Browser Feature Usage on the Modern Web

Summary

- Analysis of how frequently javascript features are used
  - "Feature" defined as browser capability accessed through JavaScript function or property.
- Considers only sites in Alexa top 10k
  - Represent ⅓ of total web page visits
- Over 50% of features are not used by sites in the Alexa 10k.

Affected by ad & tracking blocking
- Around 10% of browser features are used by websites but blocked by these extensions at least 90% of the time
- With blockers present, 83% of features are run on less than 1% of websites

Web Features
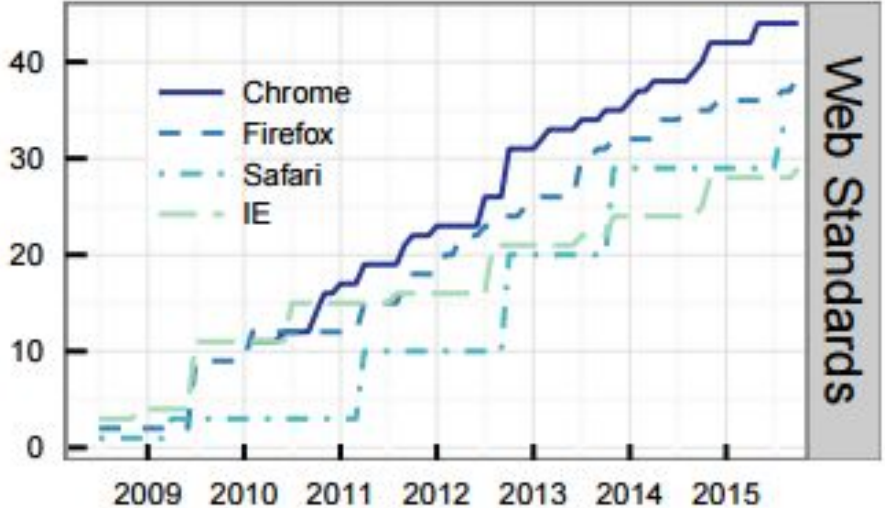
Browsers support different sets of features & standards

Data obtained from W3C documents

Feature Proliferation

- Browser makers are incentivized to keep adding new features in order to stay competitive.
- No motivation to remove old, unused features.
- Based on analysis of Firefox 46.0.1 WebIDL definitions, there are 1,392 total methods and properties.

Survey Methodology

- Visited each top 10k website 10 times
  - Used Firefox browser only.
  - 5 with unmodified browser.
  - 5 with popular tracking and ad-blocking extensions.
    - AdBlock Plus and Ghostery
- Recorded browser features that were used
  - Used a custom Firefox browser extension to record each usage.
- Ran automated crawl through sites using "monkey testing"
  - Interacted with random elements on page
  - Validation shows that after 5 rounds no new standards were discovered

Domains measured	9,733
Total website interaction time	480 days
Web pages visited	2,240,484
Feature invocations recorded	21,511,926,733

Results: Browser Standards

- Standard: Set of related browser features
    - Charted instead of features for simplicity
- Standard popularity distribution
    - 6 of 75 standards used on over 90% of websites
    - 28 standards used on 1% or fewer
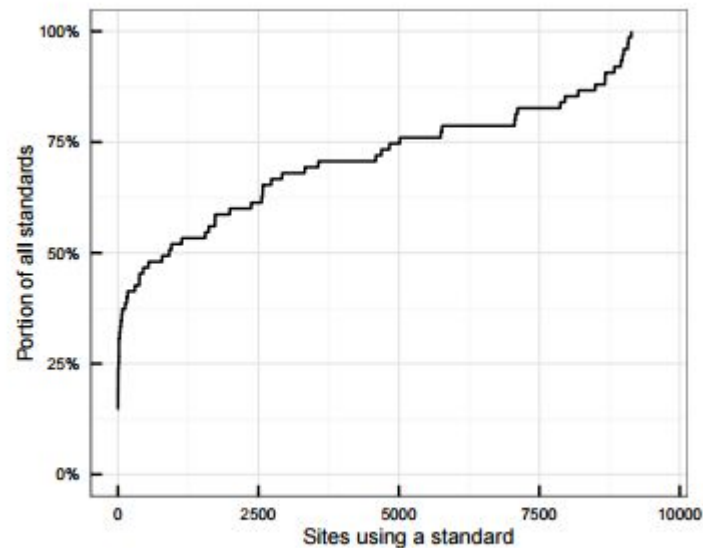   - 11 not used at all

Figure 3: Cumulative distribution of standard popularity within the Alexa 10k.

Standard Popularity versus Site Popularity

- Some standards are used by fewer, but more popular sites
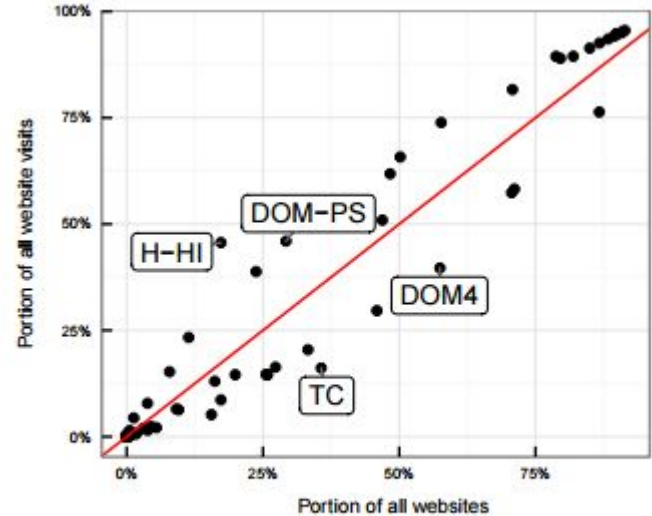- Trend is mostly linear

Figure 5: Comparison of percentage of sites using a standard versus percentage of web traffic using a standard.

Standard Creation Date

- Old standards tend to be either very popular or very unpopular (AJAX vs HTML: Plugins)
- New standards see a lot more variance, with some being highly popular and many more being unpopular
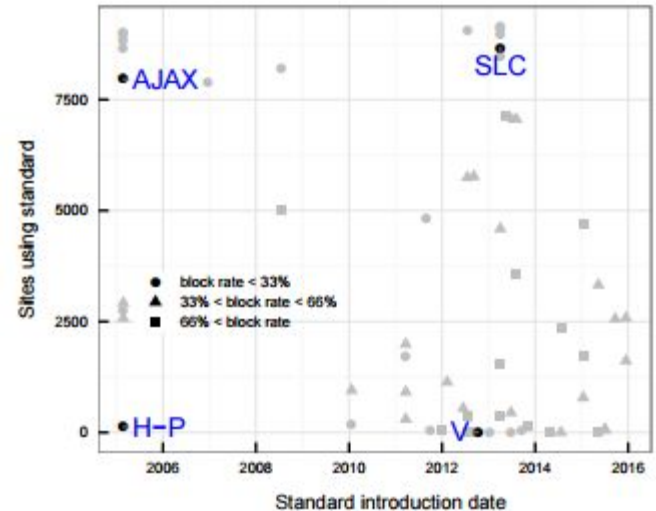

Figure 6: Comparison of a standard's availability date, and its popularity.

Ad and Tracking Blocking

- Standards that are blocked by ad blocking extensions are more likely to be blocked by tracking blocking extensions.
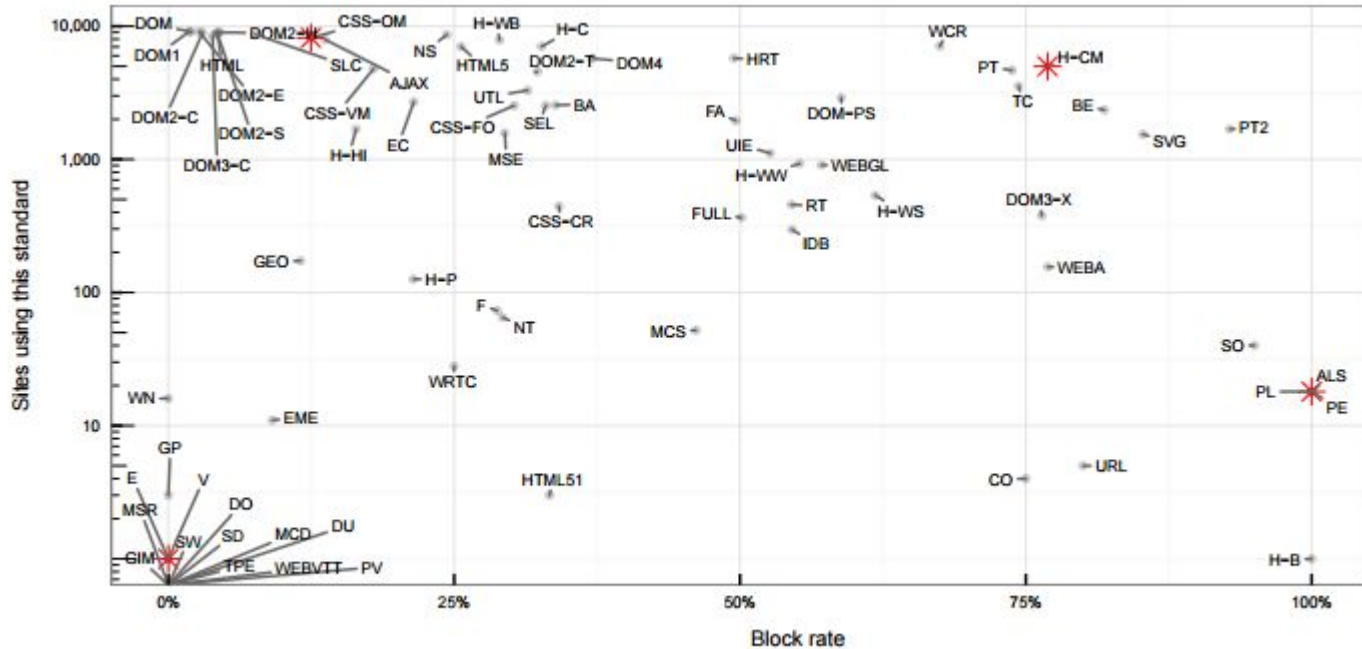

Figure 4: Popularity of standards versus their block rate, on a log scale.

Browser Vulnerabilities

- Many standards have known vulnerabilities (up to 15)
  - Popular standards like HTML: Canvas, HTML 5, and XMLHttpRequest are used on over 70% of sites but have 15, 10, and 8 vulnerabilities respectively.
- Just because a site is using a vulnerable standard doesn't mean that it is using vulnerable features
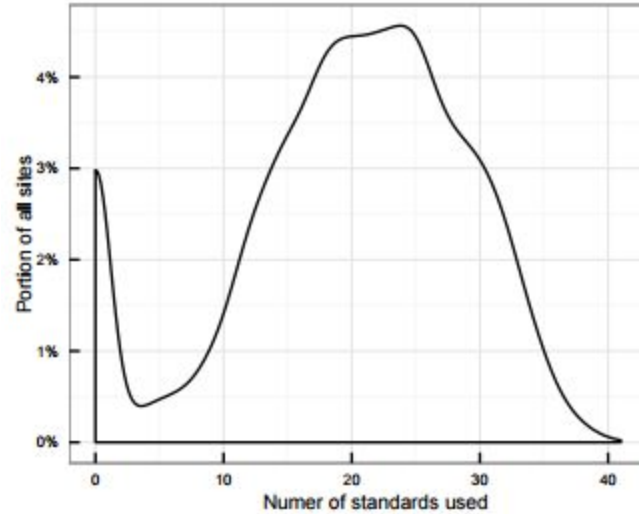
Site Complexity

Figure 8: Probability density function of number of standards used by sites in the Alexa 10k.