible number of views of data as much as possible</li> <li>• Add levels of interactivity across graphs</li> </ul>
4/30 – 5/2	Finalize the model Finalize the supporting documentation Finalize the website Create the screencast		
5/3 – 5/5	Madly <i>really</i> finalize any last details		

## Additional Sketches

