

Adam Arreguin,
Adrian Gutierrez,
Joel Staggs,
Kenny Taylor

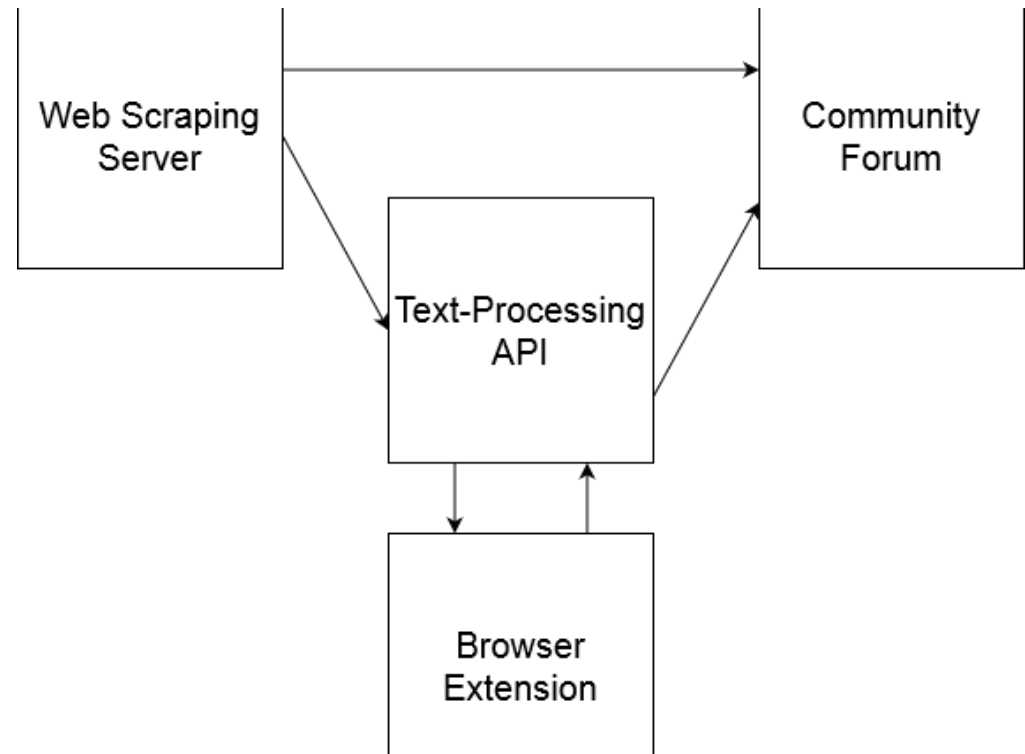
Web Crawling and Community Review to Prevent Misleading Links

The Problem

- Low quality webpages oversell content, use manipulative tactics to generate ad revenue via site visits.
- This "clickbait" journalism often results in a poor experience for the user.
- Our goal is to warn users of and lessen the presence of this type of article.

Project Overview

- Our project is composed of four main components
- A text-classification API to score webpages based on title content
- A browser extension that allows user to intuitively view site scores
- A web scraping server to retrieve large quantities of article assessments
- A community forum web application with a variety of services



- Uses MNB Classifier to assess articles using conditional probabilities
- This assessment is sent to other components upon completion

Text- Classification API

Browser Extension

- Users can hover over a link to display information about the link's content
- Upon hovering, the Text-Classification is automatically invoked, providing the user with a score
- Users can dock the text-box, click through it, or disable it entirely.

Explore Article ▾

moz://a HACKS

Search Mozilla Hacks

Learn web technology at "sofa school"

By **Val Grimm**

Posted on March 25, 2020 in [CSS](#), [Featured Article](#), [Games](#), [Learning](#), and [WebVR](#)

Lots of kids around the world are learning from home right now. In this post, I introduce free resources based on web technologies that will help them explore and learn from the safety of their living rooms. VR headsets and high-end graphics cards aren't necessary. Really, all you need is a web browser!

Create a secret hideout

[Hubs by Mozilla](#) lets you share a virtual room with friends

HACKS Blog

[read more at hacks.mozilla.org](#) →

Innovating on Web Monetization: Coil and Fire

An in-depth introduction to web monetization with Coil. Coil is really on advertising or personal data collection to support digital Reality is rolling out a Coil-based experiment for creators and participate or learn more about how ...

Posted March 31, 2020 by Anselm Hook

Learn web technology at "sofa school"

Page Reputation: Good

Site Score: 93.93%

Learn web technology at "sofa school"

Here are some web development learning resources for kids in your browser. This collection looks at social VR experiences using games that teach Grid and Flexbox, and WebXR field trips the

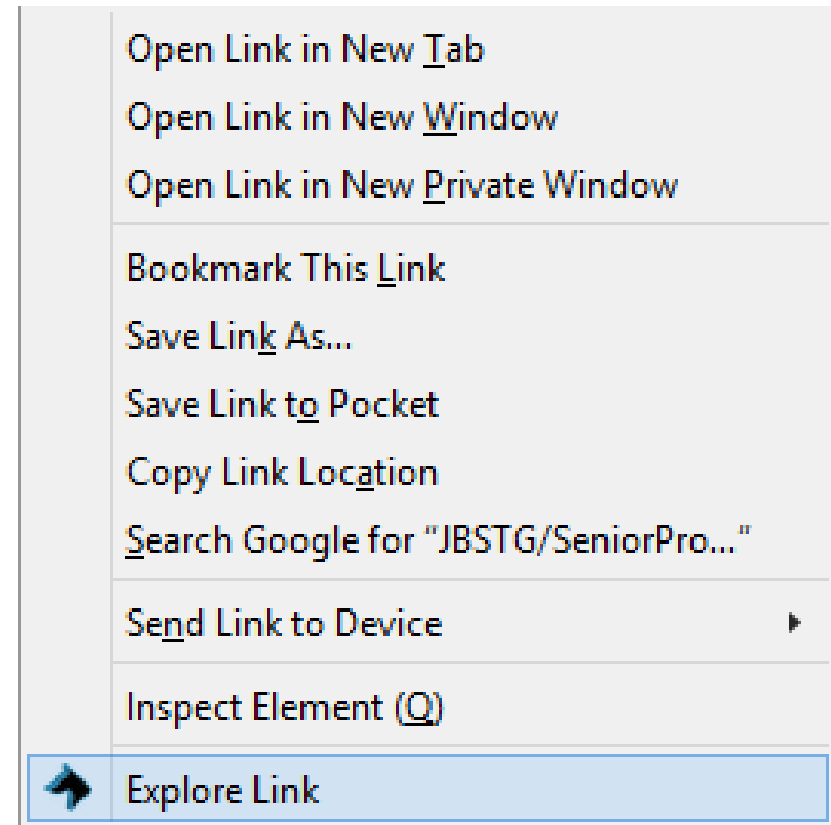
Posted March 25, 2020 by Val Grimm

Security means more with Firefox 74

The release of Firefox 74 is focused on security enhancement Resource-Policy header, and removal of TLS 1.0/1.1 support. text property features, the JS optional chaining operator, and features, along with the usual wealth of DevTools enhance

Browser Extension (Cont.)

- By clicking "Explore Link" within our extension's context menu users can:
 - Read and write comments using a remote WebSocket server.
 - Preview the site's content without needing to visit it.



Automated Web-Scraping Server

- Users retrieve or request articles scraped by the server
- Domain names can be searched from the website to display scraped scores

- Our community forum web application allows users to interact with one another along with articles and domains scored and ranked in our database
- Users also can use a scrape functionality on the website similar to the browser extension which scores a URL and stores it in the database

Community Forum Web Application

Challenges

- Accommodating various website configurations
- Balancing code changes across many interlinked components
- Tracking feature progress across components

Conclusion

- Our implementation of the **Text-Classification API**, **Automated Web-Scaping Server**, **Browser Extension** and **Community Forum Web-App** gave us an invaluable breadth of experience collaborating on and maintaining a large, distributed software platform.
- We look forward to using these skills in the future.