Site fixes - Site script request/donation #100

[site request] qualtrics.com

01/07/2022 05:32 AM - jacobk

Status:	New		Start date:	01/06/2022
Priority:	Low		Due date:	
Assignee:			% Done:	0%
Category:			Estimated time:	0.00 hour
Target version:				

Description

Yearly my school sends me a Qualtrics survey that I have to fill out in order to renew my scholarship, and sometimes I get other, optional surveys that I also want to fill out, often also on qualtrics.com.

Surveys are at SOMENAME.qualtrics.com/jfe/form/FORMID where SOMENAME is the name of the organization hosting the survey and FORMID is the ID of the individual form. Sometimes, after answering a survey, an email is sent containing the response and a link to view the response: SOMENAME.qualtrics.com/CP/Report.php?SID=FORMID&R=RESPONSEID where SOMENAME and FORMNAME are the same as above and RESPONSEID is the ID of the individual response (since a survey may have multiple responses). Since the response is also part of the email, I don't think the second scheme is very important, but some of the scripts there are already free, so it might be useful to look at?

The scripts in the first scheme are entirely proprietary as far as I can tell. Strings in the script embedded in the HTML file suggest there may be a captcha (though I don't think I've ever had to answer one so hopefully it's only in some cases). It seems that the text that would be displayed as part of the survey is baked into a script embedded in the HTML file (It's JSON data, I think). Displaying the survey seems straight-forward if you can read scripts that are part of the HTML file (hm, I wonder if LibreJS would get in the way of this...), but I'm not sure how one would go about submitting a survey. The scripts are around 500 kilobytes total.

In general, I do not want to share the survey links I have publicly, but there are lots of examples here: https://old.reddit.com/search?q=qualtrics.com&sort=new (not helpful for the second scheme, but I can send the free scripts themselves if you think they could be helpful)

I am willing to pay for the development of a fix to enable answering surveys, but I'm not sure how best to pay, how much it should cost, or how much I can pay (I can for sure pay $100, but I'm not sure if that's way too much or way too little.). I think this is a fix that would benefit multiple people, since Qualtrics surveys seem to be fairly common, so maybe the cost could be split across multiple people?

I also want to learn how fixes are developed, so I can write my own fixes (I am familiar with programming, so I've been able to modify existing fixes without too much trouble, but I'm not sure how to start writing a new fix without any pre-existing knowledge of how the website works.), so I would appreciate information about how fixes are developed as well (using qualtirics.com as an example, perhaps).

Not sure if I should mark this as "Feature" or "Support", but I decided to go with Support since it's not a feature of Haketilo itself but rather a script that Haketilo will load.

History

#1 - 01/08/2022 04:34 PM - koszko

Hi and thanks for posting!

I'll start the response with some bad news: this site uses reCAPTCHAv3. You might recall Google reCAPTCHA challenges having humans select images from a set and stuff like that. It was all reCAPTCHAv2 (or older v1). reCAPTCHAv3 works differently. It monitors person's actions on a site and sends this data to Google to analyze without ever showing any challenge. No chance of getting away without it (unless we reimplement the entire thing).

At some point we definitely want to implement a libre reCAPTCHAv2 script as something like this has already been proven possible. It might also be possible to implement libre reCAPTCHAv3 script and we'd like to do so. However, even the libre reCAPTCHAv3 client we'll eventually come up with might be considered by some to be unacceptably privacy-hostile.

Anyway, any reCAPTCHA-related work is unlikely to be started before we:

- finish Haketilo 1.0 that I am working on all the time
- add mechanism to allow Haketilo scripts to bypass CORS (and maybe also spoof some request headers)

You might want to keep the 100$ for later :)

End of bad news

As to LibreJS compatibility, it is assumed the user would disable LibreJS on the site that is meant to be handled by Haketilo. When we say those 2 extensions are compatible, we mean they won't interfere in some other, more serious ways (e.g. causing redirect loops, deleting HTTP headers the other one relies on, etc.).

You want to write a fix from scratch? I usually start by playing with the page in Developer Tools. You know, the interface you open with Ctrl+Shift+I. Look at page's structure, inspect network traffic, etc. Once you know what the new script should do to make the page work, you use the Developer Tools' Console to run the code while having site's own scripts disabled (either by Haketilo or some other extension). I found it most convenient to only write very short snippets in the Console and put anything bigger in a text editor. I'd then verify the code is working ok by copying it from Emacs to the Console (and reloading the page when necessary). Once you think it's ready, go to Haketilo settings page and add it as a new "script". Then, in "pages" tab, click to add a new "page", choose the new script as its payload and type in the URL pattern [(explained here)](#) for which this replacement script should be used. Once these are added, reload the page again to see if everything works.

If you want to make the script usable in Hydrilla, just use the same format as those [here](#). Or, simpler, send it to us to do this and include the script in the repo (this way we'll also maintain it and translate to the new Hydrilla format once we introduce it)

#2 - 01/08/2022 07:55 PM - jacobk

I thought the reCAPTHCAs that relied on normal user activity had a fallback on more traditional (picture) CAPTCHAs, which I thought would mean implementing v2 would be enough in this case (or any case that uses v3).

When writing fixes, you have to run the original, proprietary, scripts in order to see the complete original network traffic, right? Do you use a different browser instance or somehow separate the activity from your normal browsing? I'm not totally sure how browser sandboxing works, but sites can't see the data of other sites, right? So maybe it doesn't matter if I use the same browser instance?

#3 - 01/10/2022 03:46 PM - koszko

- File profiles.png added

jacobk wrote:

> I thought the reCAPTHCAs that relied on normal user activity had a fallback on more traditional (picture) CAPTCHAs, which I thought would mean implementing v2 would be enough in this case (or any case that uses v3).

Actually, it's the reCAPTCHA v2 "Invisible reCAPTCHA" that relies on user activity and falls back to pictures. The reCAPTCHA v3 has no such fallback by default - it just returns a score in a [0, 1] range. 1 means "certainly human" and 0 menas "certainly bot". It is up to the site developer/owner to decide what (if any) the fallback behavior should be in case of a low score.

[link to google's reCAPTCHA docs](#)

> When writing fixes, you have to run the original, proprietary, scripts in order to see the complete original network traffic, right?

That's true. However, in many simple cases the original network traffic is not really needed to figure out what is needed to make the site work. I figured out how to make Google Drive and Sheets display (and allow download of) content by just looking at the data in pages' sources. On Odysee, on the other hand, I had to enable the original scripts and look at network traffic (and then work went smoothly).

Do you use a different browser instance or somehow separate the activity from your normal browsing? I'm not totally sure how browser sandboxing works, but sites can't see the data of other sites, right? So maybe it doesn't matter if I use the same browser instance?

AFAIK the most optimal option is to create a separate browser profile for stuff like this. In case of Ungoogled Chromium the profiles icon is at the top right:
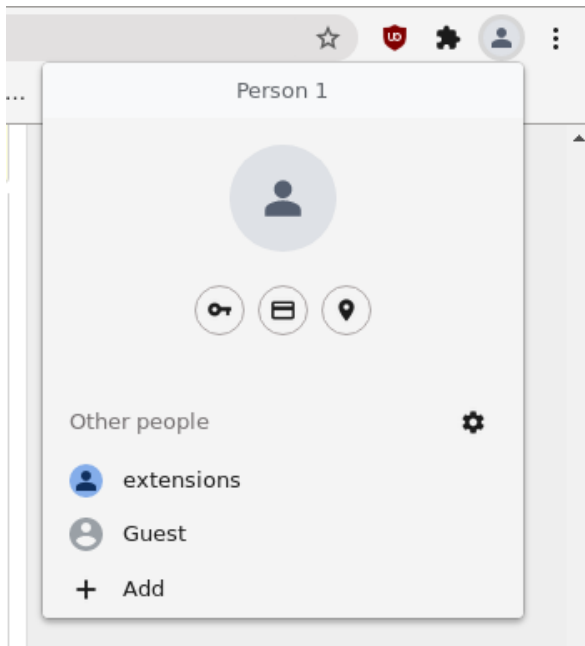
In case of Mozilla-based browsers you should go to about:profiles

#4 - 02/19/2022 11:23 AM - koszko

- Priority changed from Normal to Low

#5 - 03/03/2022 05:16 PM - koszko

- Tracker changed from Support to Site script request/donation

I just defined a new tracker "Site script request/donation" for issues like this one. There'll be no more confusion

#6 - 03/03/2022 05:21 PM - koszko

- Project changed from Haketilo to Site fixes

Files

profiles.png	18.3 KB	01/10/2022	koszko