Mining Questions Asked by Web Developers

Kartik Bajaj, Karthik Pattabiraman, Ali Mesbah

Electrical and Computer Engineering
University of British Columbia
Vancouver, BC, Canada
{kbajaj, karthikp, amesbah}@ece.ubc.ca

MSR 2014

Problem Statement

- Web applications are becoming popular over time
  - [Jazayeri et al., FOSE 2007]

- Complex integration of JavaScript, CSS, and HTML

- JavaScript code often experiences errors
  - [Ocariza et al., ISSRE 2011]

Existing Work

- Static Checkers

- Dynamic Checkers

- Console Messages

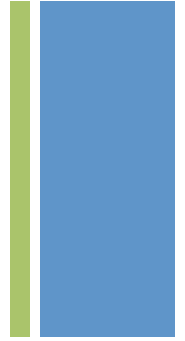- Bug Reports

Existing Work

- Static Checkers

- Dynamic Checkers
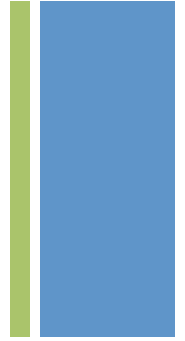
- Console Messages

- Bug Reports

No attempt to understand developer's misconception

Goal

- Understand the common challenges and/or misconceptions among web developers

Background

- **Stack Overflow**
  - QA website for programmers
  - Started in 2008
  - 4,125,638 questions asked from Jan'09 to Dec'12
  - 500,000+ questions related to web development

- **Questions directly asked/answered by developers**
  - Followed by discussion in comments
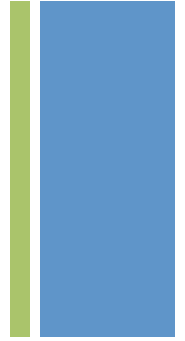
Background

- Stack Overflow
  - QA website for programmers
  - Started in 2008
  - 4,125,638 questions asked from Jan'09 to Dec'12
  - 500,000+ questions related to web development

- Questions directly asked/answered by developers
  - Followed by discussion in comments

StackOverflow provides sufficient data for analysis

Straw man Approach

- Categorization based on tags attached to each question

- Can provide an overview of major topics of discussion

Straw man Approach

- Categorization based on tags attached to each question

- Can provide an overview of major topics of discussion

How appropriate are tags?

Tags

- What is known to developer while tagging?
  - Platform
  - Target Device

- What is unknown to developer?
  - Possible ways to solve the problem
  - Correct topic of discussion

Example

UIWebView intermittently denied access to html5 database storage

▲
5
▼

☆
1

I have an iphone app that enables users to login via a native `UIView` on ios, that then fires up a `UIWebView` to display the main content. The webapp uses database storage to retain some of the content locally.

However, every now and then, the webapp will fail to load when attempting to access the database with the following message appearing in the logs.

..... sandboxd[3203] : APPNAME(3201) deny file-write-create /Databases.db
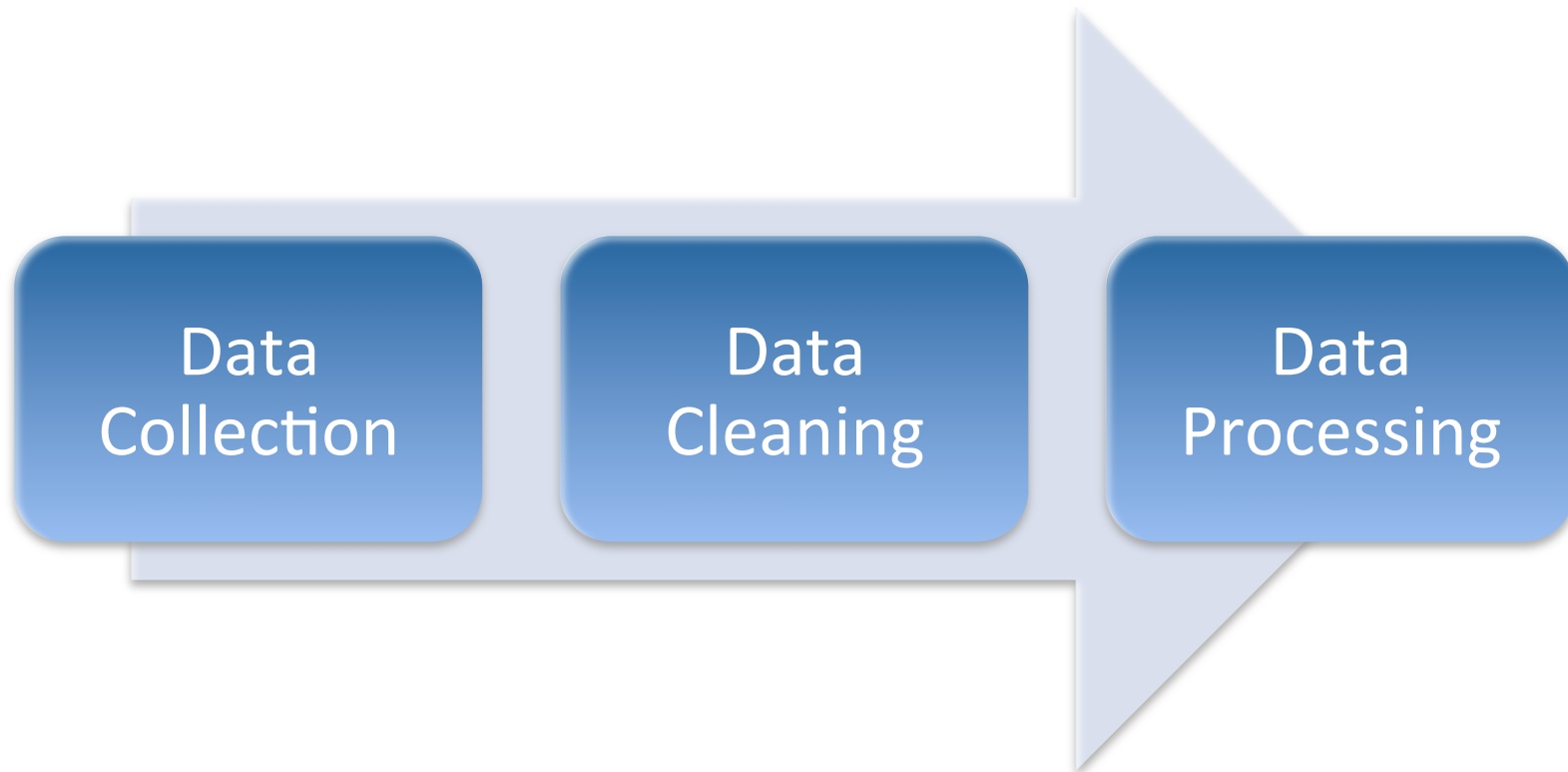
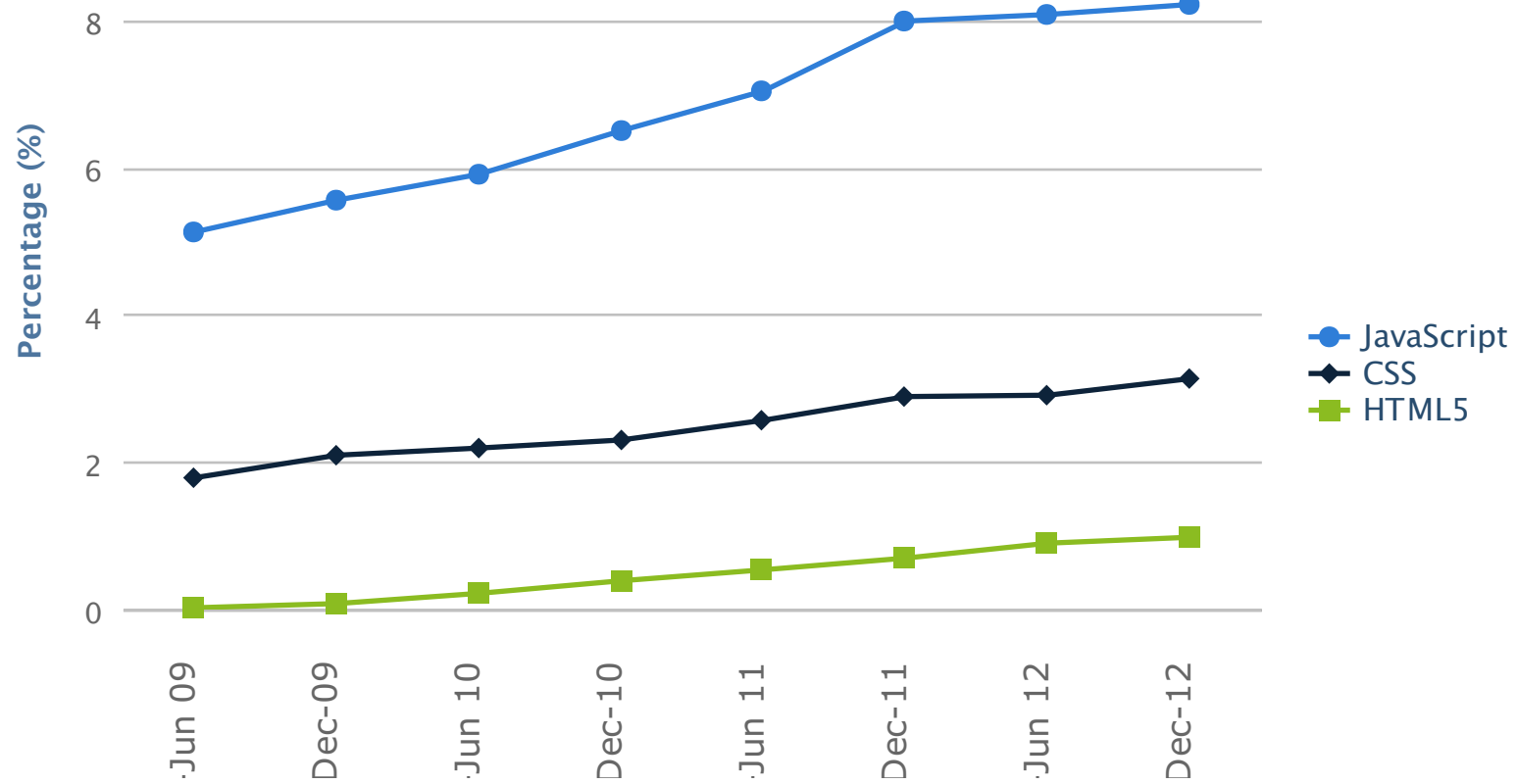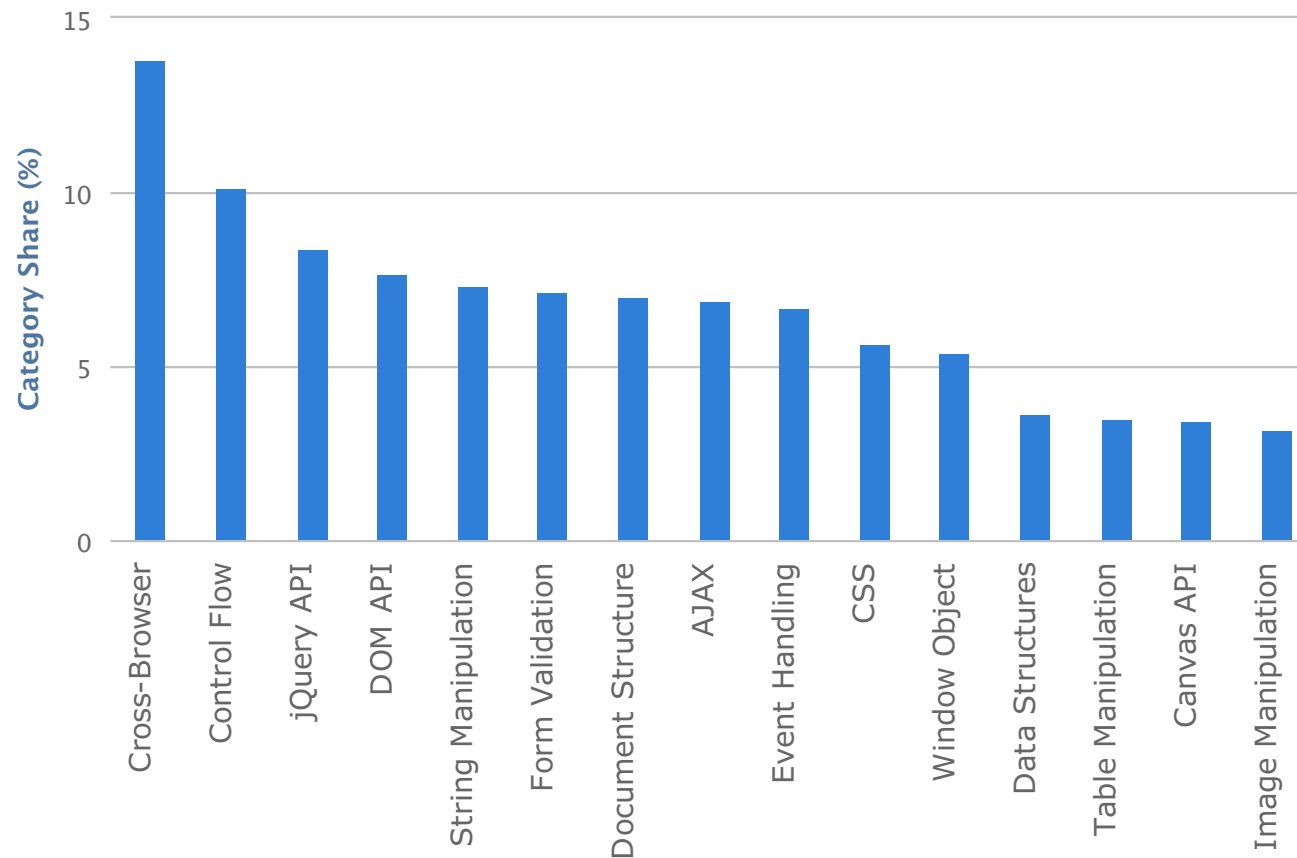| iphone | ios | html5 | uiwebview | local-storage |

Natural Language Analysis

■ Analyzing the text provided in the questions and answers provided by developers

Data Collection

Data Cleaning

Data Processing

Datasets

RQ1: Categorization of topics of discussion

RQ1: Categorization of topics of discussion

15

Cross Browser related discussions have gained maximum attention from web developers.

0

Cross-Browser
Control Flow
jQuery API
DOM API
String Manipulation
Form Validation
Document Structure
AJAX
Event Handling
CSS
Window Object
Data Structures
Table Manipulation
Canvas API
Image Manipulation

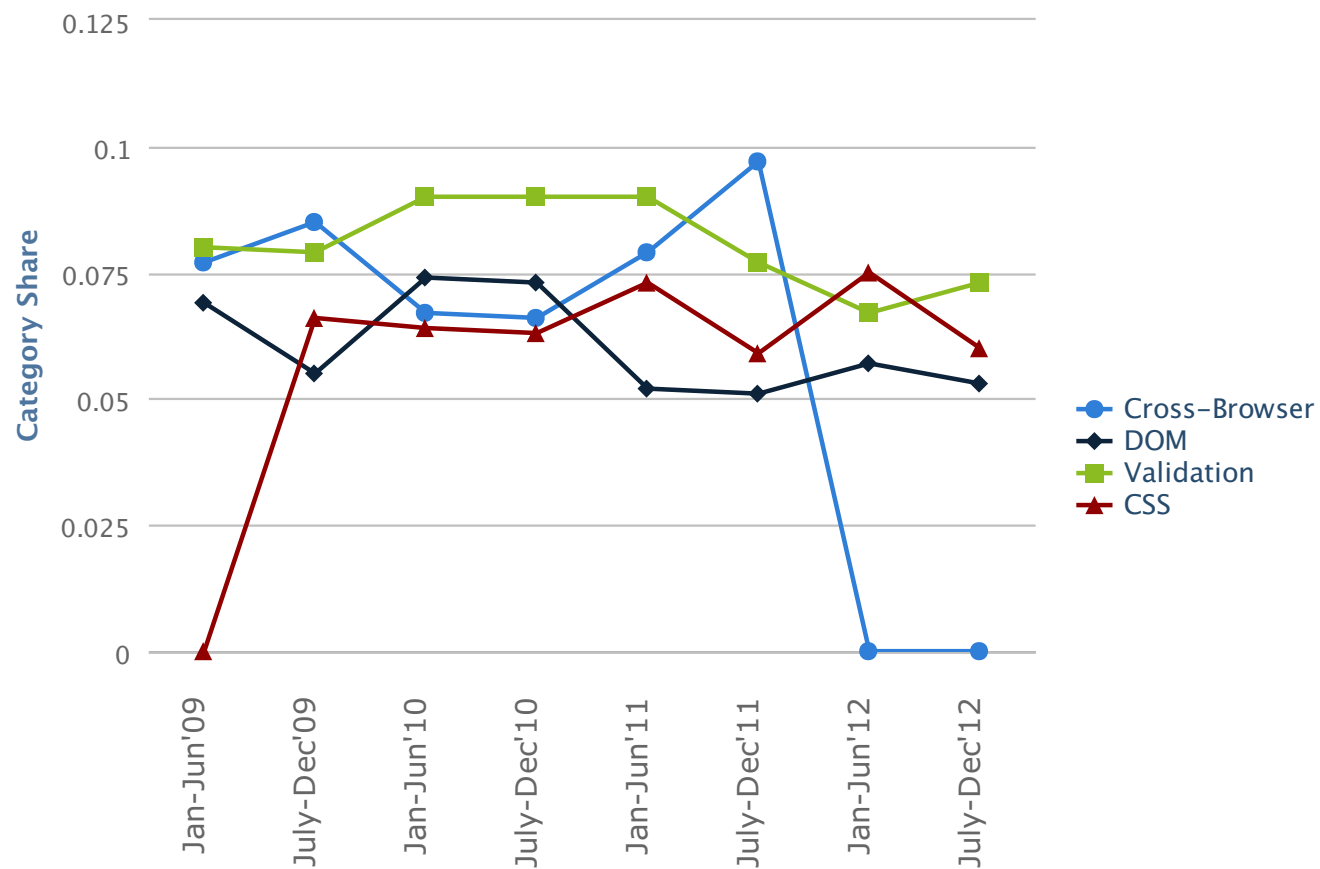# RQ2: Hot topics of discussion

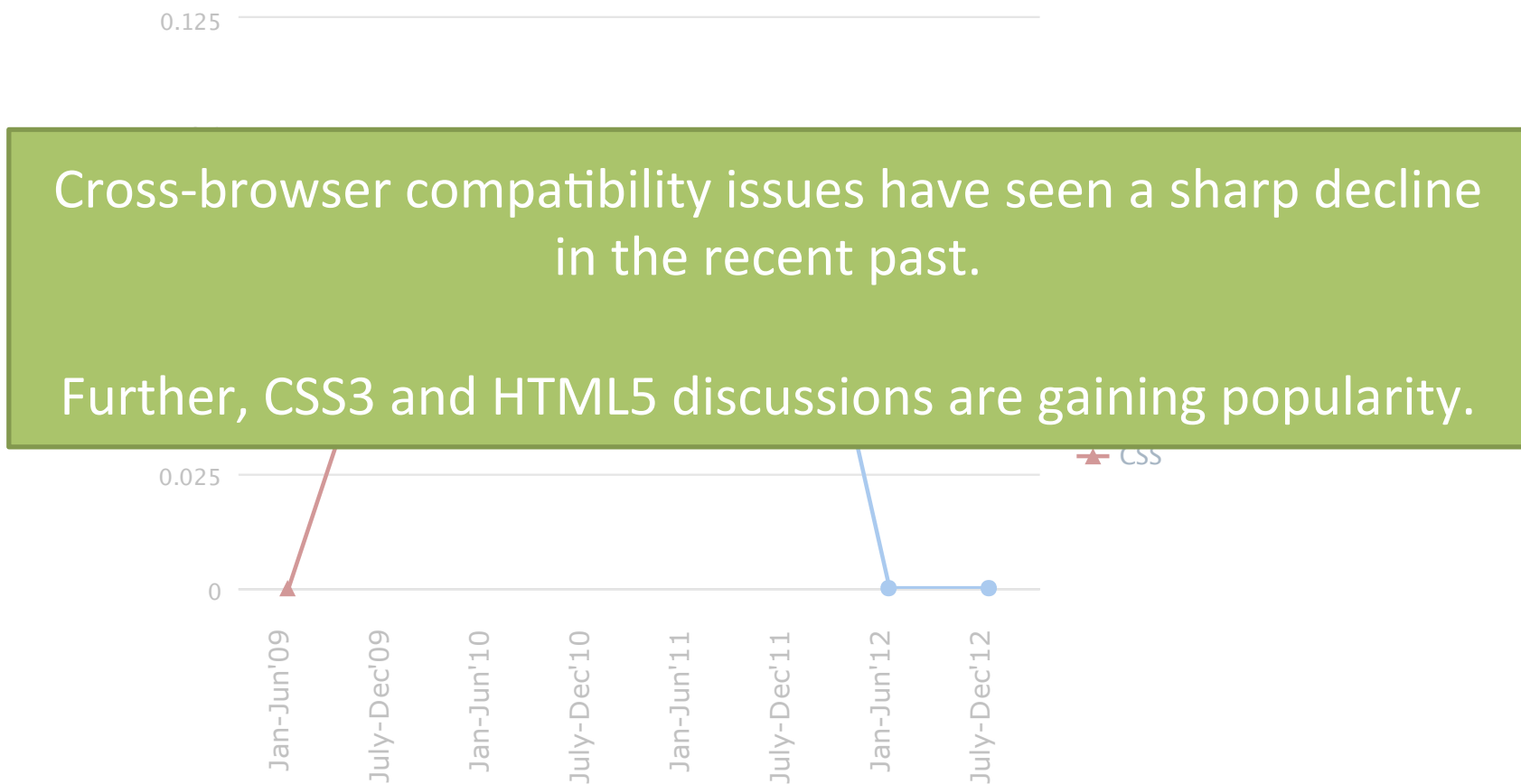| Technology	Hot Topics
JavaScript	**Document Structure, File Handling,** Cross-Browser, jQuery, DOM
HTML5	**Media**, Browser Support, **HTML5 Elements**, Canvas API, **Offline Web**
CSS	**CSS3**, **Fonts**, JavaScript, Box-Model, Layout

RQ2: Hot topics of discussion

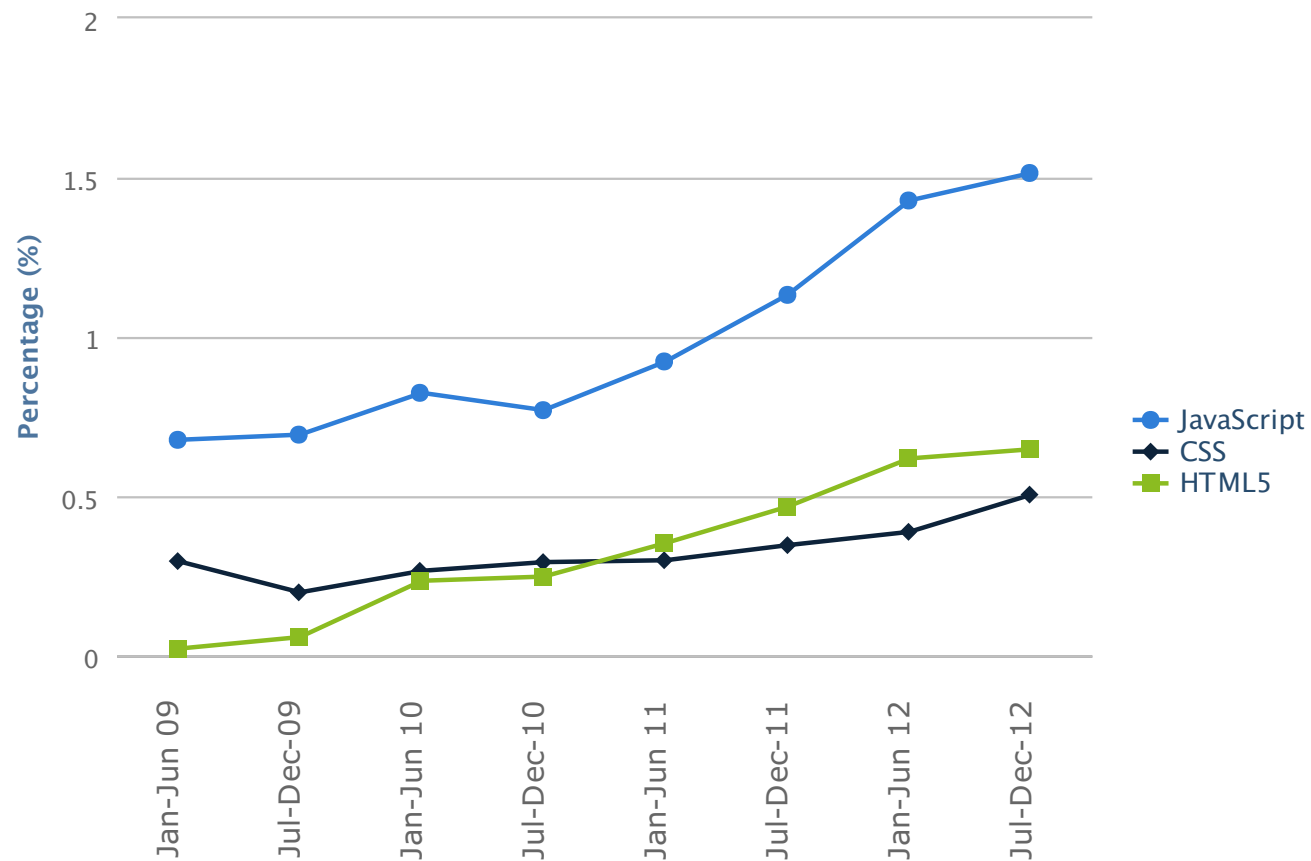View counts provide a hint towards recurrent issues faced by web developers.
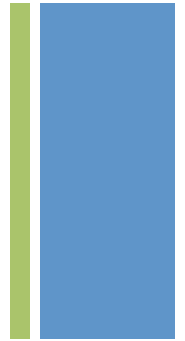
RQ3: Temporal trends over time

RQ3: Temporal trends over time

0.125

Cross-browser compatibility issues have seen a sharp decline
in the recent past.

Further, CSS3 and HTML5 discussions are gaining popularity.

0.025   CSS

0

Jan-Jun'09   July-Dec'09   Jan-Jun'10   July-Dec'10   Jan-Jun'11   July-Dec'11   Jan-Jun'12   July-Dec'12
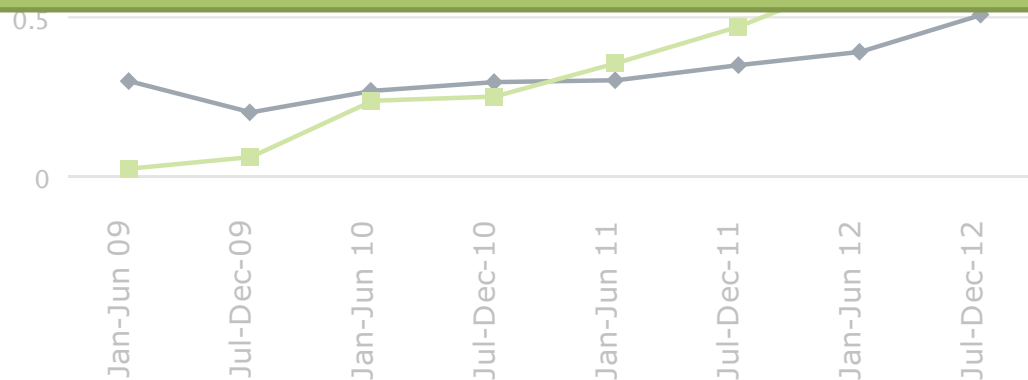
RQ4: Prevalence of web in mobile development

RQ4: Prevalence of web in mobile development

2

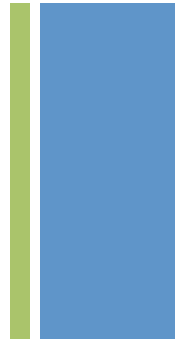Discussions related to Mobile development are seeing an increasing share of web technologies.

0.5

0

Jan-Jun 09　Jul-Dec-09　Jan-Jun 10　Jul-Dec-10　Jan-Jun 11　Jul-Dec-11　Jan-Jun 12　Jul-Dec-12

Popularity

■ Questions that gained maximum attention from the community

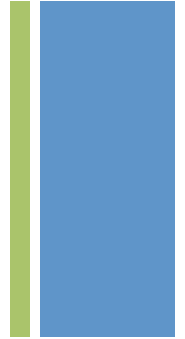■ Post Score
  ■ Up votes
  ■ Down votes

Popularity

- Questions that gained maximum attention from the community

- Post Score
  - Up votes
  - Down votes

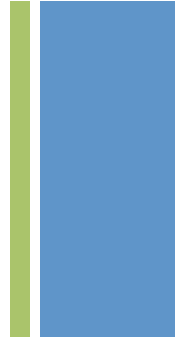Is post score accurate representation of user participation?

StackOverflow Post Score

■ Missing details

- How many users up voted or down voted the question?

- How many users tried to answer the question?

- Is there an accepted answer for this question?

- How many users are following the question?

- How many users are participating in the discussion?

StackOverflow Post Score

- Missing details
  - How many users up voted or down voted the question?

We devised a new metric to rank Stack Overflow questions based on these factors.

RQ5: Technical challenges

I'd like to store a JavaScript object in HTML5 `localStorage`, but my object is apparently being converted to a string.

I can store and retrieve primitive JavaScript types and arrays using `localStorage`, but objects don't seem to work. Should they?

I have noticed I am getting a "CSS Explosion". It is becoming difficult for me to decide how to best organize and abstract data within the CSS file.

RQ5: Technical challenges

Programmers get confused about new features, suggesting that the available API resources for these features is far from ideal.

Maintaining web code, such as CSS, is complex without proper tool support.

Implications

- **Finding 1, 3 (Categorization and temporal trends)**

  - *Developers* can shift their focus away from cross browser issues.

- **Finding 4 (Prevalence of mobile applications)**

  - Can guide *tool developers* to build better mobile development tools.

- **Finding 5 (Technical Challenges)**

  - Can guide *standardization communities* to focus on areas that need improvement.

Thank You!!

Contributions

1. Categorized web related discussions.
2. Identified temporal trends.
3. Devised a metric to rank StackOverflow questions.

Analyzed Data: http://www.ece.ubc.ca/~kbajaj/so/data.zip

Mining Questions Asked by Web Developers

Kartik Bajaj

Karthik Pattabiraman

Ali Mesbah

Electrical and Computer Engineering
University of British Columbia
Vancouver, BC, Canada
{kbajaj, karthikp, amesbah}@ece.ubc.ca

and JavaScript to interact with that layout. JavaS

functionality of a web appl