## Gold Sponsors

**AUTOMATTIC** 





Silver Sponsors





## **Bronze Sponsors**







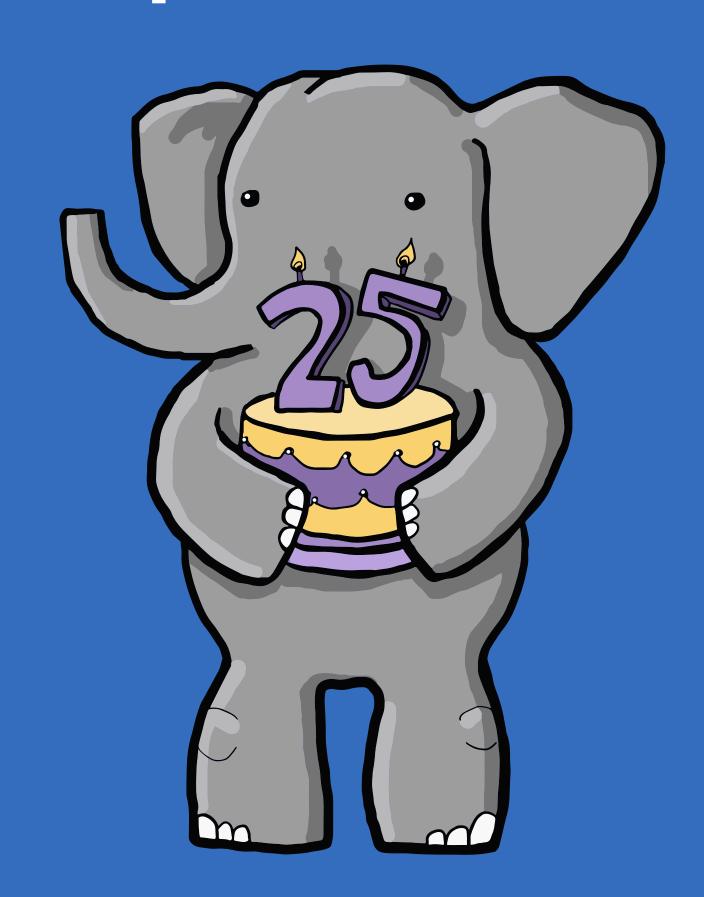








## Thanks to Our Sponsors



## TOOLS AND TIPS FOR GLUING TOGETHER THE OPEN WEB

php[world] 2019



@ChrisHardie

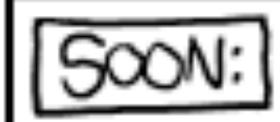




## HOW STANDARDS PROLIFERATE: (SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION: THERE ARE 14 COMPETING STANDARDS.





SITUATION: THERE ARE 15 COMPETING STANDARDS.

https://xkcd.com/927/



VICE VICELAND I-D IMPACT UNITED STATES EN



Video Podcasts News

## MOTHERBOARD

**TECH BY VICE** 

## Yahoo Groups Is Winding Down and All Content Will **Bermanently Removed**

Users won't be able to upload new content to the site after October 28 and have until December 14 to archive their content, Yahoo said in an announcement.



# "A web by and for all its users, not select gatekeepers or governments." —Mark Surman

https://www.yearofopen.org/november-open-perspective-what-is-open-web/what-is-the-open-web-and-why-is-it-important-submitted-by-mark-surman-executive-director-of-the-mozilla-foundation/







## Part Two: Maria speaks of transition, settling into Richmond, and her dreams of bringing more Mexican culture to her adopted city!

Recorded August 28, 2018 Archived August 28, 2018 5:00 minutes



ID: APP518310



#### **DESCRIPTION**

Part one of two: Maria sits down with Alison and translator Kaitlyn to talk about family, her move to Richmond, and ways she believes we can incorporate more Mexican culture into the town she has grown to love and appreciate!

### **PARTICIPANTS**

Kaitlyn Blansett

Maria Orduño-Valle

```
<!DOCTYPE html>
<html lang="en-US">
   <head>
      </or>
      <meta name="viewport" content="width=device-width, initial-scale=1" />
      <script type="text/javascript">
         WebFontConfig = {
            custom: {
              families: ['gotham', 'mercury'],
              urls: [ 'https://archive.storycorps.org/wp-content/themes/storycorps-twenty-sixteen/assets/css/storycorps---fonts.css' ]
        };
         (function() {
            var wf = document.createElement('script');
           wf.src = ('https:' == document.location.protocol ? 'https' : 'http') + '://ajax.googleapis.com/ajax/libs/webfont/1/webfont.j
           wf.type = 'text/javascript';
           wf.async = 'true';
            var s = document.getElementsByTagName('script')[0];
            s.parentNode.insertBefore(wf, s);
        })();
      </script>
      <noscript>
         </or>
      </noscript>
```



```
<html lang="en-US">
    <head>
         <link rel="shortcut icon" href="https://archive.storycorps.org/wp-content/themes/archive-child/assets/images/favicon.ico" />
         <meta name="viewport" content="width=device-width, initial-scale=1" />
         <script type="text/javascript">
              WebFontConfig = {
                   custom: {
                        families: ['gotham', 'mercury'],
                        urls: [ 'https://archive.storycorps.org/wp-content/themes/storycorps-twenty-sixteen/assets/css/storycorps---fonts.css'
              };
              (function() {
                   var wf = document.createElement('script');
                   wf.src = ('https:' == document.location.protocol ? 'https' : 'http') + '://ajax.googleapis.com/ajax/libs/webfont/1/webfont.j
                   wf.type = 'text/javascript';
                   wf.async = 'true';
                   var s = document.getElementsByTagName('script')[0];
                   s.parentNode.insertBefore(wf, s);
              })();
         </script>
         <noscript>

         </noscript>
```



## REST API Developer Endpoint Reference # https://developer.wordpress.org/rest-api/reference/

Resource	Base Route
Posts	/wp/v2/posts
Post Revisions	/wp/v2/posts/ <id>/revisions</id>
Categories	/wp/v2/categories
<u>Tags</u>	/wp/v2/tags
<u>Pages</u>	/wp/v2/pages
Page Revisions	/wp/v2/pages/ <id>/revisions</id>
Comments	/wp/v2/comments
<u>Taxonomies</u>	/wp/v2/taxonomies
<u>Media</u>	/wp/v2/media

```
_id: 535385,
- author: {
     _id: 215129,
     name: "Alison Zajdel",
     },
 default_source: "storage",
- audio: {
     url: "https://api.soundcloud.com/tracks/312582030?secret token=s-Y55Kk",
     length: 23590,
     source: "soundcloud"
- interview audio: {
     url: "https://s3-us-west-2.amazonaws.com/storycorpsarchive/APP/2018/11/535385/Bottorff.mp3?X-Amz-Content-Sha256=UNSIGNED-PAYLOAD&X-Amz-Algorithm=AWS4
     Credential=AKIAIKDMAUBDBYPHPKSA%2F20191014%2Fus-west-2%2Fs3%2Faws4 request&X-Amz-Date=20191014T163552Z&X-Amz-SignedHeaders=host&X-Amz-Expires=604800
     Signature=f04a37ce94596140f73f5ca2b0598f6b916643a90bc4075e3f714776f5530ed7",
     length: 2208,
     source: "storage"
 time: "",
 status: "publish",
 activity: [ ],
 title: "" If you want something to happen, you just have to go make it happen. " – Jon and Aria-Blair about the way they experience our
 description: "Jon, Aria-Blair, and their toddler daughter join Alison for a conversation about their lives in Richmond, their two businesses (Black Dog
 safe spaces for women and creative places for artists in Richmond.",
- additional_participants: [
   - {
        first_name: "Aria",
        last_name: "-Blair"
     },
        first_name: "Jon",
        last_name: "Bottorff"
        first_name: "Alison",
        last_name: "Zajdel"
                                         https://api.storycorps.me/wp-json/interviews?filter[places]=richmond,indiana
 questions: [ ],
```

```
<?php
$storycorps_archive_html = file_get_contents( 'https://api.storycorps.me/wp-json/interviews?filter[places]=richmond,india
$storycorps_archive = json_decode( $storycorps_archive_html, true );
$xml = new SimpleXMLElement('<rss/>');
$xml->addAttribute("version", "2.0");
$channel = $xml->addChild("channel");
$channel->addChild("title", 'Richmond Indiana interviews on StoryCorps');
$channel->addChild("link", "https://archive.storycorps.org/user/joesmith/interviews/" );
$channel->addChild("description", "Richmond Indiana interviews on StoryCorps");
$channel->addChild("language", "en-us");
foreach ( $storycorps_archive as $interview ) {
        $item = $channel->addChild("item");
        $item->addChild("link", $interview['url'] );
        $interview_date = strtotime( $interview['created_date'] );
        $interview_date_for_title = date( 'F j, Y', $interview_date );
        $item->addChild("pubDate", date( 'r', $interview_date ) );
        if ( ! empty( $interview['additional_participants'] ) && is_array( $interview['additional_participants'] ) ) {
                $interview_subjects = array();
                foreach ( $interview['additional_participants'] as $participant ) {
                        if ( 'Joe' === $participant['first_name'] && 'Smith' === $participant['last_name'] ) {
                                continue;
                        $interview_subjects[] = trim( $participant['first_name'] ) . ' ' . trim( $participant['last_name']
                                           https://gist.github.com/ChrisHardie/73247372aaa00332fa5136e27b325f86
```









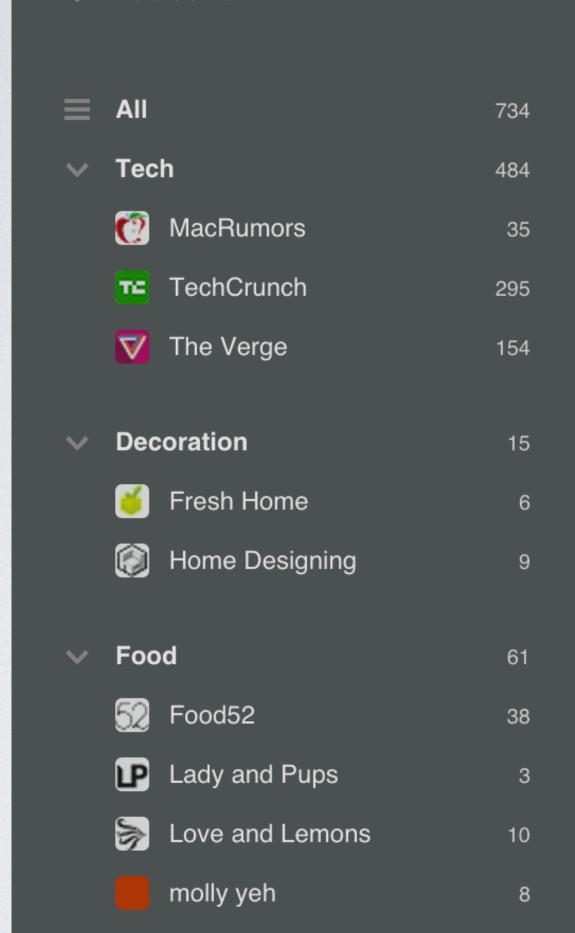
Saved For Later

Shared Collections NEW

Unpin

Team

+ Add Content



smitten kitchen

Competition

edwin@feedly.com

via Google / Logout



## Apple TV Gains Updated NFL Channel With Game Pass Integration

The Apple TV's existing NFL Now channel was today revamped, changing the name to "NFL" and adding support for Game Pass subscriptions. Through the updated channel, NFL fans who have a Game Pass subscription can watch on-100+ MacRumors / by Juli Clover / 2h



### Tep Is An Adorable Fitness Tracking App That Works Like A Tamagotchi

Remember the Tamagotchi? Those little monsters were great. A new iOS app called Tep created a Tamagotchi-like app for your phone to help you stay motivated when it comes to working out. Move around if you want to feed your 400+ TechCrunch / by Romain Dillet / 4h



## Apple Seeds Eighth Beta of OS X El Capitan to Developers, Sixth Beta to Public Testers

Apple today released the eighth beta of OS X El Capitan to developers for testing purposes, nearly two weeks after releasing the seventh El Capitan beta and more than two months after unveiling the operating system at its 2015 300+ MacRumors / by Juli Clover / 4h



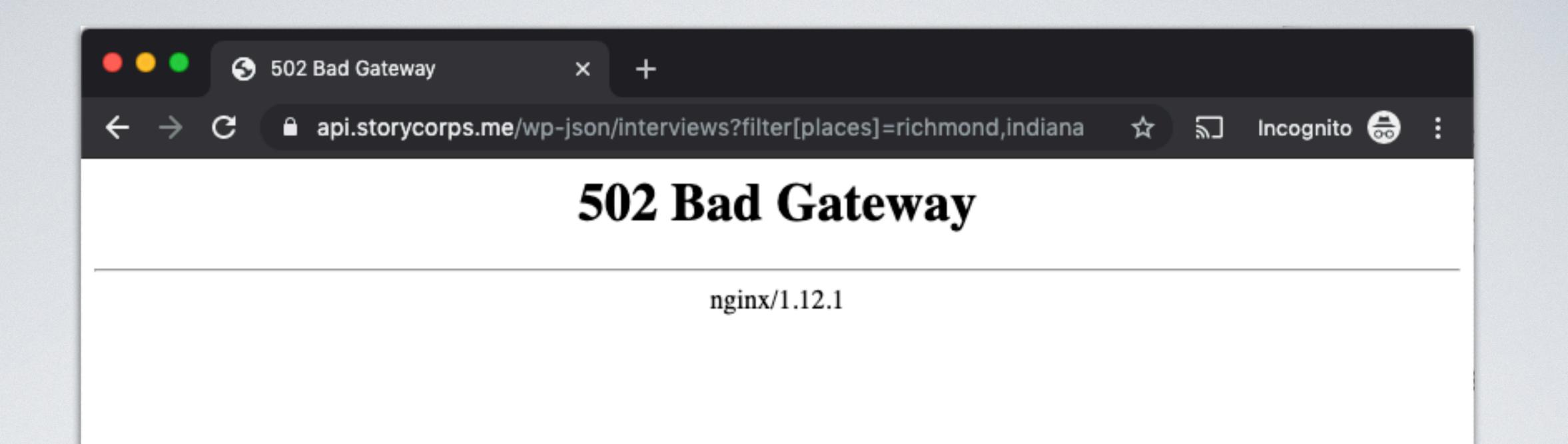
#### Are you still using Apple Music?

Apple Music has officially been available for two months now, and in that time it's had a few ups and downs. Despite some pesky, persisting bugs, Apple Music has quickly gained 11 million subscribers and counting during the trial 1K The Verge / by Micah Singleton / 5h

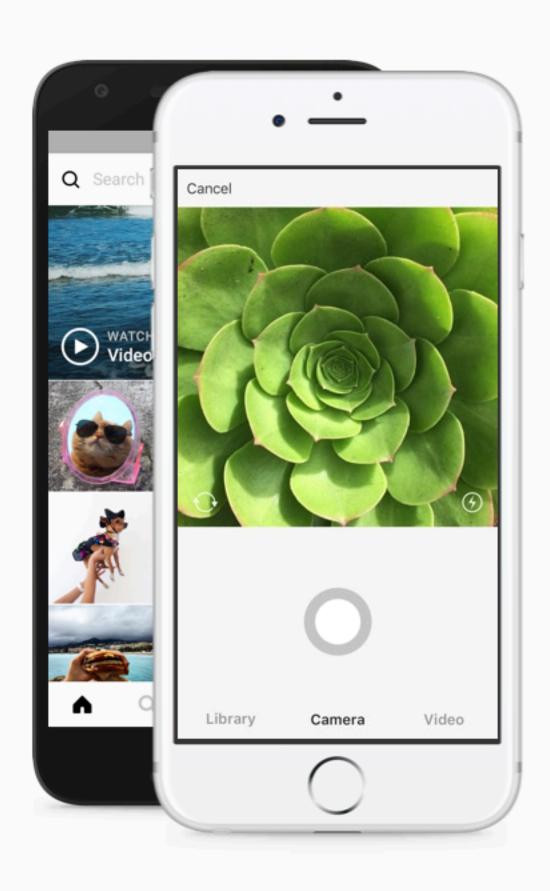
## Eatsa, A Futuristic Restaurant Where Robot Cubbies Serve Quinoa

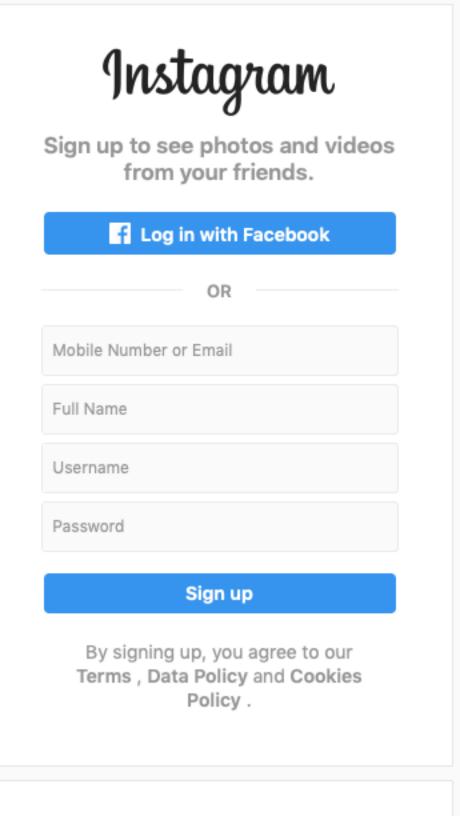
500+ TechCrunch / by Josh Constine / 5h











Have an account? Log in

Get the app.





```
<?php
$target_accounts = array(
       'bikes_as_transportation',
);
foreach ( $target_accounts as $target_user ) {
       $target_user = urlencode( $target_user );
       $results_array = scrape_insta( $target_user );
       $xml = new SimpleXMLElement('<rss/>');
       $xml->addAttribute("version", "2.0");
       $channel = $xml->addChild("channel");
       $channel->addChild("title", sprintf( "%s's photos on Instagram", $target_user ) );
       $channel->addChild("link", sprintf( "https://www.instagram.com/%s/" , $target_user ) );
       $channel->addChild("description", "Instagram Photos");
       $channel->addChild("language", "en-us");
       for (\$i = 0; \$i < 10; \$i++) {
               $feed_item = $results_array['entry_data']['ProfilePage'][0]['graphql']['user']['edge_owner_to_timeline_media']['edges'][$i]['node']
               $item = $channel->addChild("item");
               $item->addChild("link", sprintf("https://www.instagram.com/p/%s/", $feed_item['shortcode'] ) );
               if ( ! empty( $feed_item['edge_media_to_caption']['edges'][0]['node']['text'] ) ) {
                       $item->addChild("title", $feed_item['edge_media_to_caption']['edges'][0]['node']['text'] );
               } else {
                       $item->addChild("title", sprintf( "Image %s from %s", $feed_item['shortcode'], $target_user ) );
               $mediaGroup = $item->addChild('media:group', '', 'http://search.yahoo.com/mrss/');
               $thumbnail = $mediaGroup->addChild( 'media:thumbnail' );
               $thumbnail->addAttribute( 'url', $feed_item['thumbnail_src'] );
                               https://gist.github.com/ChrisHardie/5d8d8dd0a6fe3e1c67df2a7e4444594a7
```

### Posts

Sep 5, 2019

Instagram has blocked geo-pages

Aug 28, 2019

Instagram becomes a paid feature

Aug 27, 2019

Deprecation of Google Plus and Facebook feeds

Aug 27, 2019

Instagram outage

Apr 3, 2019

Classical Twitter search is not available anymore

Jul 4, 2018

Google and Twitter result type option added

Jul 3, 2018

Google authentication added

Apr 14, 2018

Instagram feeds work

Apr 13, 2018

Troubles with Instagram, again.

Apr 9, 2018

Instagram works, the server is stable

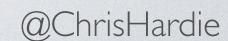
Mar 30, 2018

Troubles with Instagram feeds

Feb 22, 2018

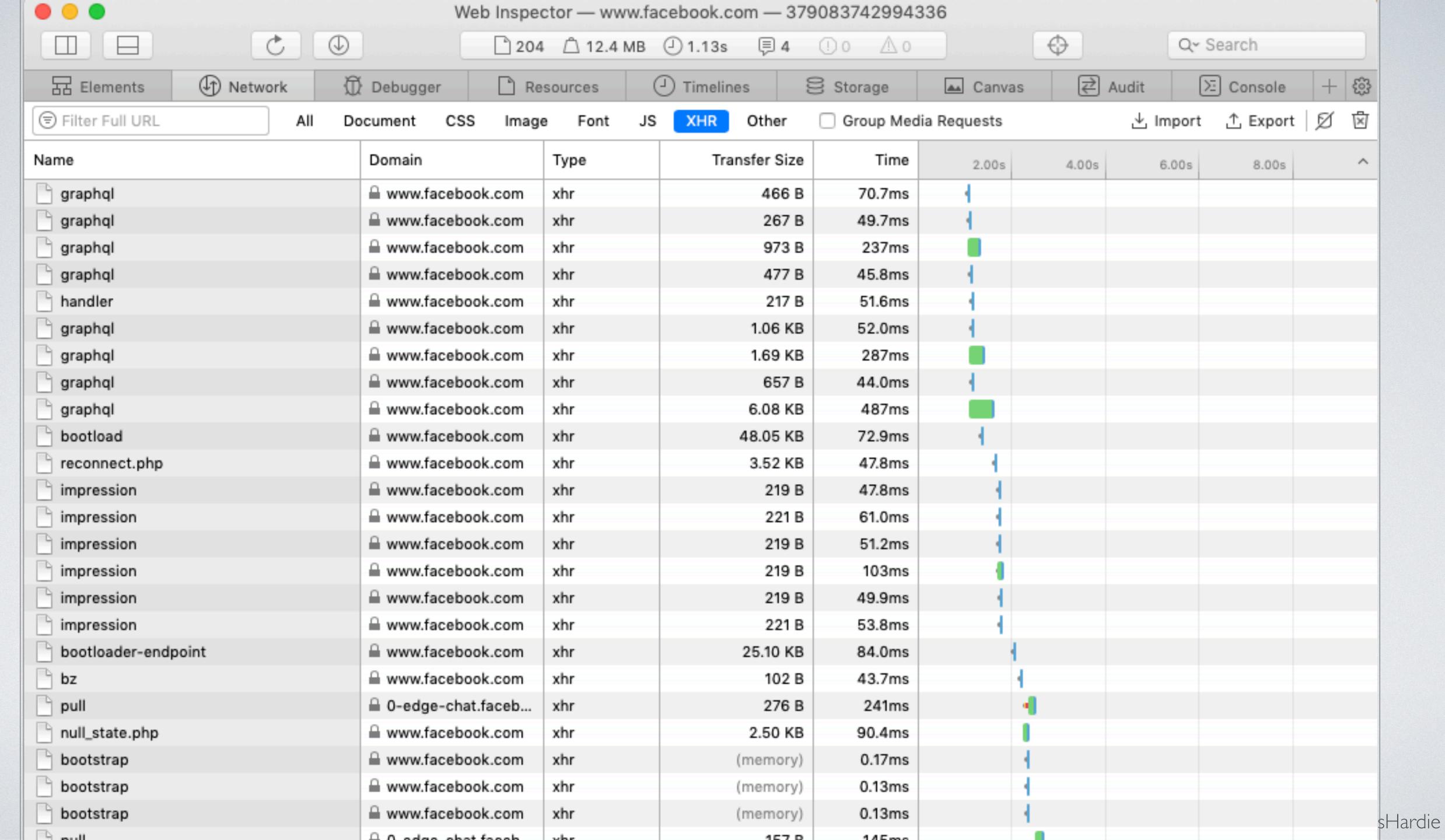


http://queryfeed.net/pages/

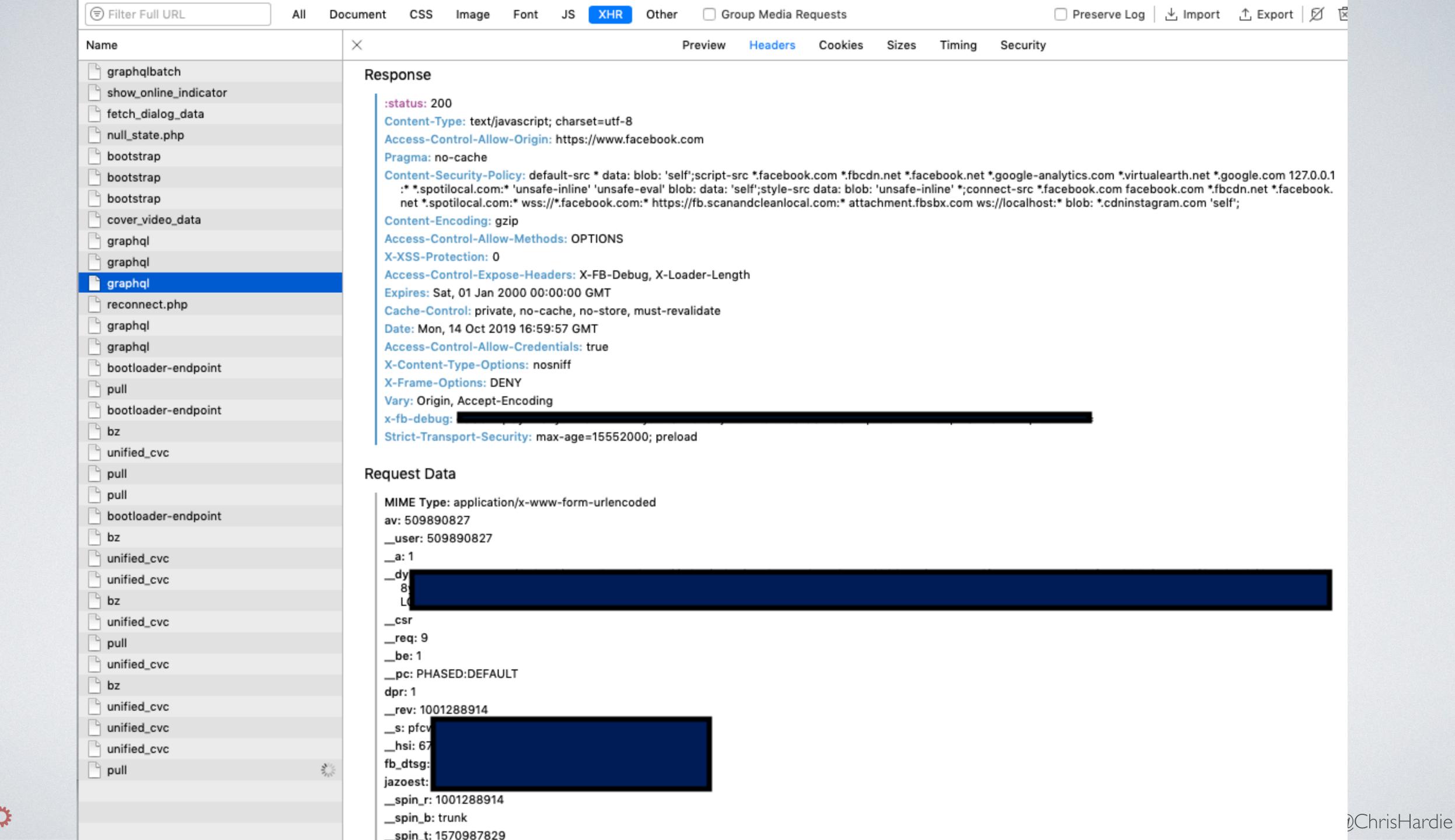


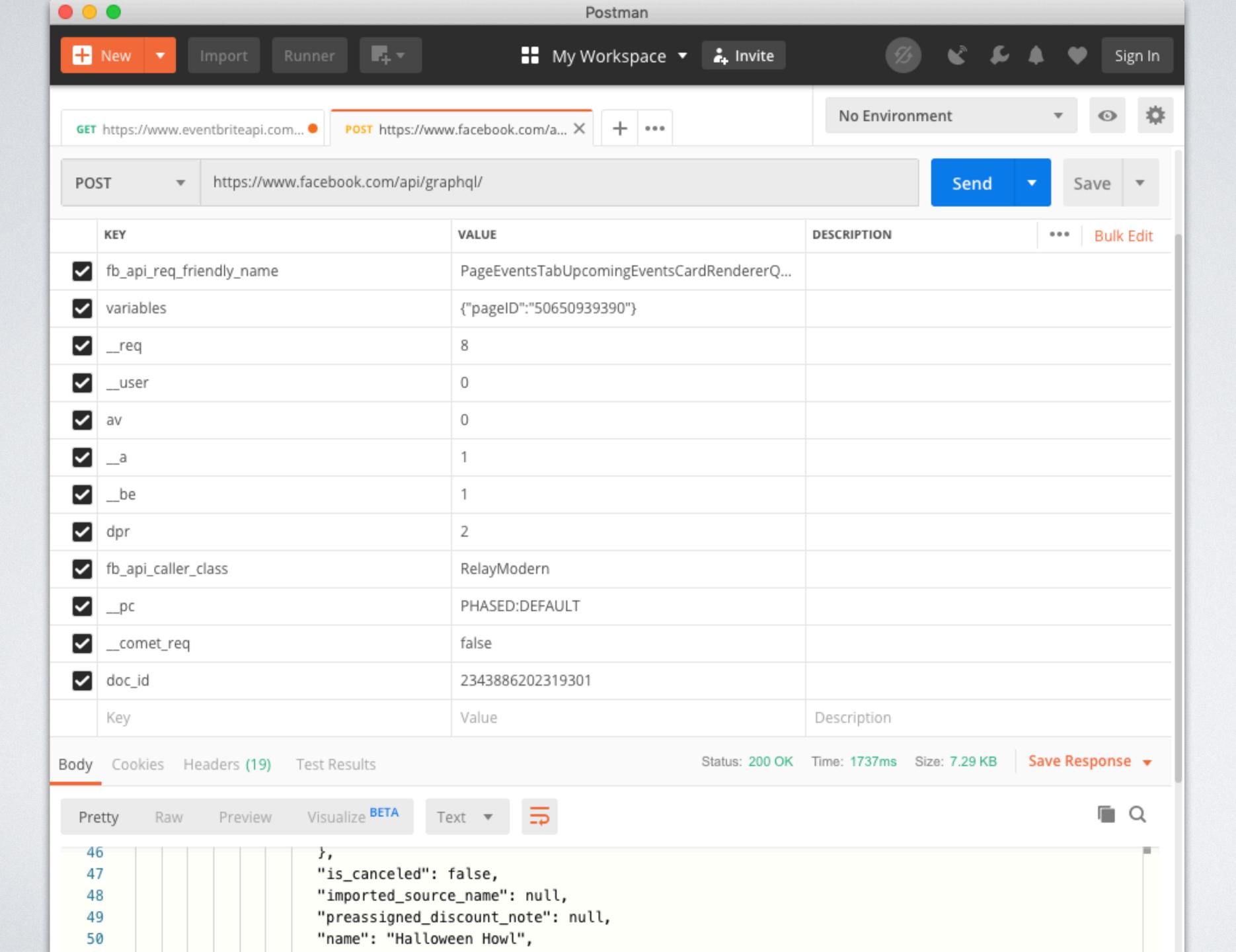






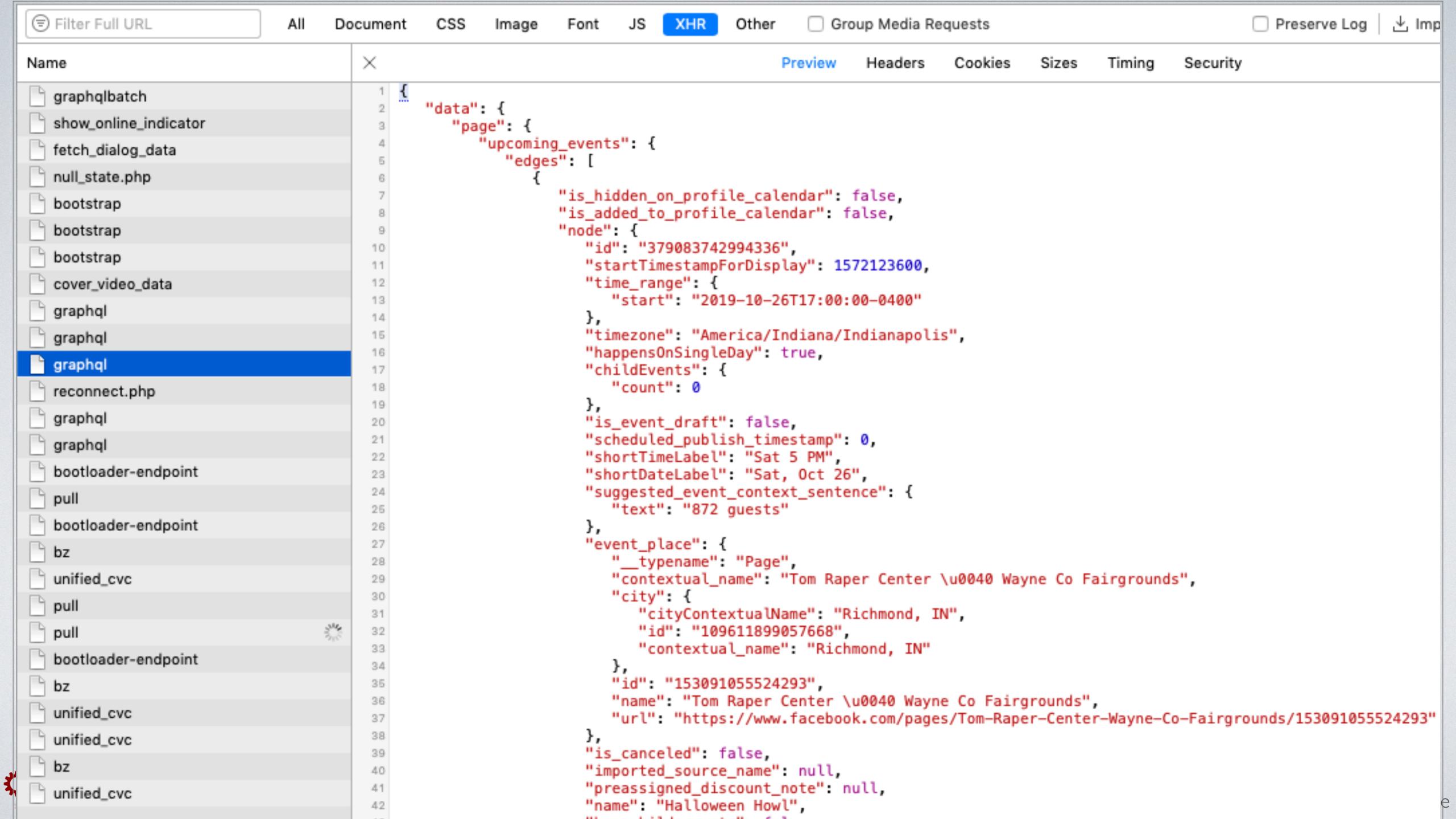








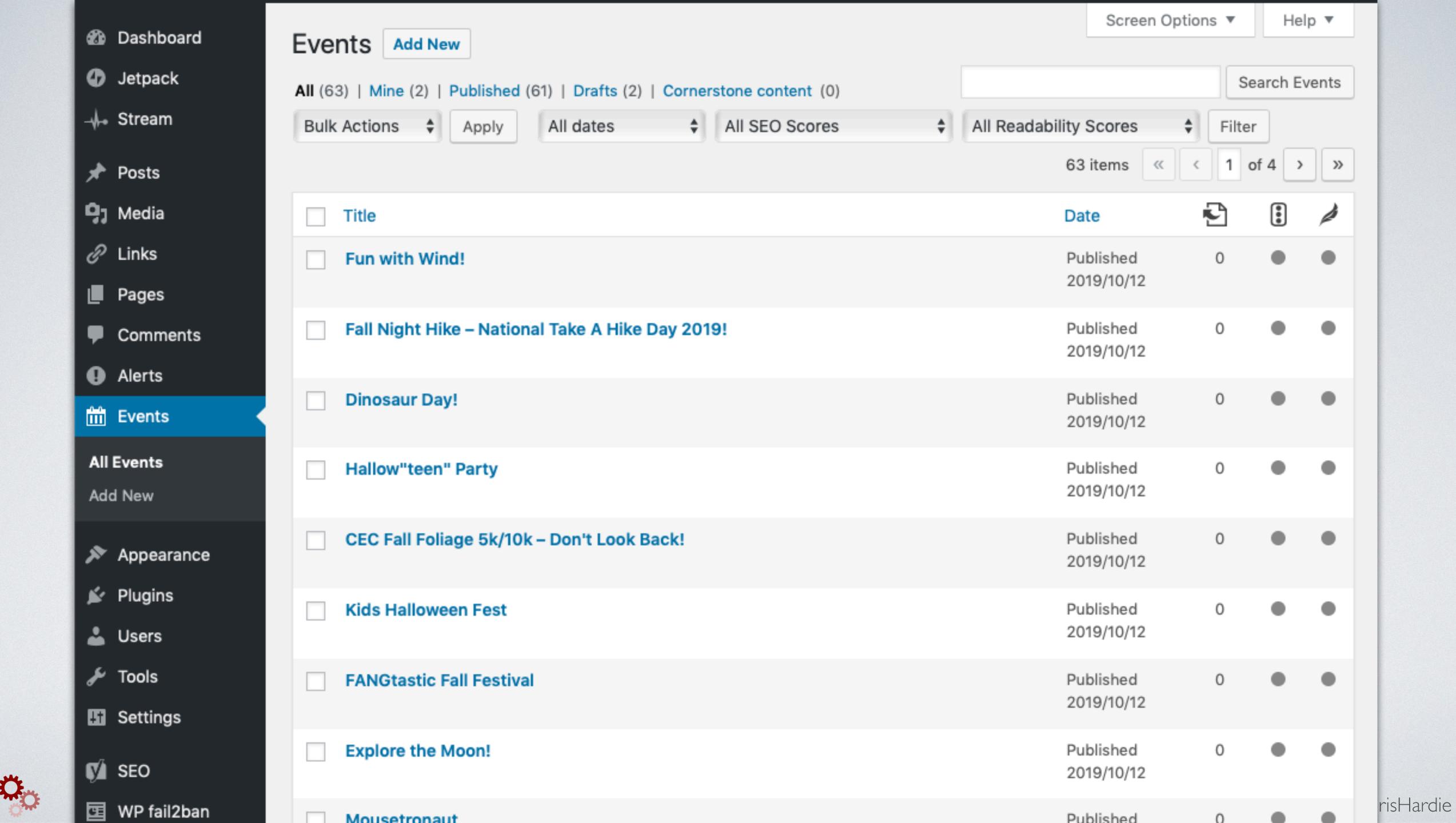




```
require 'vendor/autoload.php';
13
14
    use GuzzleHttp\Client;
15
16
    define( 'FB_UPCOMING_EVENT_DOC_ID', 2343886202319301 );
17
18
    $facebook_pages = array(
19
            '50650939390' => array(
                    'page_name' => 'Richmond Parks Department',
20
                    'page_url' => 'https://www.facebook.com/richmondparks/',
                    'category' => 'outdoor',
            ),
23
24
25
26
    $facebook_base_params = array(
27
            'fb_api_req_friendly_name' => 'PageEventsTabUpcomingEventsCardRendererQuery',
            '__req'
                                      => '8',
28
                                      => '0',
29
            '__user'
            'av'
                                      => '0',
30
                                      => '1',
            '__a'
31
                                      => '1',
            '__be'
32
                                      => '2',
33
            'dpr'
            'fb_api_caller_class'
34
                                      => 'RelayModern',
35
            '__pc:PHASED'
                                      => 'DEFAULT',
            '__comet_req'
                                      => 'false',
36
38
                                https://gist.github.com/ChrisHardie/3962b3d4fcb61f0f074c26d02edb2e68
    $client = new Client(
                    // Base URI is used with relative requests
41
                                                                                                                                    @ChrisHardie
                    'base_uri' >> 'https://www.facebook.com',
42
```

```
if ( ! $existing_post && ! $existing_post_by_name ) {
   // Create the new event post
   $new_event_post = wp_insert_post(
       wp_slash(
           array(
               'post_type' => 'event',
               'post_name' => 'facebook-' . esc_attr( $event->id ),
               'post_title' => esc_html( $event->name ),
               'post_content' => esc_html( $event_text ),
               'post_status' => 'publish',
               'post_date_gmt' => esc_html( $post_date ),
               'meta_input' => array(
                   'event_source_details_url' => esc_url( $event_url ),
                   'event_url'
                                            => esc_url( $event_url ),
                   'event_source_slug' => 'facebook',
                   'event_source_id' => esc_attr( $event->id ),
                   'event_start_time' => esc_attr( $event_start_time ),
                                            => esc_attr( $event_end_time ),
                   'event_end_time'
    );
```





₿ CFAA —

## Web scraping doesn't violate antihacking law, appeals court rules

Employer analytics firm can keep scraping public LinkedIn profiles, court says.

TIMOTHY B. LEE - 9/9/2019, 5:00 PM



Scraping a public website without the approval of the website's owner isn't a violation of the Computer Fraud and Abuse Act, an appeals court ruled on Monday. The ruling comes in a legal battle that pits Microsoft-owned LinkedIn against a small data-analytics company called hiQ Labs.





HiQ scrapes data from the public profiles of LinkedIn users, then uses the data to help companies better understand their own workforces. After tolerating hiQ's scraping activities for several years, LinkedIn sent the company a cease-and-desist letter in 2017 demanding that hiQ stop harvesting data from LinkedIn profiles. Among other things, LinkedIn argued that hiQ was violating the Computer Fraud and Abuse Act, America's main anti-hacking law.

This posed an existential threat to hiQ because the LinkedIn website is hiQ's main source of data about clients' employees. So hiQ sued LinkedIn, seeking not only a declaration that its scraping activities were not hacking but also an order banning LinkedIn from interfering.

A trial court sided with hiQ in 2017. On Monday, the 9th Circuit Appeals Court agreed with the lower court, holding that the Computer Fraud and Abuse Act simply doesn't apply to information that's available to the general public.

"The CFAA was enacted to prevent intentional intrusion onto someone else's computer specifically computer hacking," a three-judge panel wrote. The court notes that members debating the law repeatedly drew analogies to physical crimes like breaking and entering. In the 9th Circuit's view, this implies that the CFAA only applies to information or computer systems that were private to start with—something website owners typically signal with a password requirement.



```
# Update manually scraped feeds
15,45 * * * * /home/chris/feeds/bin/update-feeds.sh
# Check for stale feeds
0 8 1,23 * * /home/chris/bin/check-47374info-stale-feeds.sh
# Force 47374.info to update its feeds
*/10 * * * * /usr/bin/curl --silent https://47374.info?update_feedwordpress=1
```

## #!/bin/sh

```
/home/chris/feeds/bin/feed-scraper.pl -p richmondin
/home/chris/feeds/bin/feed-scraper.pl -p waynecounty
/home/chris/feeds/bin/feed-scraper.pl -p wthr
/home/chris/feeds/bin/feed-scraper.pl -p reidhealth
/home/chris/feeds/bin/feed-scraper.pl -p hoosieropportunity
/usr/bin/php /home/chris/feeds/bin/newsapi-feed/newsapi-results-to-rss.php
/home/chris/feeds/bin/get-and-fix-wctv-feed.sh
/home/chris/feeds/bin/get-and-fix-palitem-feed.sh
```







After an initial import, all of these importers can also optionally check each hour and automatically download new content as well, keeping

home-on-the-web, WordPress? Read more about this technique/ap-

proach to data ownership.

import indieweb

**Advanced View** 

keyring social sync

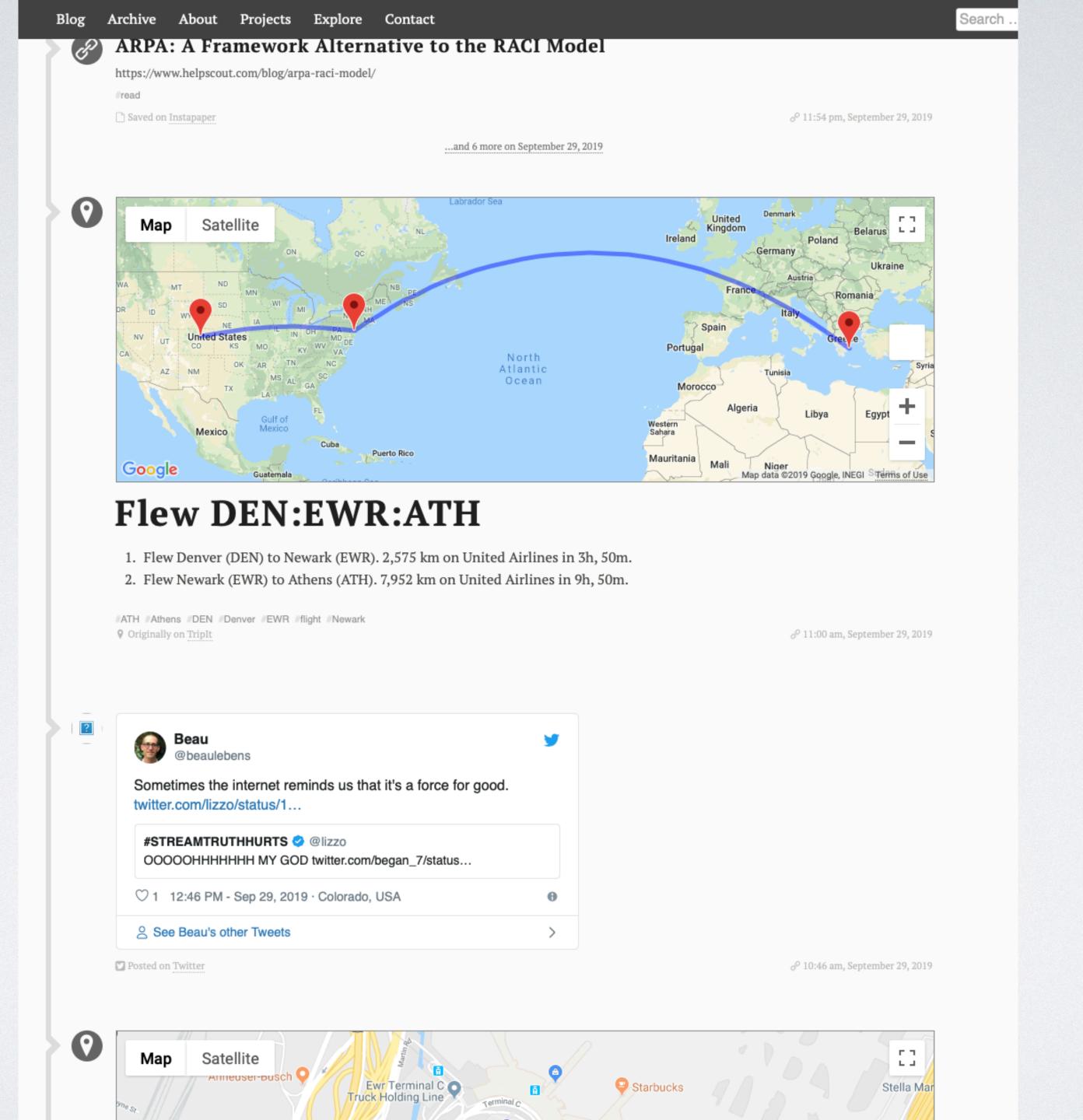
Tags:

## Importers included currently:

- Delicious
- Fitbit
- Flickr
- Foursquare/Swarm
- Instagram
- Instapaper
- Jetpack/WordPress.com
- Moves
- Nest Cameras
- Pinterest
- Pocket
- Strava
- Triplt
- Twitter

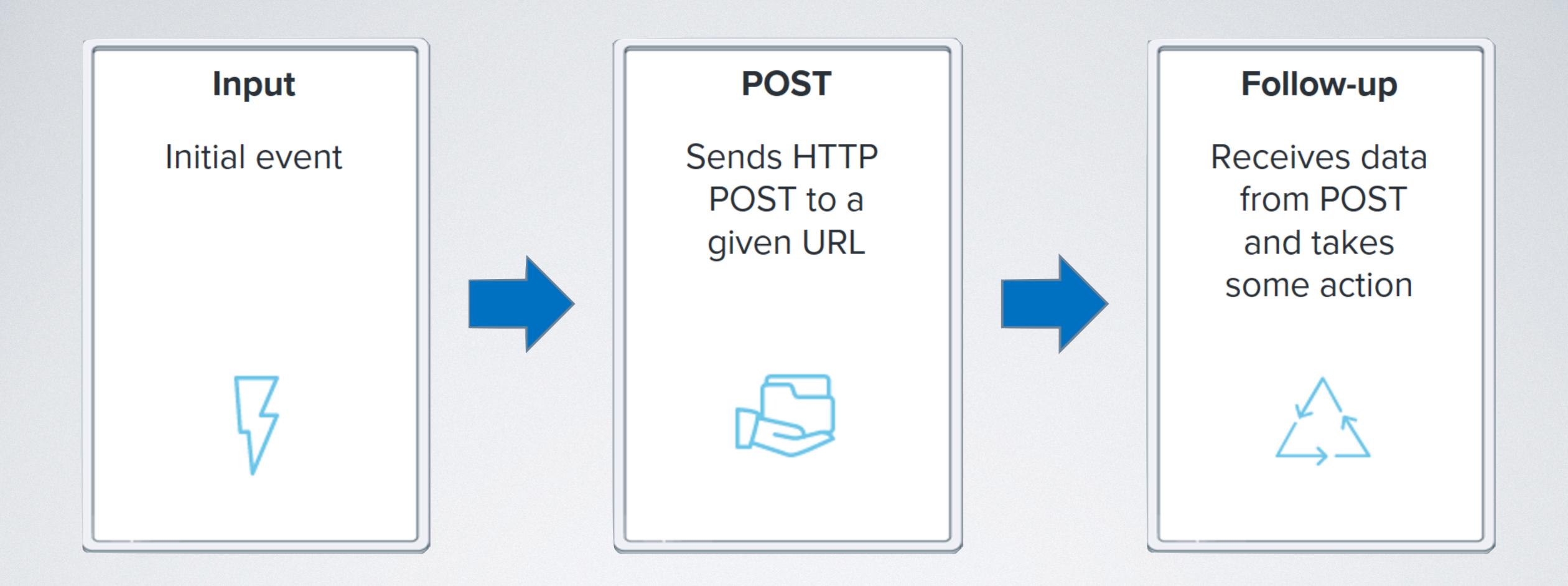
You can potentially <u>write your own importers</u> as well, using the base class included.











https://tray.io/blog/what-are-webhooks-for-non-technical-people







Looks like the dog is going for a w-a-l-k.

5:57 AM - 23 Nov 2015















Mailbox activity detected.

5:55 AM - 17 Sep 2016















It's raining at home - Scattered Thunderstorms and 72F outside.

2:12 PM - 17 Sep 2016

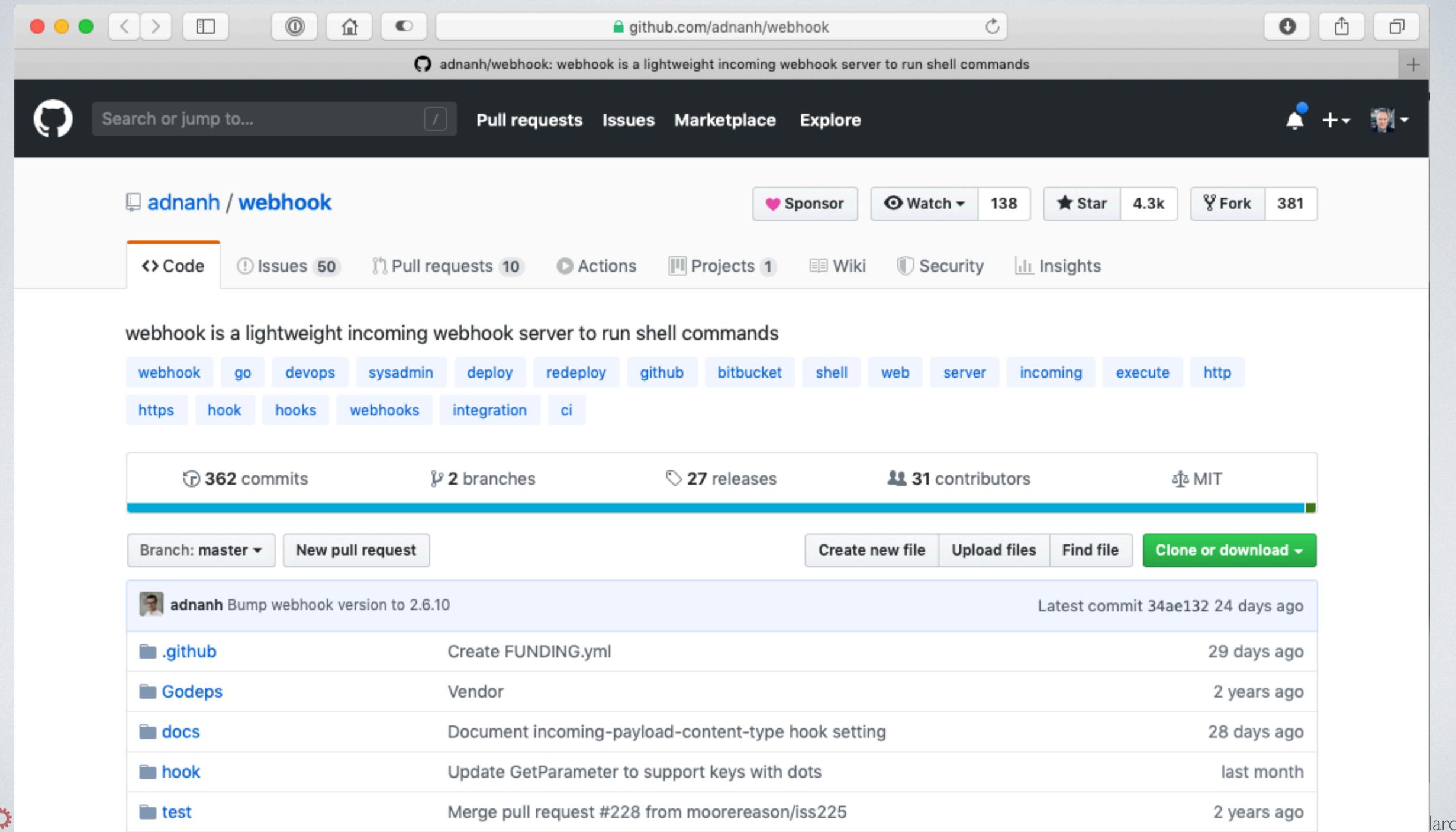






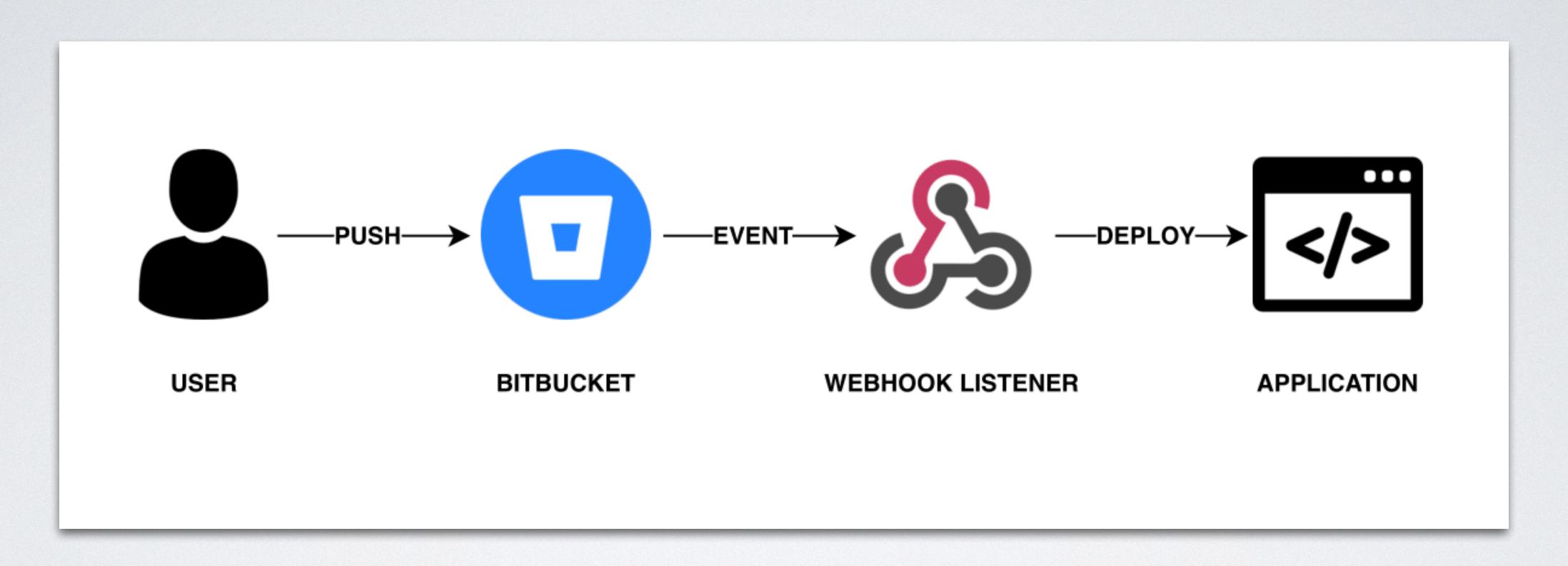








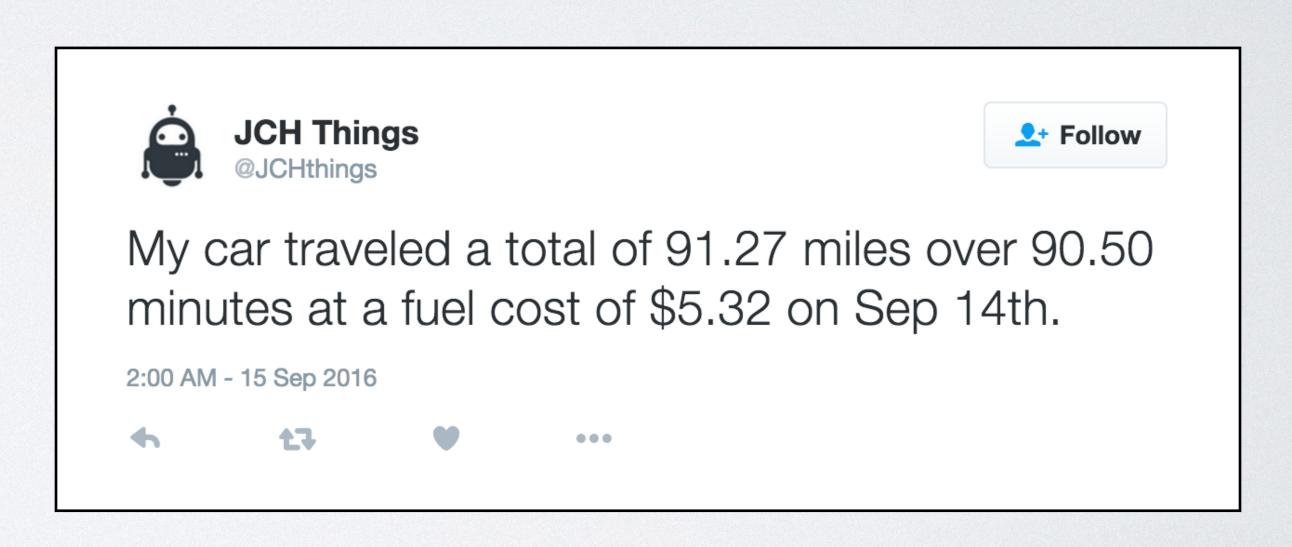
lardie



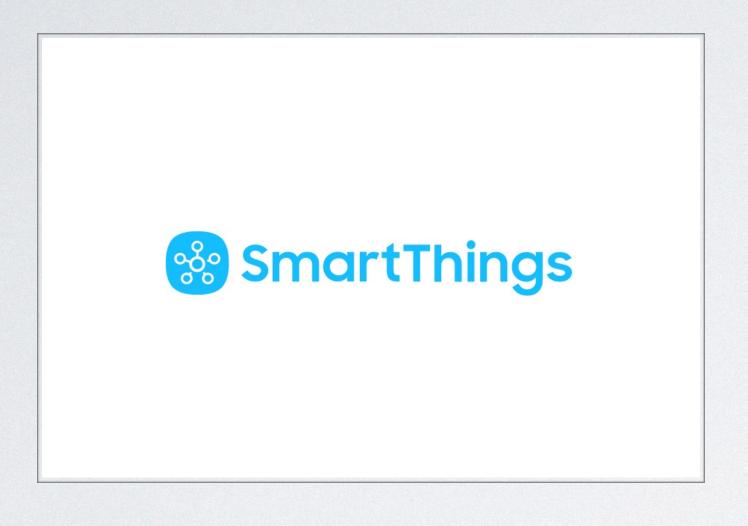
https://medium.com/@pedrokayser00/an-automated-deployment-with-bitbucket-webhook-listener-9394eb2f81c0







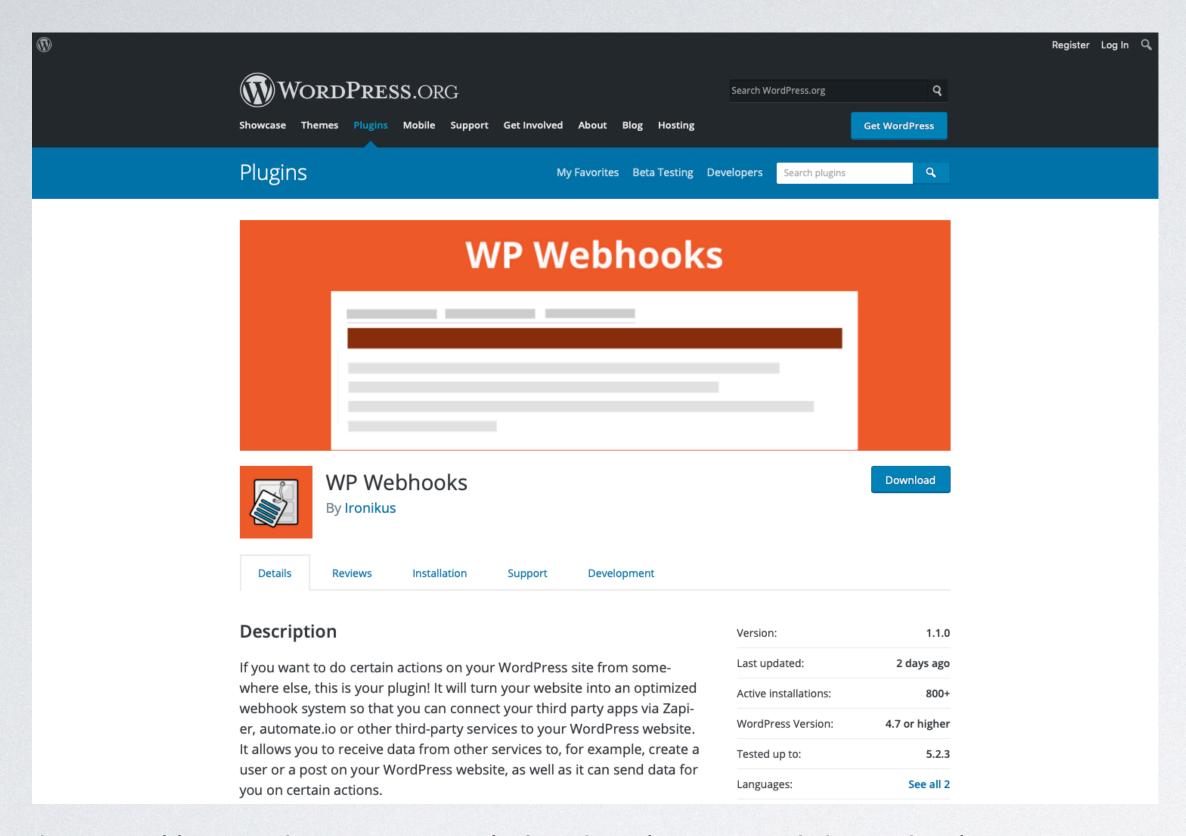






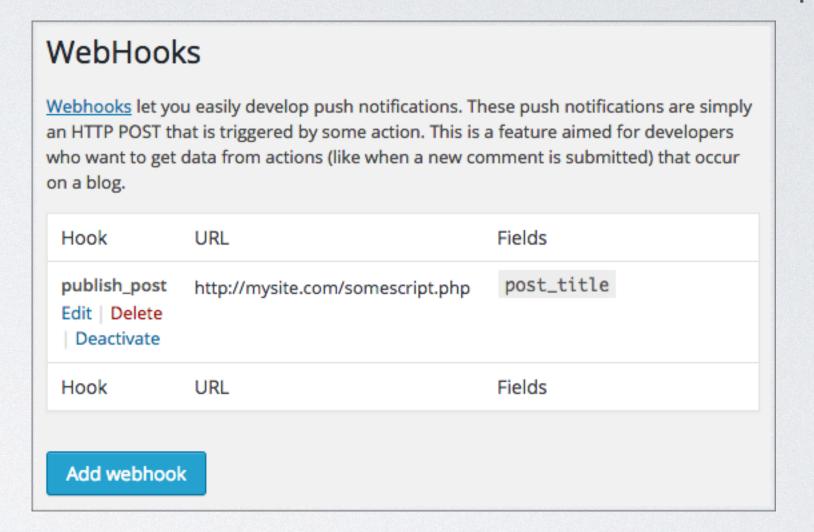
https://github.com/harperreed/SmartThings-webhook





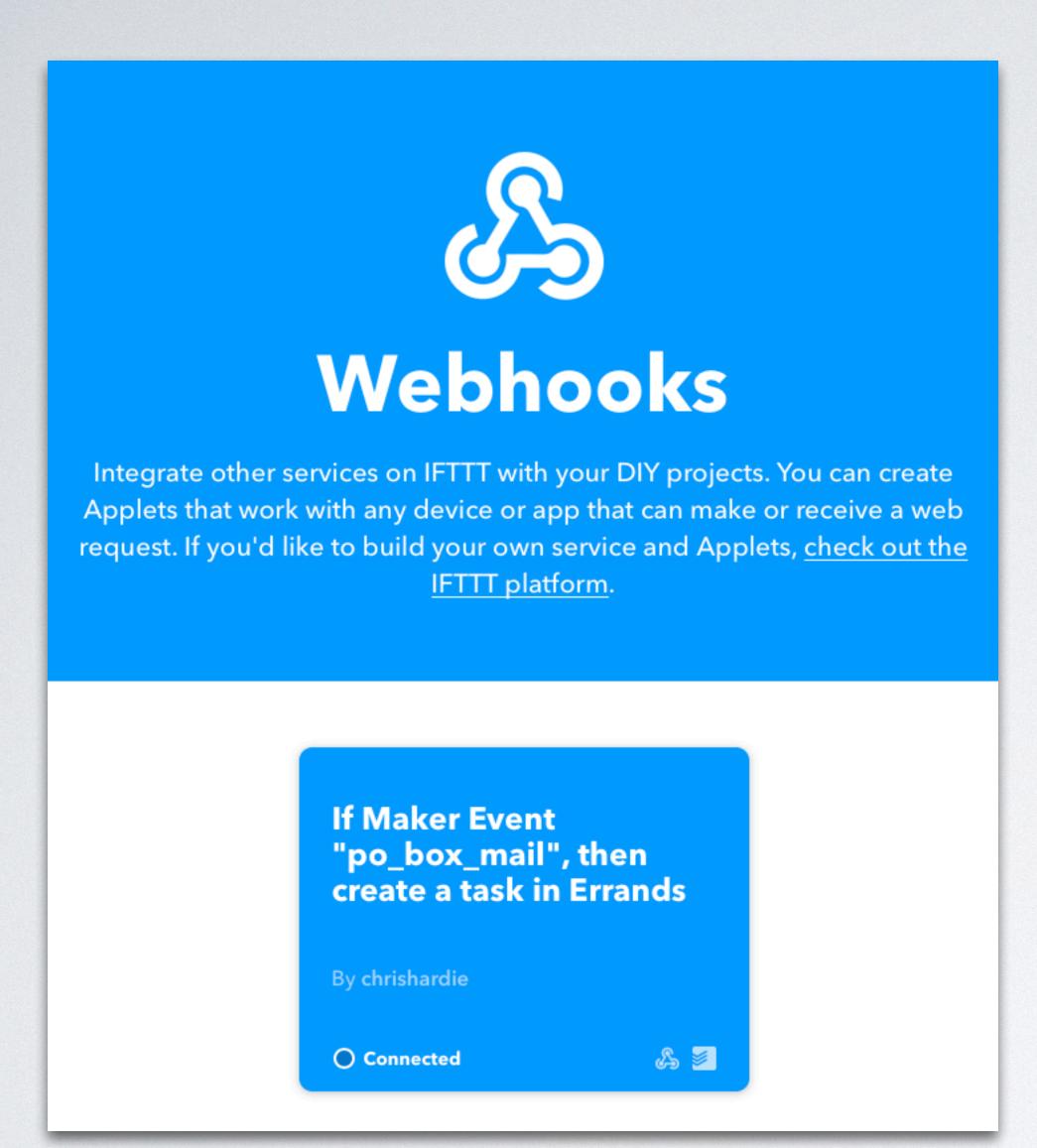
https://wordpress.org/plugins/wp-webhooks/

#### WordPress.com webhook support



https://en.support.wordpress.com/webhooks/





https://ifttt.com/maker\_webhooks





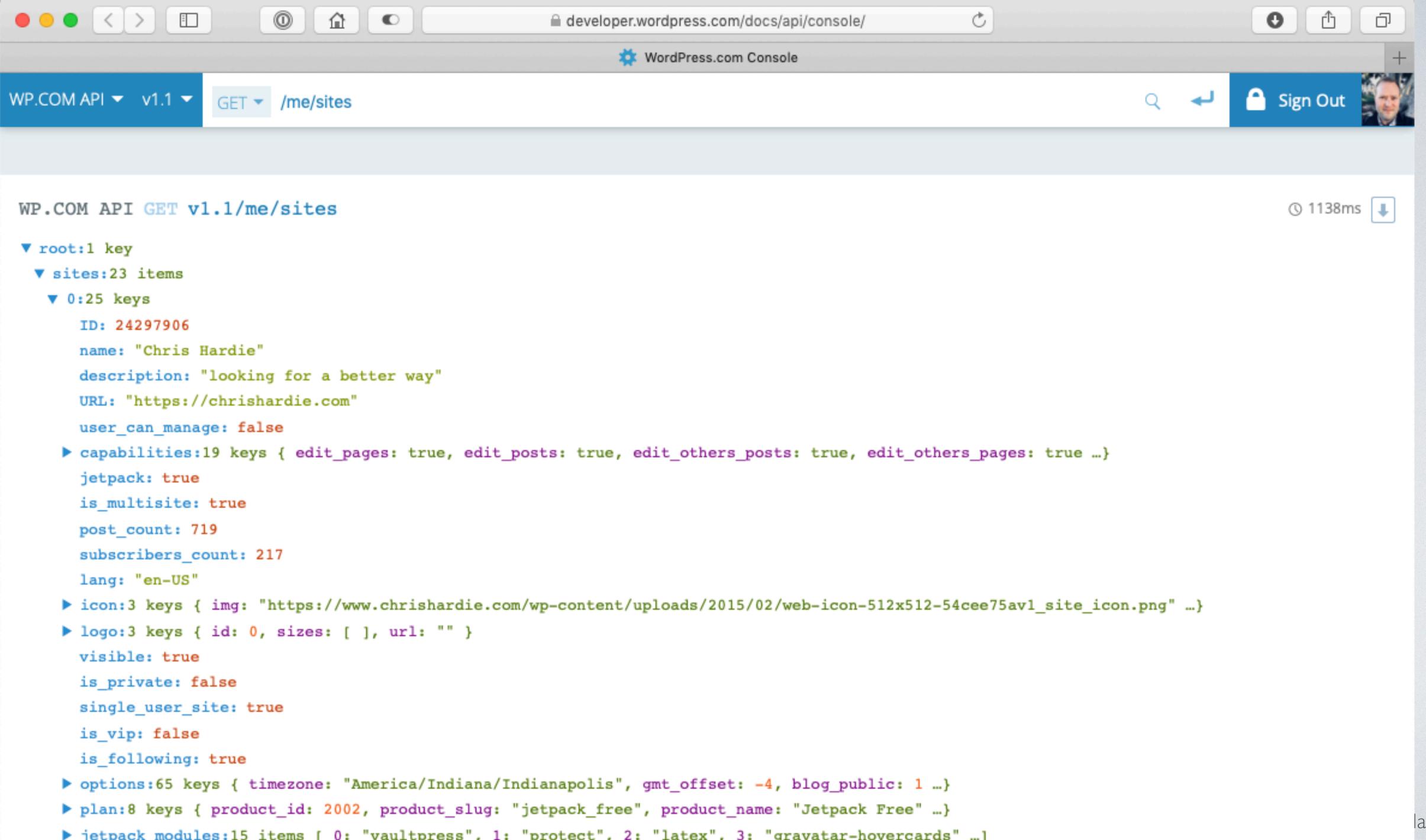
## OPEN WEB CONSIDERATIONS



### API

- Robust
- Well documented
- Supported and maintained
- · Testing console, client libraries, example queries







lardie

### RSS / STRUCTURED DATA FEED

- · Let users consume content their own way
- Easily findable



```
https://chrishardie.com/feed/ X
   C a chrishardie.com/feed/
<?xml version="1.0" encoding="UTF-8"?><rss versi</pre>
        xmlns:content="http://purl.org/rss/1.0/m
        xmlns:wfw="http://wellformedweb.org/Comm
        xmlns:dc="http://purl.org/dc/elements/1.
        xmlns:atom="http://www.w3.org/2005/Atom"
```



#### EXPORTS

- Full data export, including media
- · Common format, easily converted



### DOCUMENTATION

- Even if it's self-explanatory
- Encourage an ecosystem around your tool
- · Consider behind-the-scenes "how we did it" docs too



# OPEN SOURCE IT

• Even if it's ugly or broken



#### USERS

- Take the time to research a service's commitment to the open web
- Ask hard questions
- Share what you find





# THANKYOU. QUESTIONS?

ChrisHardie.com

@ChrisHardie

automattic.com/jobs

joind.in/talk/eeae9

