Exploring the potential of the Web Speech API in

KARAOKE

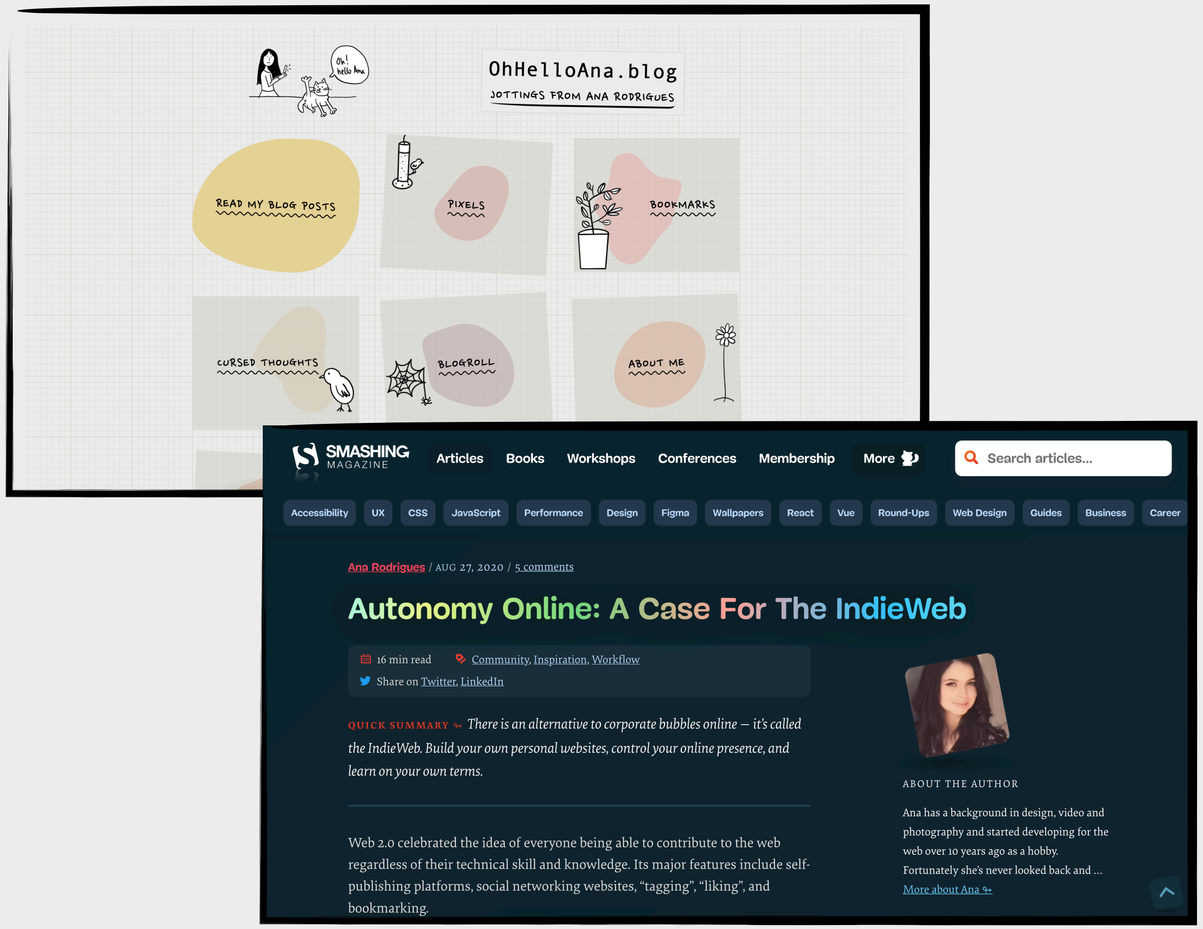
FFConf 2023

Oh Hello!

I work as a front-end developer at hactar.is.

I spend most of my free time on my blog.

I participate in the IndieWeb community.

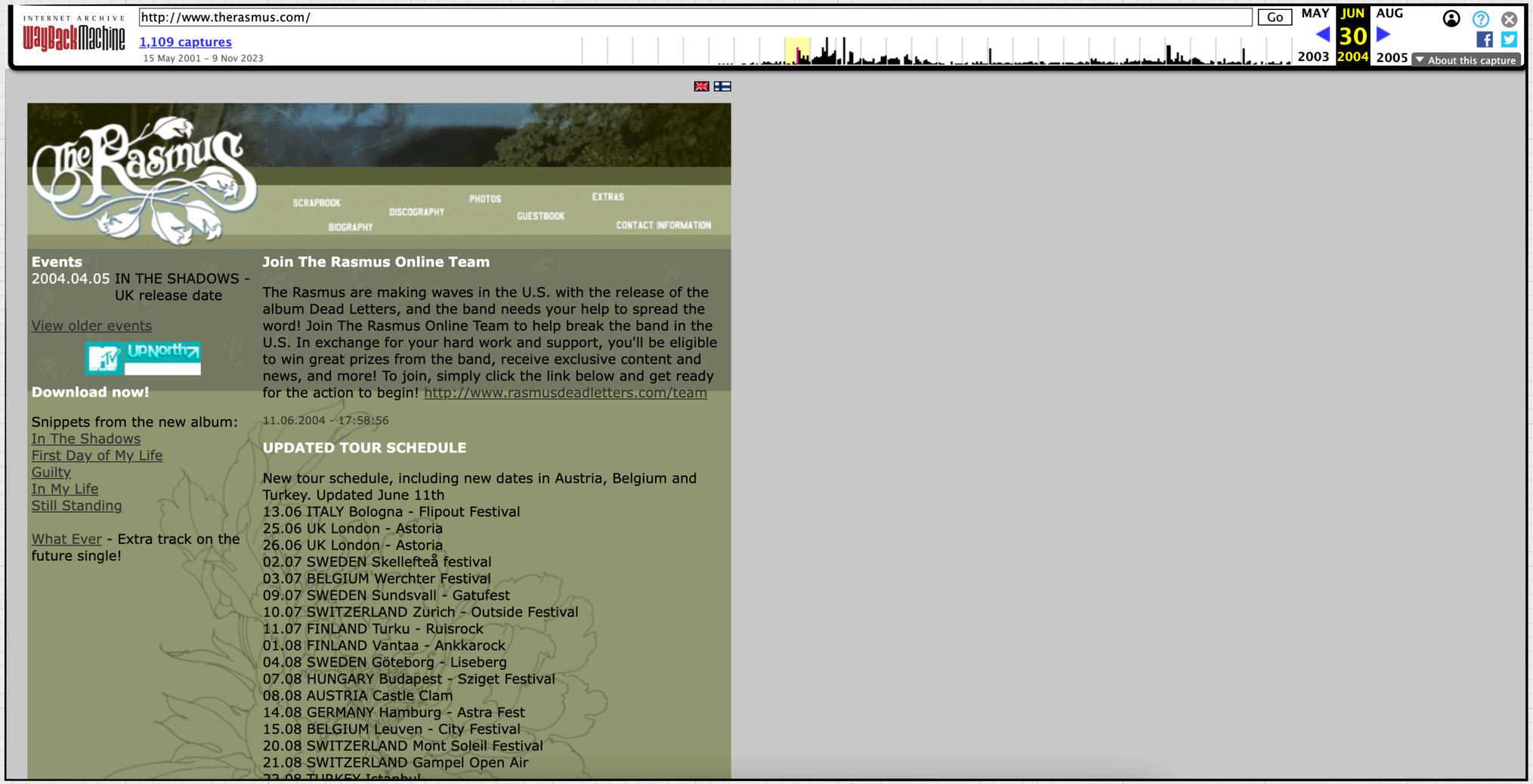


Oh Hello!

I do not represent any browser vendor.

I haven’t been to Karaoke since before the pandemic.

I am a big fan of The Rasmus (not a stan).



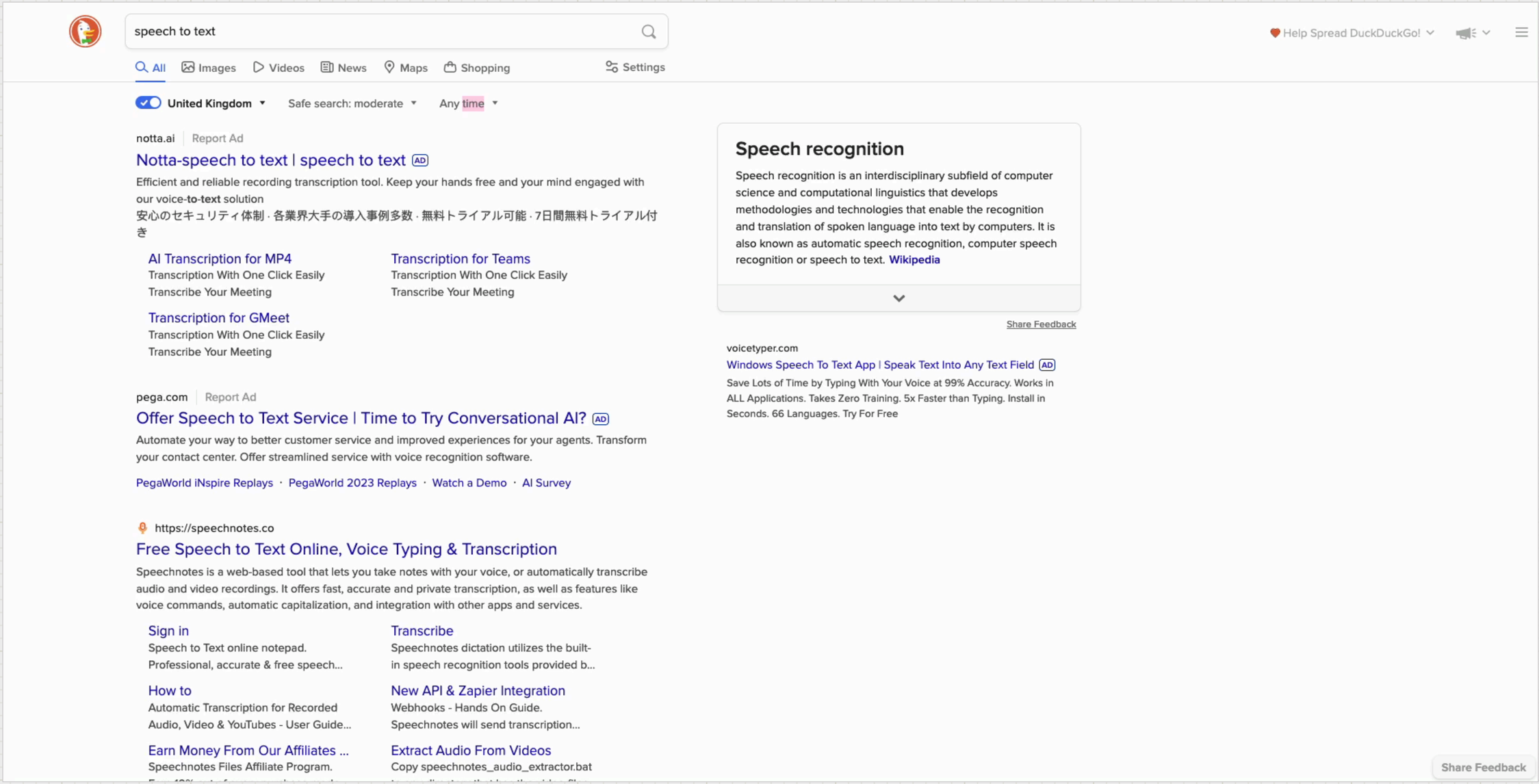
https://web.archive.org/web/20040630073555/http://www.therasmus.com/

I had a problem that I needed to fix

Whenever I went karaoke,
they only had “In the Shadows” available.

What if we had more than just lyrics on a screen?

What if we could gamify the experience?


speech to text
Help Spread DuckDuckGo!
All
Images
Videos
News
Maps
Shopping
Settings
United Kingdom
Safe search: moderate
Any time
notta.ai
Report Ad
Notta-speech to text | speech to text
AD
Efficient and reliable recording transcription tool. Keep your hands free and your mind engaged with our voice-to-text solution
安心のセキュリティ体制 · 各業界大手の導入事例多数 · 無料トライアル可能 · 7日間無料トライアル付き
AI Transcription for MP4
Transcription With One Click Easily
Transcribe Your Meeting
Transcription for Teams
Transcription With One Click Easily
Transcribe Your Meeting
Transcription for GMeet
Transcription With One Click Easily
Transcribe Your Meeting
pega.com
Report Ad
Offer Speech to Text Service | Time to Try Conversational AI?
AD
Automate your way to better customer service and improved experiences for your agents. Transform your contact center. Offer streamlined service with voice recognition software.
PegaWorld iNspire Replays · PegaWorld 2023 Replays · Watch a Demo · AI Survey
https://speechnotes.co
Free Speech to Text Online, Voice Typing & Transcription
Speechnotes is a web-based tool that lets you take notes with your voice, or automatically transcribe audio and video recordings. It offers fast, accurate and private transcription, as well as features like voice commands, automatic capitalization, and integration with other apps and services.
Sign in
Speech to Text online notepad. Professional, accurate & free speech...
Transcribe
Speechnotes dictation utilizes the built-in speech recognition tools provided b...
How to
Automatic Transcription for Recorded Audio, Video & YouTubes - User Guide...
New API & Zapier Integration
Webhooks - Hands On Guide. Speechnotes will send transcription...
Earn Money From Our Affiliates ...
Speechnotes Files Affiliate Program.
Extract Audio From Videos
Copy speechnotes_audio_extractor.bat
Speech recognition
Speech recognition is an interdisciplinary subfield of computer science and computational linguistics that develops methodologies and technologies that enable the recognition and translation of spoken language into text by computers. It is also known as automatic speech recognition, computer speech recognition or speech to text. Wikipedia
Share Feedback
voicetyper.com
Windows Speech To Text App | Speak Text Into Any Text Field
AD
Save Lots of Time by Typing With Your Voice at 99% Accuracy. Works in ALL Applications. Takes Zero Training. 5x Faster than Typing. Install in Seconds. 66 Languages. Try For Free
Share Feedback

So let's talk about the Web Speech API

In particular, the speech recognition.

W3C Editor's Draft

Speech Input API Specification

Editor's Draft 18 October 2010

Latest Editor's Draft:
http://dev.w3.org/...

Editors:
Satish Sampath, Google Inc.
Bjorn Bringert, Google Inc.

Copyright © 2010 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark and document use rules apply.

Abstract

This specification extends HTML and defines an API that provides speech recognition and input to web pages.

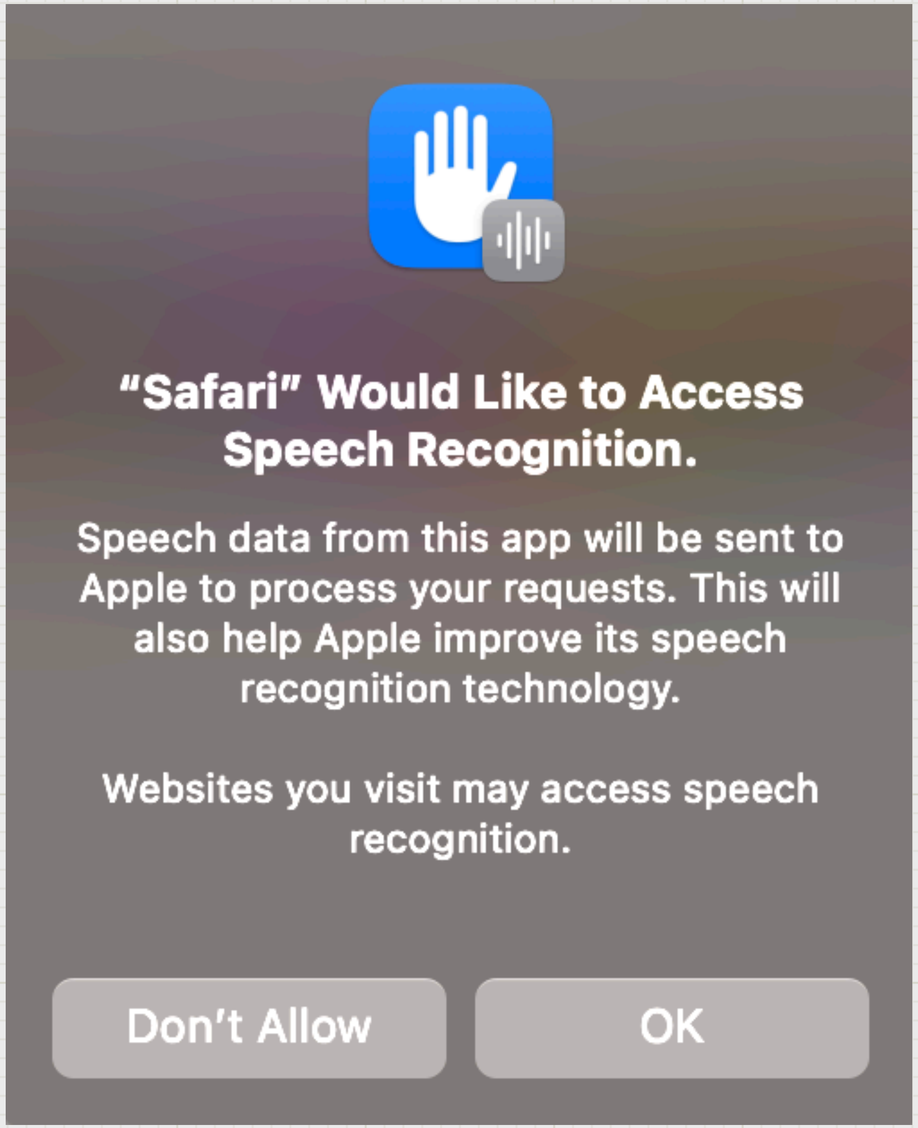
Status of This Document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the lates technical report can be found in the W3C technical reports index at http://www.w3.org/TR/.

This document is an API proposal from Google Inc. to the HTML Speech Incubator Group. If you wish to make comments regarding this document, please send them to public-x (subscribe, archives).

All feedback is welcome.

https://www.w3.org/2005/Incubator/htmlspeech/2010/10/google-api-draft.html


"Safari" Would Like to Access Speech Recognition.
Speech data from this app will be sent to Apple to process your requests. This will also help Apple improve its speech recognition technology.
Websites you visit may access speech recognition.
Don't Allow
OK

> **Note:** On some browsers, like Chrome, using Speech Recognition on a web page involves a server-based recognition engine. Your audio is sent to a web service for recognition processing, so it won't work offline.

https://developer.mozilla.org/en-US/docs/Web/API/SpeechRecognition

Filter

Web Speech API

▾ Guides

Using the Web Speech API

▾ Interfaces

SpeechGrammar

SpeechGrammarList

SpeechRecognition

SpeechRecognitionAlternative

SpeechRecognitionErrorEvent

SpeechRecognitionEvent

SpeechRecognitionResult

SpeechRecognitionResultList

SpeechSynthesis

SpeechSynthesisErrorEvent

SpeechSynthesisEvent

SpeechSynthesisUtterance

SpeechSynthesisVoice

api.SpeechRecognition

Report problems with this compatibility data on GitHub

	Chrome	Edge	Firefox	Opera	Safari	Chrome Android	Firefox for Android	Opera Android	Safari on iOS	Samsung Internet	WebView Android
SpeechRecognition	✓ 33	✓ 79	✗ No	✓ 20	✓ 14.1	✓ 33	✗ No	✓ 20	✓ 14.5	✓ 2.0	✓ 4.4.3
SpeechRecognition() constructor	✓ 33	✓ 79	✗ No	✗ No	✓ 14.1	✓ 33	✗ No	✗ No	✓ 14.5	✓ 2.0	✓ 37
abort	✓ 33	✓ 79	✗ No	✓ 20	✓ 14.1	✓ 33	✗ No	✓ 20	✓ 14.5	✓ 2.0	✓ 4.4.3
audioend event	✓ 33	✓ 79	✗ No	✓ 20	✓ 14.1	✓ 33	✗ No	✓ 20	✓ 14.5	✓ 2.0	✓ 4.4.3
audiostart event	✓ 33	✓ 79	✗ No	✓ 20	✓ 14.1	✓ 33	✗ No	✓ 20	✓ 14.5	✓ 2.0	✓ 4.4.3
continuous	✓ 33	✓ 79	✗ No	✓ 20	✓ 14.1	✓ 33	✗ No	✓ 20	✓ 14.5	✓ 2.0	✓ 4.4.3
end event	✓ 33	✓ 79	✗ No	✓ 20	✓ 14.1	✓ 33	✗ No	✓ 20	✓ 14.5	✓ 2.0	✓ 4.4.3
	✓	✓	✗	✓	✓	✓	✗	✓	✓	✓	✓

In this article

Web Speech Concepts and Usage

Web Speech API Interfaces

Errors

Examples

Specifications

Browser compatibility

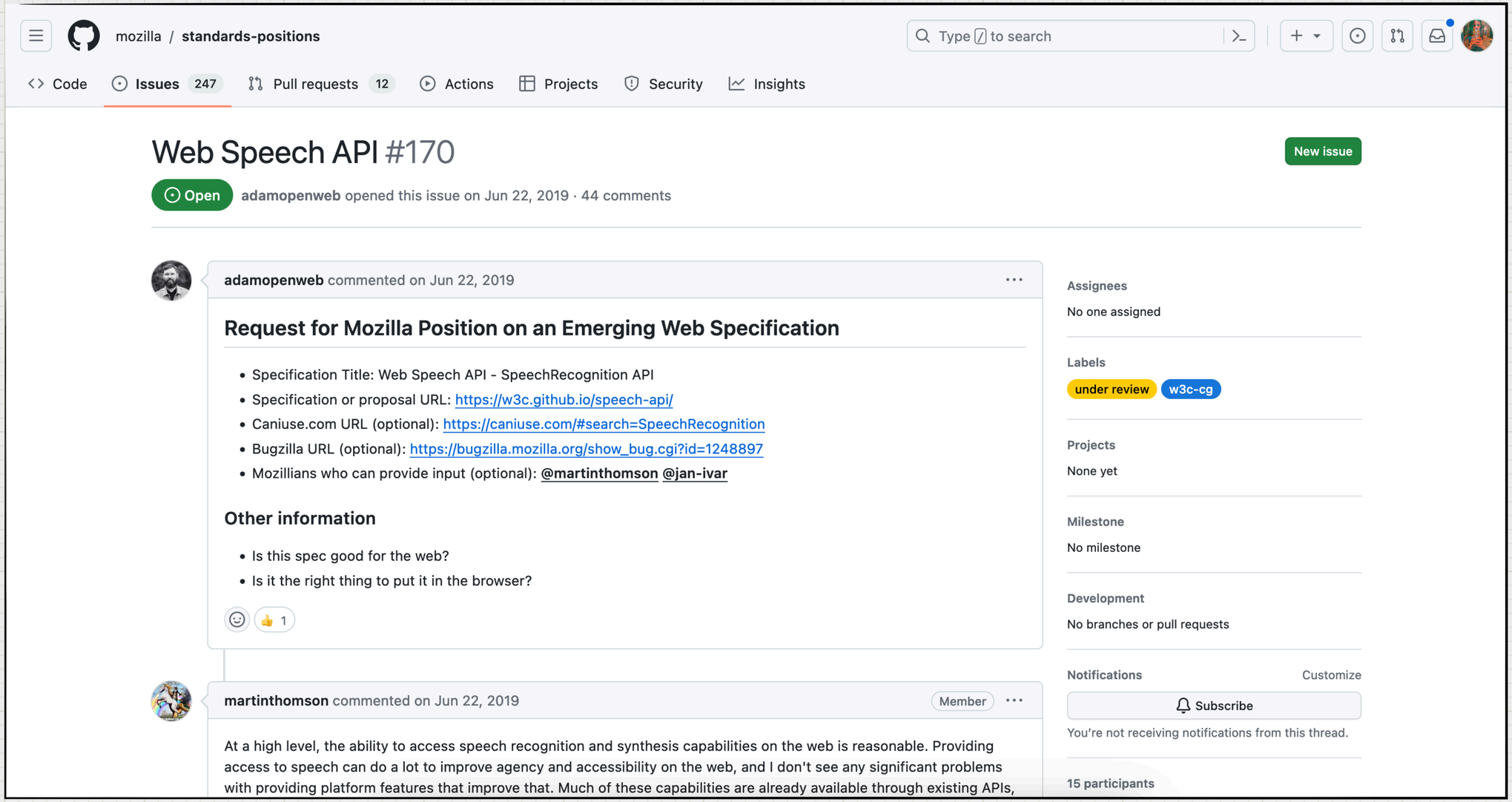
See also

GitLab

Simplify software development with the One DevOps Platform. Start your free 30 day trial today!

Mozilla ads

Don't want to see ads?



https://github.com/mozilla/standards-positions/issues/170

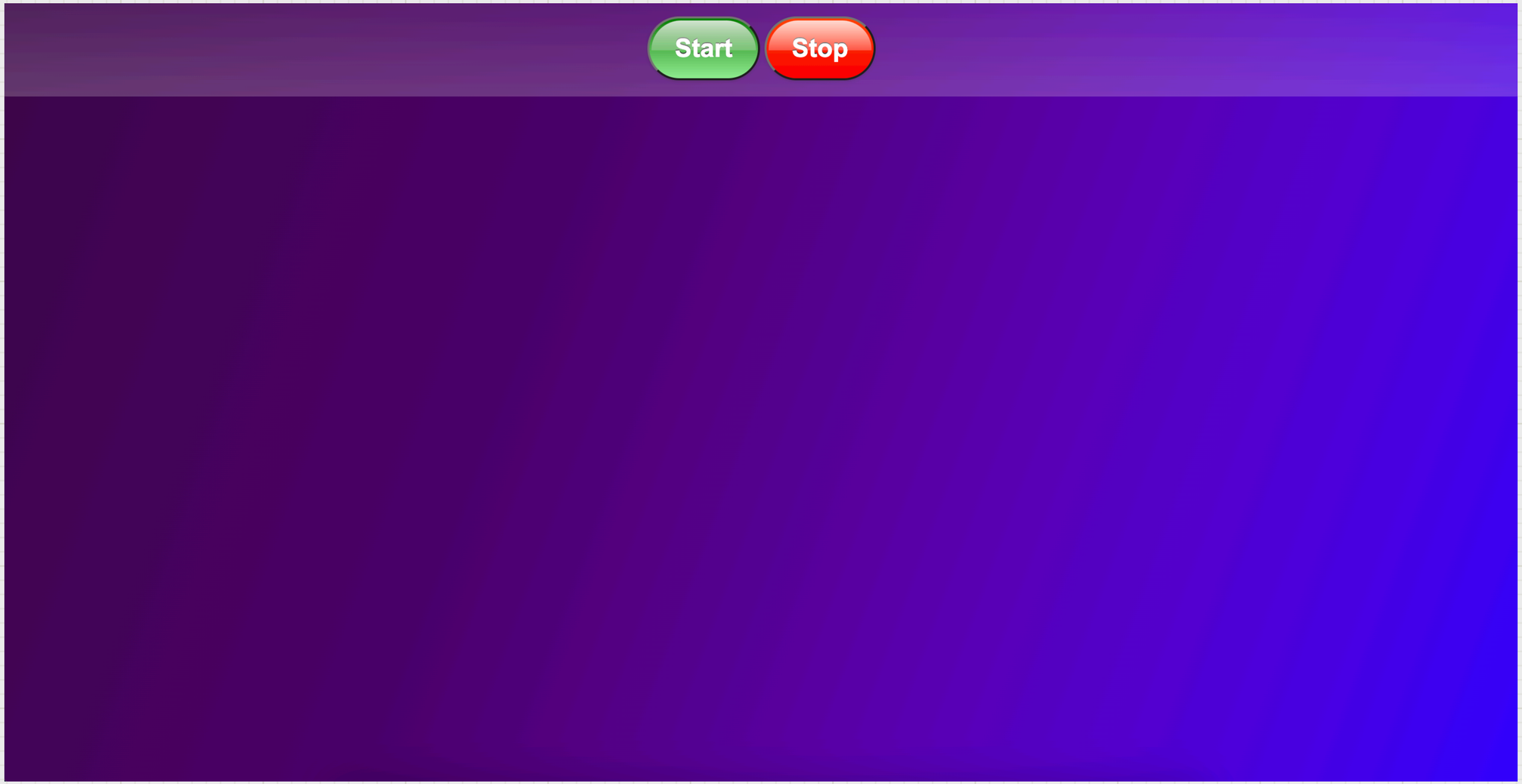
Demo

```
<div class="controls">
    <button id="play">Start</button>
    <button id="stop">Stop</button>
</div>

                                                    play>

<script type="text/javascript" src="karaoke.js"></script>
```

```
<div class="controls">
    <button id="play">Start</button>
    <button id="stop">Stop</button>
</div>

<div class="realTimeTranscription">
</div>

                                                    play>

<script type="text/javascript" src="karaoke.js"></script>
```

```
<div class="controls">
    <button id="play">Start</button>
    <button id="stop">Stop</button>
</div>

<div class="realTimeTranscription">
</div>

<audio id="myAudio" class="karaoke__audio" preload="auto" autoplay>
    <source src="song.mp3" type="audio/mpeg"/>
    <p>Your browser does not support the audio element.</p>
</audio>



<script type="text/javascript" src="karaoke.js"></script>
```

```
<div class="controls">
    <button id="play">Start</button>
    <button id="stop">Stop</button>
</div>

<div class="realTimeTranscription">
</div>

<audio id="myAudio" class="karaoke__audio" preload="auto" autoplay>
    <source src="song.mp3" type="audio/mpeg"/>
    <p>Your browser does not support the audio element.</p>
</audio>

<ul id="lyrics" class="karaoke__lyrics"></ul>
<script type="text/javascript" src="karaoke.js"></script>
```


Start
Stop

```
window.SpeechRecognition = window.SpeechRecognition || window.webkitSpeechRecognition;

const recognition = new SpeechRecognition();
recognition.interimResults = true;
recognition.lang = 'en-GB';
recognition.continuous = false;

let p = document.createElement('p');
const realTimeTranscription = document.querySelector('.realTimeTranscription');
realTimeTranscription.appendChild(p);

let transcript;

recognition.addEventListener('result', e => {
  transcript = Array.from(e.results)
    .map(result => result[0])
    .map(result => result.transcript)
    .join('');

    p.textContent = transcript;

    if (e.results[0].isFinal) {
      p = document.createElement('p');
      realTimeTranscription.appendChild(p);
    }
});

recognition.addEventListener('end', recognition.start);
```

```
window.SpeechRecognition = window.SpeechRecognition || window.webkitSpeechRecognition;

const recognition = new SpeechRecognition();
recognition.interimResults = true;
recognition.lang = 'en-GB';
recognition.continuous = false;

let p = document.createElement('p');
const realTimeTranscription = document.querySelector('.realTimeTranscription');
realTimeTranscription.appendChild(p);

let transcript;

recognition.addEventListener('result', e => {
  transcript = Array.from(e.results)
    .map(result => result[0])
    .map(result => result.transcript)
    .join('');

    p.textContent = transcript;

    if (e.results[0].isFinal) {
      p = document.createElement('p');
      realTimeTranscription.appendChild(p);
    }
});

recognition.addEventListener('end', recognition.start);
```

```
var lyrics = [
    { text: "Twinkle twinkle little star", start: "5", end: "11.2" },
    { text: "How I wonder what you are", start: "11.9", end: "17" },
    { text: "Up above the world so high", start: "17.7", end: "23.1" },
    { text: "Like a diamond in the sky", start: "23.3", end: "29.2" },
    { text: "Twinkle twinkle little star", start: "30", end: "35.1" },
    { text: "How I wonder what you are", start: "35.5", end: "42" }
];
```

What is happening here?

- When the speech recognition is initiated, the song will also start.
- We'll check if the current time falls between the start and end of that lyric line.
- If yes, let's add a class showing that it is the current lyric line.

```
startbtn.addEventListener("click", () => {
    recognition.start();
    music.play().then(() => {
        music.addEventListener("timeupdate", (e) => {
            lyrics.forEach((songLine, index) => {
                tempTranscript = transcript;
                var currentLine = lyricsList.children[index].firstChild;

                if (e.target.currentTime >= songLine.start &&
                    e.target.currentTime <= songLine.end) {
                    currentLine.classList.add("karaoke__lyrics-item--
selected");
                    window['line'+index] = tempTranscript;
                }

                //...
```

```
startbtn.addEventListener("click", () => {
    recognition.start();
    music.play().then(() => {
        music.addEventListener("timeupdate", (e) => {
            lyrics.forEach((songLine, index) => {
                tempTranscript = transcript;
                var currentLine = lyricsList.children[index].firstChild;

                if (e.target.currentTime >= songLine.start &&
                    e.target.currentTime <= songLine.end) {
                    currentLine.classList.add("karaoke__lyrics-item--
selected");
                    window['line'+index] = tempTranscript;
                }

                //...
```

- If the current time is bigger than the “end” time of a lyric line, add a class that marks it as the past.

- Once that line is in the “past” let’s do some checks.

So what's the verdict?

- Okay for small voice notes;
- Not perfect for singing;
- Unlikely to be okay for rap;

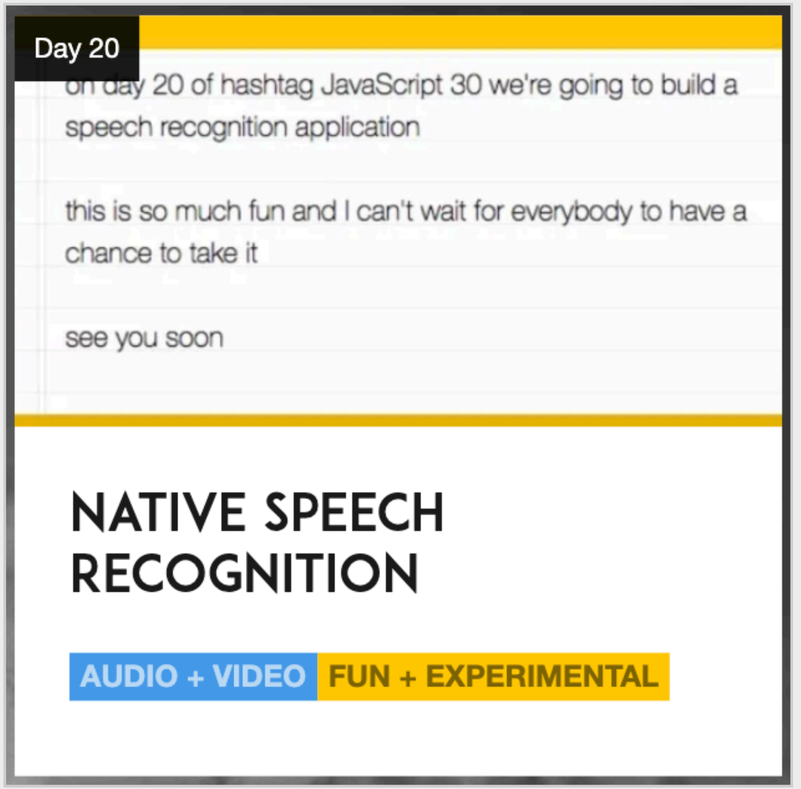
Tony Edwards - Beats, Rhymes and Unit Tests - Frontend Love 2020

JSWORLD Conference
10.3K subscribers
Subscribe
1
Share
Download
Clip

https://www.youtube.com/watch?v=x_L1eQT6TyA

By Stephanie Eckles

https://12daysofweb.dev/2021/speech-api/



By Wes Bos

https://javascript30.com

WEB CAPTIONER

Web Captioner will sunset and be open sourced on October 31, 2023.

July 26, 2023
By Curt Grimes

Web Captioner will sunset and stop ope
continue to access Web Captioner at ht

Since its start in 2017, this project has p
about accessibility and provide some to

As I promised in the past, if I were to su
Web Captioner source available on my (
31, 2023 after Webcaptioner.com is sun

GO TO WEB CAPTIONER Available

https://webcaptioner.com

README.md

speech-recognition-aws-polyfill

bundlephobia inaccessible | vulnerabilities no longer available | npm v0.3.0

A polyfill for the experimental browser Speech Recognition API which falls back to AWS Transcribe.

Features

- Works without a server (browser-only)
- Supports the following browsers/versions (~94% coverage)

Note: this is not a polyfill for `MediaDevices.getUserMedia()` - check the support table in the link above.

Who is it for?

This Library is a good fit if you are already using AWS services (or you would just prefer to use AWS).

A polyfill also exists at: /antelow/speech-polyfill, which uses Azure Cognitive Services as a fallback. However, it seems to have gone stale with no updates for ~2 years.

Prerequisites

- An AWS account
- ognito identity pool (unauthenticated or authenticated) with the `TranscribeStreaming` permission.

Setup Guide

1. In the AWS console, visit the Cognito section and click **Manage Identity Pools**

ceuk CEUK
JamesBrill

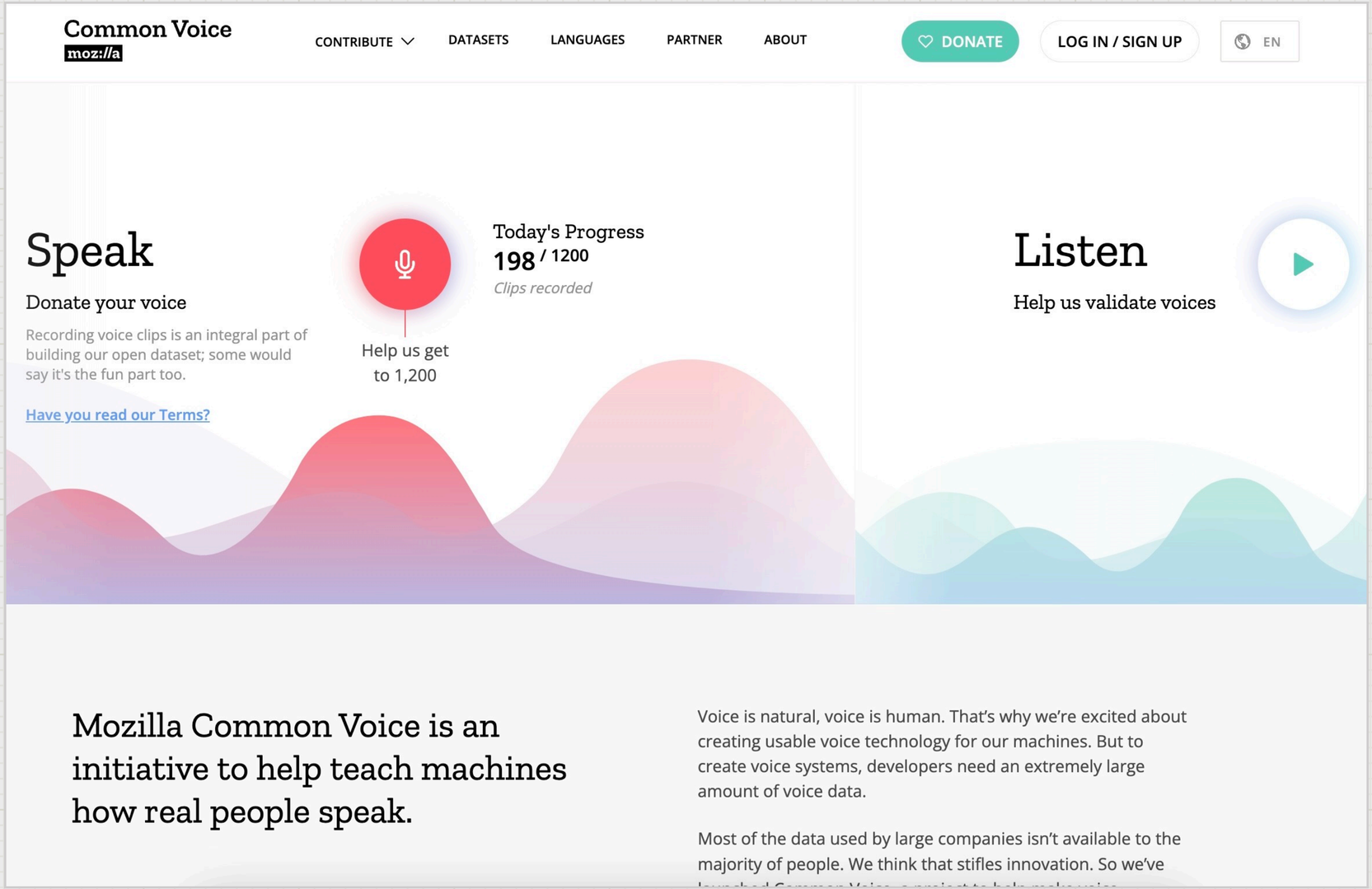
Languages

TypeScript 96.3% JavaScript 3.7%

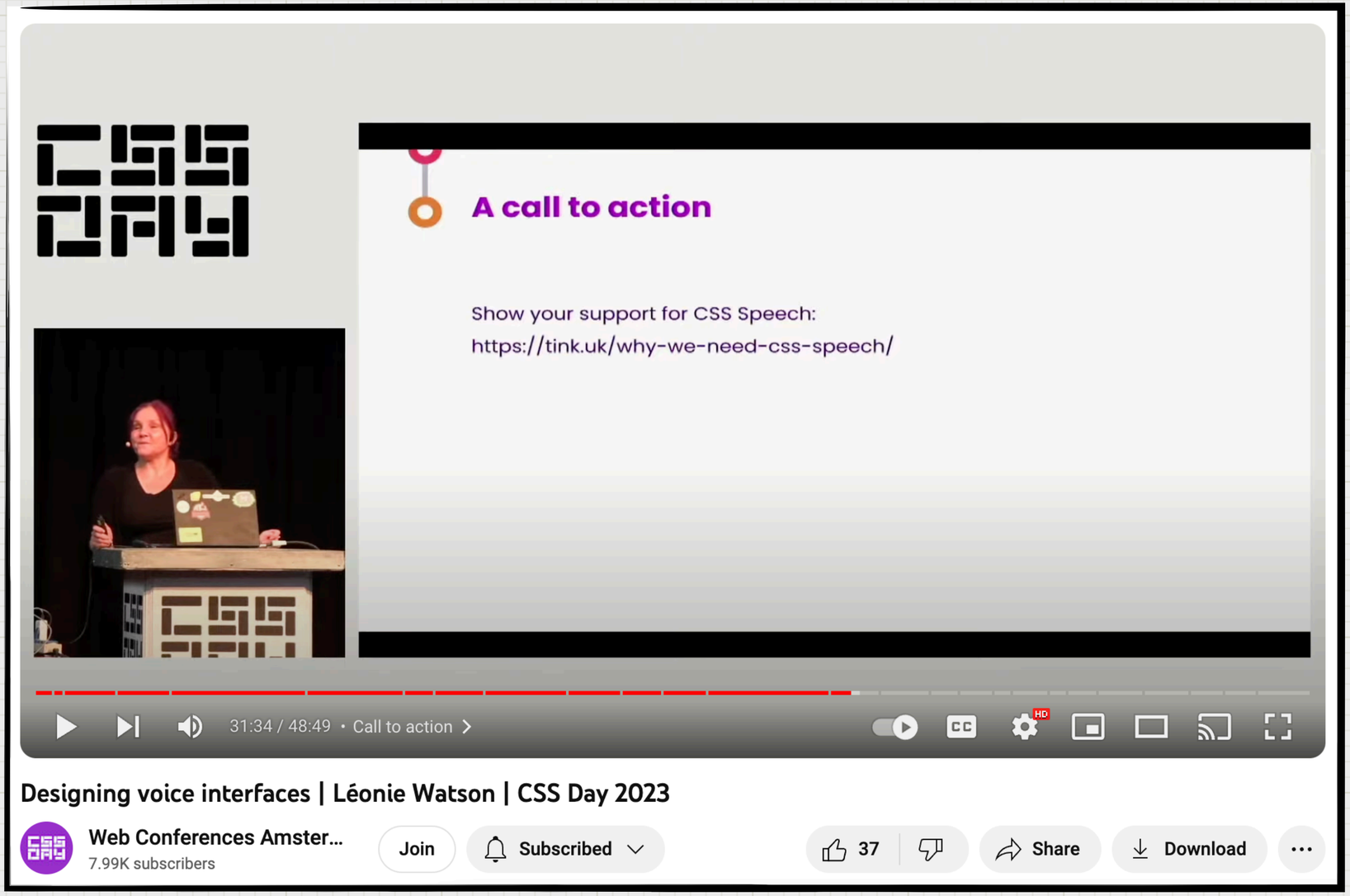
https://github.com/ceuk/speech-recognition-aws-polyfill

What's next?

I don't know but I would like to see these types of APIs grow as they open a world of FREE creative opportunities.



https://commonvoice.mozilla.org/en



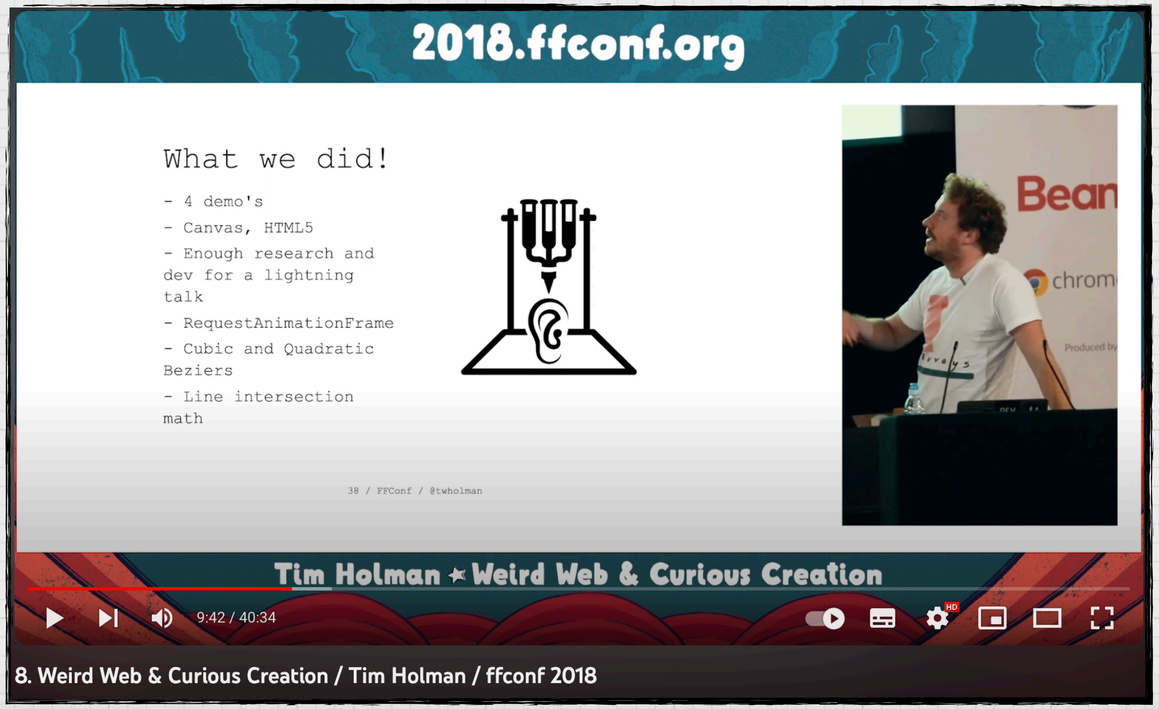
https://www.youtube.com/watch?v=6X23I9yHqd4

Are you going to use the Web Speech API at work?

Probably not.

“Look at what inspires you”

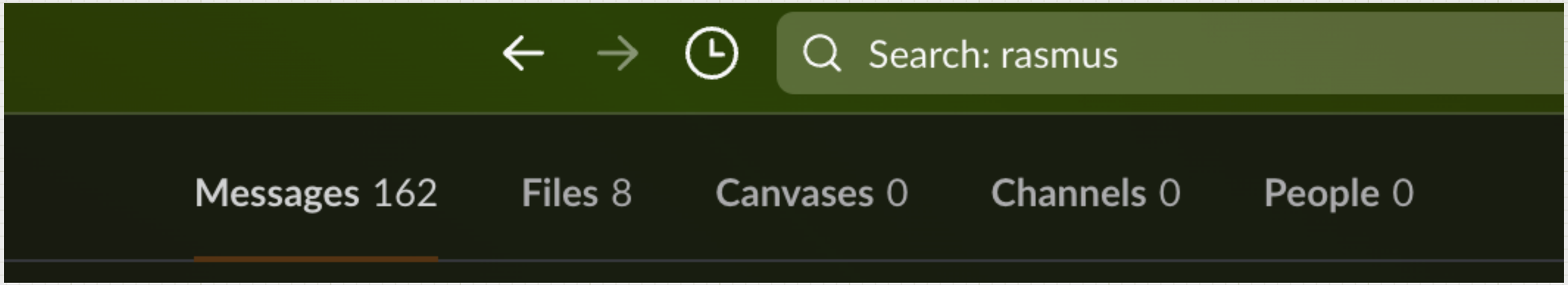
- Tim Holman, FFConf 2018



What we did!

- One working demo!
- Experiment with a browser API: WebSpeechApi and HTMLMediaElement
- Lots of javascript
- Content for a talk
- CSS animations
- Read and understand API Specifications

You can learn by building “useless” things.


Search: rasmus
Messages 162
Files 8
Canvases 0
Channels 0
People 0

Start Stop

alright let's do this

hello everyone i'm so happy to be here

truly feels like the first day of my life

today is about exploring the potential of the web speech api in karaoke

hello my name is anna kendrick

front-end developer at factor you can find them in websites where i also have a cancer via my personal website

spend as much time as possible turning to my blog

i also participated in the indian community and i've

Thank you!

THE RASMUS AND I 💖

Find me on my website:
ohhelloana.blog

MY RASMUS PLAYLIST