Federal Elections Data Report

Tin Le

Dr. Reardon

CPSC502.04

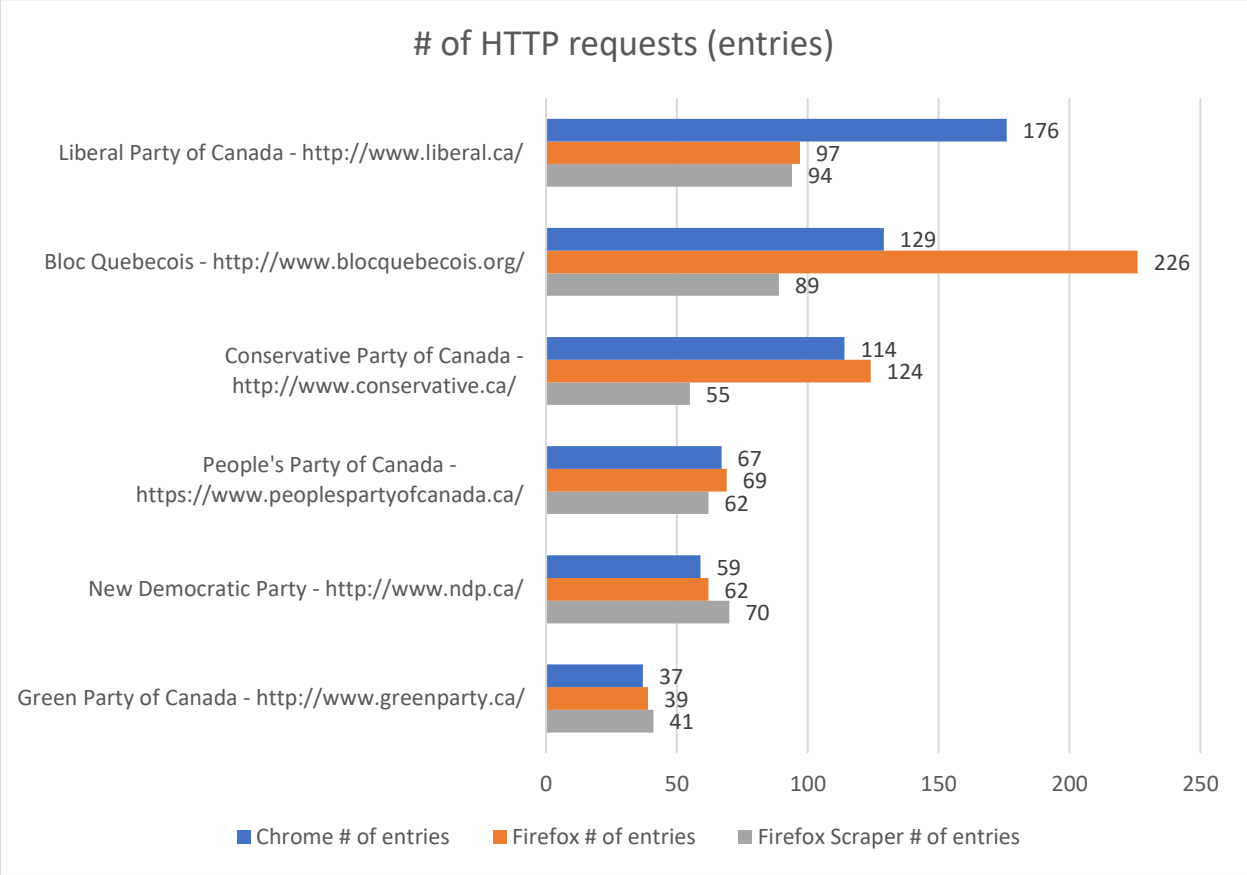
Updated 09/01/2022

Contents

General Notes

- 6 federal parties/corresponding 2021 campaign websites analyzed:
  - *Bloc Quebecois* - http://www.blocquebecois.org/
  - *Conservative Party of Canada* - http://www.conservative.ca/
  - *Green Party of Canada* - http://www.greenparty.ca/
  - *Liberal Party of Canad*a - http://www.liberal.ca/
  - *New Democratic Party* - http://www.ndp.ca/
  - *People's Party of Canada* - https://www.peoplespartyofcanada.ca/
- Chrome/Firefox HAR files for all parties were manually collected on 18/09/2021, Firefox scraper HAR files for all parties were automatically collected on 20/09/2021.
  - Only 1 set of HAR files were collected for each browser.
- Relevant Python scripts created for this part of the project:
  - **getJSONsBUP.py**: Python script that automates a clean slate Firefox browser and visits a set of websites + downloads HAR files in JSON format for each website. Used for collecting data on federal elections. Referred to as the "Firefox scraper" in this report.
  - **getJSONsChrome.py**: Same as the above script but automates a clean slate Chrome browser instead. Was not developed in time for the federal elections, and therefore wasn't used in this portion of the project.
  - **analyzeJSON.py**: Data analysis program. Used to sift through collected HAR files in JSON format and returns important information in .txt files. Also has been recently updated to:
    - detect and mark requests to known tracker sites (based on https://github.com/duckduckgo/tracker-radar/tree/main/domains)
    - detect and mark potential tracking pixels
    - generate .txt files in csv format that contains important metadata for each HAR file analyzed. The data in these csv files is then used to create spreadsheets.
  - **generateDomainsList.py**: Generates a .txt file containing all the tracker domains in the DuckDuckGo database of trackers. The resulting .txt is used by **analyzeJSON.py** to efficiently detect and mark trackers.
- **Important Issues**
  - HAR files generated by scraper program seem to not include any tracking pixel entries (no tracking pixels detected by HAR analysis program). Still trying to figure out why this is the case.
  - Occasional inconsistencies across data sets

Section 1: Number of HTTP requests



Notes/Observations:

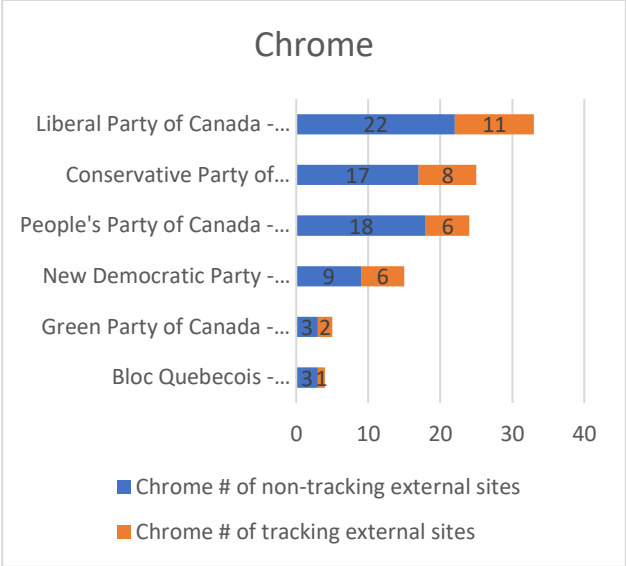
- Occasional inconsistency regarding the number of requests depending on the browser used. Examples:
    - Liberal: significantly more requests with Chrome
    - Bloc Quebecois: significantly more requests with Firefox
    - Conservative: significantly less requests with Firefox scraper program

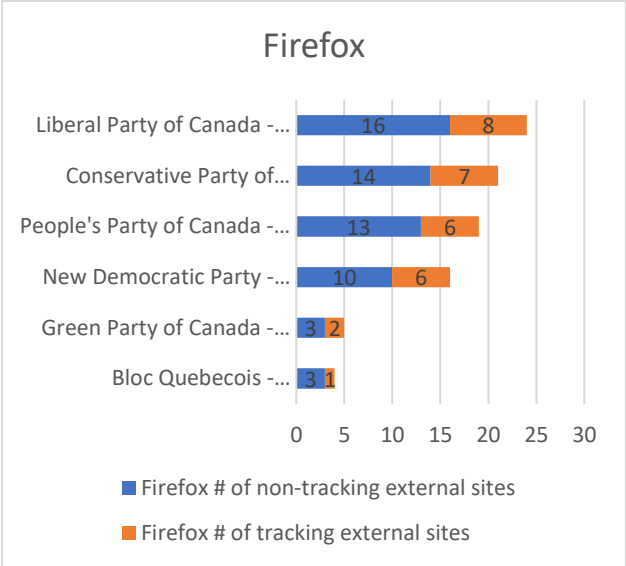
Section 2: Tracking/non-tracking external sites

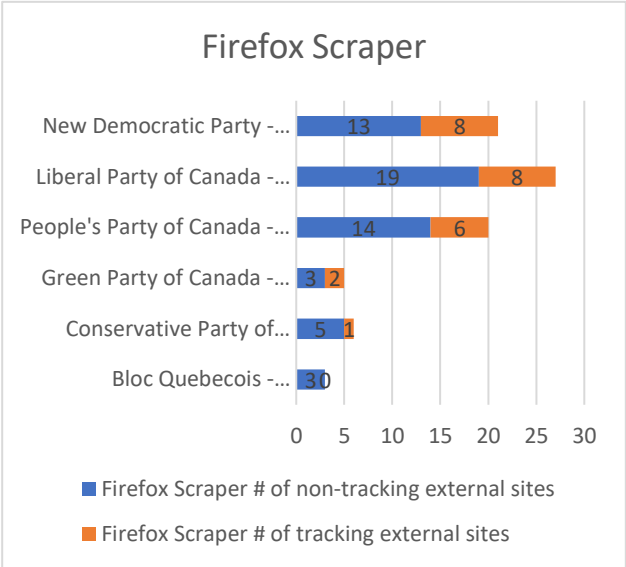
- An HTTP request's URL is considered an external site if it does not contain the party's website's hostname.
- An external site can be non-tracking or tracking. Tracking sites are based off https://github.com/duckduckgo/tracker-radar/tree/main/domains. Any external site that is found in this database is considered a tracking site.

Notes/Observations:

- Occasional inconsistencies between scraper program results and manual testing results
    - Ex. Significant reduction in Conservative external sites/tracking external sites in scraper program results. Will investigate this (is this a data collection issue or a data analysis issue?).
- Concern: some sites are not considered as a tracker.
    - Ex. Liberal makes requests to nova.collect.igodigital.com and static.ads-twitter.com, but these are not marked as trackers because they are not part of the DuckDuckGo tracker database.



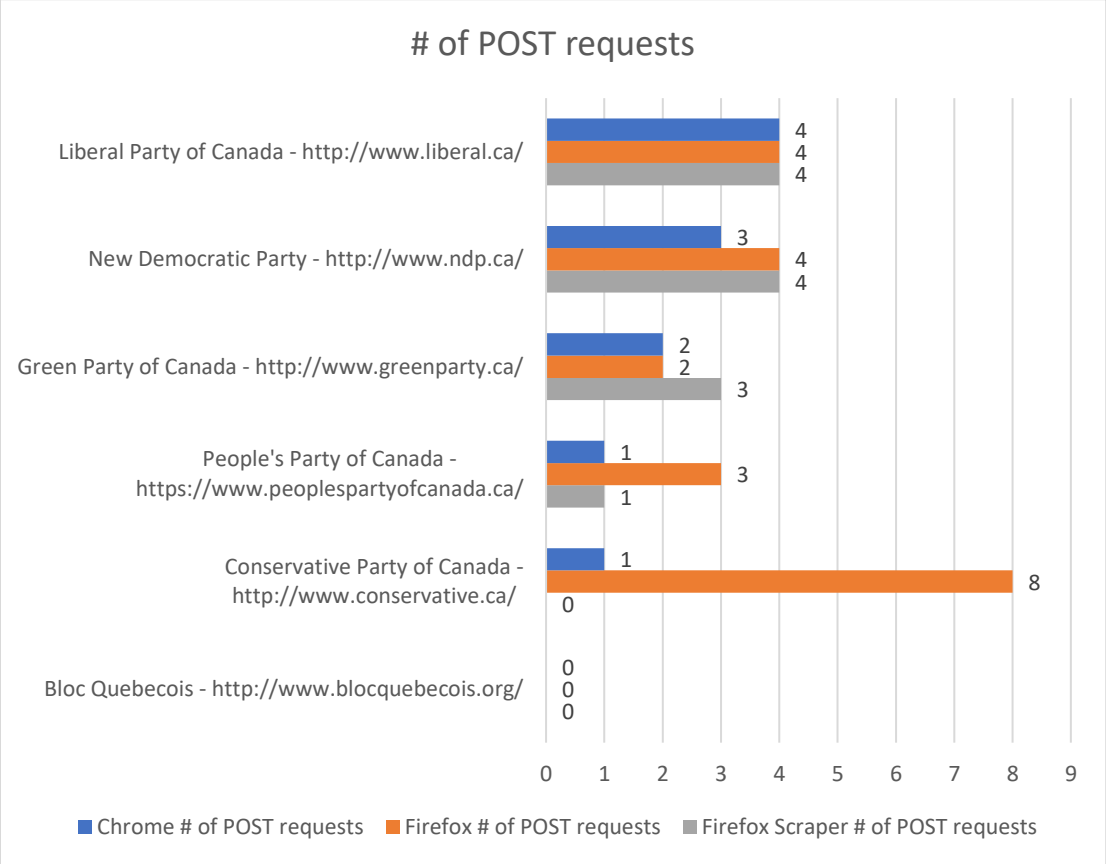




Tracking domains across all browsers (Chrome, Firefox, Firefox scraper)

Candidate - Website	Tracking domains
Liberal Party of Canada - http://www.liberal.ca/	fonts.googleapis.com googletagmanager.com youtube.com google-analytics.com googleadservices.com t.co sc-static.net facebook.com google.com google.ca gstatic.com
Bloc Quebecois - http://www.blocquebecois.org/	facebook.com
Conservative Party of Canada - http://www.conservative.ca/	fonts.googleapis.com googletagmanager.com google-analytics.com facebook.com ajax.googleapis.com youtube.com google.com gstatic.com
Green Party of Canada - http://www.greenparty.ca/	google-analytics.com facebook.com
New Democratic Party - http://www.ndp.ca/	fonts.googleapis.com googleadservices.com google-analytics.com googletagmanager.com sc-static.net facebook.com
People's Party of Canada - https://www.peoplespartyofcanada.ca/	ajax.googleapis.com d3n8a8pro7vhmx.cloudfront.net fonts.googleapis.com googletagmanager.com google-analytics.com facebook.com

Section 3: POST requests



Unique POST request URLs across all browsers (Chrome, Firefox, Firefox scraper)

- Liberal Party of Canada - http://www.liberal.ca/
  - https://tr.snapchat.com/p
  - https://www.facebook.com/tr/
  - https://www.youtube.com/youtubei/v1/log_event?alt=json&key=AIzaSyAO_FJ2SlqU8Q4STEHLGCilw_Y9_11qcW8
  - https://bat.bing.com/actionp/0?ti=5065809&tm=gtm002&Ver=2&mid=8f8548bd-6f17-4737-9a80-a6ea55b53ff5&sid=f96c90a018ce11ec8e2597e499186bc3&vid=f96c9cc018ce11ecb9d0dbc0babd0cbc&vids=1&evt=pageHide
  - https://www.youtube.com/api/stats/atr?ns=yt&el=embedded&cpn=Iv2e_Ad5ADqI95qI&docid=7lymaXqMttE&ver=2&cmt=0&fs=0&rt=0&euri=https%3A%2F%2Fliberal.ca%2F&lact=704&cl=397162147&mos=0&volume=100&cbr=Firefox&cbrver=92.0&c=WEB_EMBEDDED_PLAYER&cver=1.20210915.1.2&cplayer=UNIPLAYER&cos=Windows&cosver=10.0&cplatform=DESKTOP&hl=en_US&cr=CA&len=59&fexp=23858057%2C23983296%2C24001373%2C24002022%2C24002025%2C24002923%2C24004644%2C24004785%2C24007246%2C24080738%2C24082662%2C24096481&muted=0&vis=3
  - https://www.google-analytics.com/j/collect?v=1&_v=j93&a=118745650&t=pageview&_s=1&dl=https%3A%2F%2Fliberal.ca%2F&ul=en-us&de=UTF-8&dt=Liberal%20Party%20of%20Canada&sd=24-

bit&sr=1536x864&vp=1266x689&je=0&_u=aGDAAEADQAAAAC~&jid=824242921&gjid=1597550721&cid=790192864.1632119044&tid=UA-7046703-1&_gid=328408528.1632119044&_r=1&gtm=2wg9f0KXPGPPD&z=1963491942

    - https://stats.g.doubleclick.net/j/collect?t=dc&aip=1&_r=3&v=1&_v=j93&tid=UA-7046703-1&cid=790192864.1632119044&jid=824242921&gjid=1597550721&_gid=328408528.1632119044&_u=aGDAAEACQAAAAC~&z=2014642048

- Bloc Quebecois - http://www.blocquebecois.org/: N/A
- Conservative Party of Canada - http://www.conservative.ca/
    - https://www.youtube.com/youtubei/v1/log_event?alt=json&key=AIzaSyAO_FJ2SlqU8Q4STEHLGCilw_Y9_11qcW8
    - https://www.youtube.com/api/stats/atr?ns=yt&el=embedded&cpn=AWR0ECf4CqEPKyst&docid=bP95jGFpGGg&ver=2&cmt=0&fs=0&rt=0&euri=https%3A%2F%2Fwww.conservative.ca%2F&lact=1287&cl=397162147&mos=0&volume=100&cbr=Firefox&cbrver=92.0&c=WEB_EMBEDDED_PLAYER&cver=1.20210915.1.2&cplayer=UNIPLAYER&cos=Windows&cosver=10.0&cplatform=DESKTOP&hl=en_US&cr=CA&len=30&fexp=23858057%2C23983296%2C24001373%2C24002022%2C24002025%2C24002922%2C24004644%2C24004785%2C24007246%2C24046641%2C24080738%2C24082662%2C24096481%2C24101842&muted=0&vis=3
    - https://www.facebook.com/tr/
- Green Party of Canada - http://www.greenparty.ca/
    - https://www.google-analytics.com/j/collect?v=1&_v=j93&aip=1&a=1234112541&t=pageview&_s=1&dl=https%3A%2F%2Fwww.greenparty.ca%2Fen&ul=en-us&de=UTF-8&dt=Green%20Party%20of%20Canada%20%7C&sd=24-bit&sr=1536x864&vp=465x674&je=0&_u=QACAAAABAAAAAC~&jid=&gjid=&cid=1483909931.1632004245&tid=UA-1403066-1&_gid=82104910.1632004245&_slc=1&cd1=anonymous%20user&z=1225319316 (total of 6 similar google-analytics requests across all browsers)
    - https://stats.g.doubleclick.net/j/collect?t=dc&aip=1&_r=3&v=1&_v=j93&tid=UA-1403066-1&cid=1547214099.1632119032&jid=1334808141&gjid=1853048819&_gid=1531522319.1632119032&_u=YEBAAAAAAAAAAC~&z=906744475
- New Democratic Party - http://www.ndp.ca/
    - https://www.google-analytics.com/j/collect?v=1&_v=j93&a=1238374888&t=pageview&_s=1&dl=https%3A%2F%2Fwww.ndp.ca%2F&ul=en-us&de=UTF-8&dt=Canada%27s%20NDP&sd=24-bit&sr=1536x864&vp=465x674&je=0&_u=AACAAEABAAAAAC~&jid=&gjid=&cid=1836956990.1632004320&tid=UA-30109525-1&_gid=1681279553.1632004320&_slc=1&z=2011867269 (total of 7 similar google-analytics requests across all browsers)
    - https://tr.snapchat.com/p
    - https://stats.g.doubleclick.net/j/collect?t=dc&aip=1&_r=3&v=1&_v=j93&tid=UA-30109525-1&cid=1805177928.1632119062&jid=317824344&gjid=533943130&_gid=908443977.1632119063&_u=IADAAEACQAAAAC~&z=542161259
- People's Party of Canada - https://www.peoplespartyofcanada.ca/
    - https://v1.addthisedge.com/live/prender?sid=614669481a635a8d&uid=6146693d75bef1f6&ids=8796-8788-8822-8797&rev=v8.28.8-wp&delay=0
    - https://m.addthis.com/live/red_lojson/100eng.json?sh=0&ph=2563&ivh=689&dt=6310&pdt=209&ict=&pct=0&perf=widget%7C209%7C186%2Clojson%7C660%7C166%2Csh%7C663%7C36&rndr=render_toolbox%7C753&cmenu=null&ppd=0&ppl=0&fbe=&xmv=0&xms=0&xmlc=0&jsfw=jquery%2Cgoogleanalytics%2Cfacebook&jsfwv=jquery-3.4.0%2Cgoogleanalytics-ga.js%2Cfacebook-

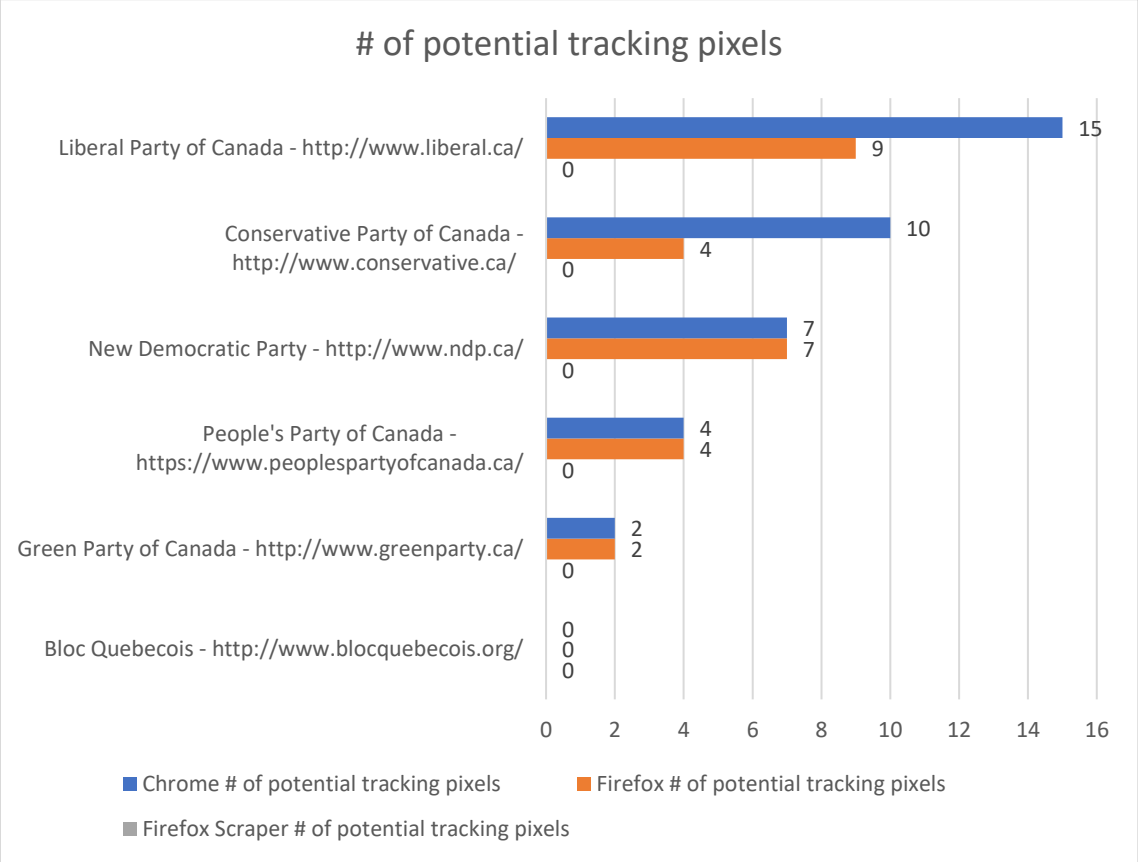
unk&al=men&scr=0&scv=0&apiu=1&ba=3&sid=614667847ee7ccff&rev=v8.28.8-wp&pub=ra-603933391e00bb49&dp=www.peoplespartyofcanada.ca&fp=&pfm=0&icns=

- https://www.google-analytics.com/g/collect?v=2&tid=G-GNV7V8CQ6V&gtm=2oe9f0&_p=1588826587&sr=1536x864&ul=en-us&cid=689828172.1632003973&_s=1&dl=https%3A%2F%2Fwww.peoplespartyofcanada.ca%2F&dt=People%27s%20Party%20of%20Canada&sid=1632003979&sct=1&seg=0&en=user_engagement&_fv=1&_ss=1
- https://www.google-analytics.com/g/collect?v=2&tid=G-GNV7V8CQ6V&gtm=2oe9f0&_p=1710696476&sr=1536x864&ul=en-us&cid=689828172.1632003973&_s=1&dl=https%3A%2F%2Fwww.peoplespartyofcanada.ca%2F&dt=People%27s%20Party%20of%20Canada&sid=1632003979&sct=1&seg=0&en=scroll&epn.percent_scrolled=90
- https://www.google-analytics.com/j/collect?v=1&_v=j93&a=750885277&t=pageview&_s=1&dl=https%3A%2F%2Fwww.peoplespartyofcanada.ca%2F&ul=en-us&de=UTF-8&dt=People%27s%20Party%20of%20Canada&sd=24-bit&sr=1536x864&vp=1266x689&je=0&_u=YEBAAEABAAAAAC~&jid=345732949&gjid=877021850&cid=1442580662.1632119070&tid=UA-127838477-4&_gid=472265440.1632119070&_r=1&gtm=2wg9f0P2KNSXV&z=56606457

Unique POST request domains across all browsers (Chrome, Firefox, Firefox scraper)

Candidate - Website	POST request domains
Liberal Party of Canada - http://www.liberal.ca/	tr.snapchat.com www.facebook.com www.youtube.com bat.bing.com www.google-analytics.com stats.g.doubleclick.net
Bloc Quebecois - http://www.blocquebecois.org/	N/A
Conservative Party of Canada - http://www.conservative.ca/	www.youtube.com www.facebook.com
Green Party of Canada - http://www.greenparty.ca/	www.google-analytics.com stats.g.doubleclick.net
New Democratic Party - http://www.ndp.ca/	www.google-analytics.com tr.snapchat.com stats.g.doubleclick.net
People's Party of Canada - https://www.peoplespartyofcanada.ca/	v1.addthisedge.com m.addthis.com www.google-analytics.com

Section 4: Tracking pixels



Notes/Observations:

- Any HTTP response with a mimeType containing the string “image” and a content size above 0 and below 100 is marked as a potential tracking pixel by the data analysis program.
- Potential tracking pixel requests will be investigated to confirm whether they actually are tracking or not. Details of those that are will be included in the final report.
- Potential tracking pixels are never detected in HAR files generated by the Firefox and Chrome scraper programs, even when manual collection of a HAR file for the same website indicates that tracking pixels are present. This has also been observed in the data for the Calgary municipal elections. Will investigate this (is this a data collection issue or a data analysis issue?).

Overall observations and next steps

- Bloc Quebecois seems fairly clean. Liberal seems to track the most, and Conservative and PPC also track a fair bit.
- The graphs above and a similar report will be created for the Calgary municipal elections, covering candidates' websites for each of the 4 positions.
    - The final report will include more details for each tracking instance (ex. Info about each tracking site, personal information that is collected by each tracking pixel)
- Occasionally, there are inconsistencies with results depending on whether the data was collected manually using Chrome, Firefox, or automatically using the scraper. I suggest that for the next set of websites to be analyzed, we collect multiple sets of data to avoid issues with outliers.
- Must investigate the issue with my programs not detecting any tracking pixels in the HAR files generated by the scrapers. It is possible that automatic collection with the scraper somehow affects the data: manually collected HAR files are much larger than the automatically collected ones. Will try to figure out the reason for this.