Question 1:

Describe your experience creating your collection with the four different tools.
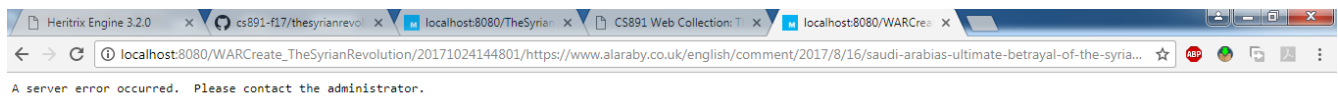
Answer:

I was able to create my collection very easily using Archive-it. The tool is great because it automatically crawls and collects so much data, much more than needed if the crawler is set to "standard+". It was so easy to add seeds, just copy all seeds and paste them.

WebRecorder is a good tool. Archiving is made easy with WebRecorder. It records all pages that have been navigated to and downloaded. I liked Archive-it better because it was faster, and pages did not need to be manually navigated and downloaded to be archived.

WARCreate is similar to WebRecorder in one way; the page has to be manually navigated in order to be archived, WARC file saved. The page is one click away from being archived, but it did not work on a couple of pages. It took a little bit to respond and actually download the WARC file. The extension is easy to install on Google Chrome. Just click one button to download the WARC file. It is very straight forward and the files can be later replayed in other software like WAIL.

I was able to create collections and add seeds to collections using WAIL. I also ran the crawler, and it showed a success message, but no pages were downloaded on the first crawl; I am not sure why. I was finally able to get it to work properly and download pages. I was also able to import WARC files I downloaded using WARCreate, however, the browser generated a server error when trying to reply the pages, view in wayback.

Question 2:

Compare and contrast the tools in terms of capture quality, ease of use, etc.

Answer:

Archive-it: By far, the most effective and pleasing tool out of the four tools I tried. At first, it feels like it is difficult to use because of its settings that are needed to run a crawl, however, once you get the lingo of web archiving and master the settings, it is rather easy.

Capture quality is good, but WebRecorder is better. It captures lots of pages if set to Standard+ crawling.

WebRecorder: The tool is neat and clean. Easy to use, however, the page has to be navigated and downloaded completely in order to be archived.

Capture quality is the best out of all four tools I have tried.

WARCreate: The extension is super easy to add to Google Chrome, and it costs one click to save the WARC file for any opened page on the computer. I was able to save the WARC files I needed, but I was not able to replay them in WAIL. The browser showed a server error. It would be nice if a success/error message was displayed because the extension did not work on a couple of links, and was slow to respond on all of them.

WAIL: It has a user friendly GUI, and it is easy to use, for the most part. I was able to add collections and seeds easily, however, I ran the crawler the first time and it showed a success message, but no pages were archived. The second crawl was successful.

Crawl URL(s)	Status	For Collection	Timestamp	Discovered	Queued	Downloaded	Actions
http://mashable.com/2017/03/13/cries-from-syria/#_Ej1wCXV5qqE	Ended	TheSyrianRevolution	Oct 24 2017 10:23am	1422	114	1307	⋮
http://www.atlanticcouncil.org/blogs/syriasource/how-jabhat-al-nusra-hijacked-the-syrian-revolution	Ended	TheSyrianRevolution	Oct 24 2017 10:31am	1027	130	917	⋮
https://www.alaraby.co.uk/english/comment/2017/8/16/saudi-arabias-ultimate-betrayal-of-the-syrian-revolution	Running	TheSyrianRevolution	Oct 24 2017 10:30am	0	0	0	⋮
http://isj.org.uk/interview-lessons-of-the-syrian-revolution/	Ended	TheSyrianRevolution	Oct 10 2017 1:24am	0	0	0	⋮
https://www.thenation.com/article/report-from-deraa-cradle-of-the-syrian-revolution/	Ended	TheSyrianRevolution	Oct 10 2017 1:24am	0	0	0	⋮
http://foreignpolicy.com/2016/03/09/the-syrian-revolution-is-not-a-holy-war/	Ended	TheSyrianRevolution	Oct 10 2017 1:24am	0	0	0	⋮
https://www.globalsecurity.org/military/world/war/syria.htm	Ended	TheSyrianRevolution	Oct 10 2017 1:24am	0	0	0	⋮
https://gowans.wordpress.com/2016/10/22/the-revolutionary-distemper-in-syria-that-wasnt/	Ended	TheSyrianRevolution	Oct 10 2017 1:23am	0	0	0	⋮
https://www.ft.com/content/f2d9bbc8-bdbc-11e2-890a-00144feab7de	Ended	TheSyrianRevolution	Oct 10 2017 1:23am	0	0	0	⋮
https://www.theatlantic.com/international/archive/2015/10/syrian-civil-war-guide-isis/410746/	Ended	TheSyrianRevolution	Oct 10 2017 1:23am	0	0	0	⋮
https://www.counterpunch.org/2016/08/26/anti-imperialism-and-the-syrian-revolution/	Ended	TheSyrianRevolution	Oct 10 2017 1:23am	0	0	0	⋮
http://www.dsausa.org/the_case_for_solidarity_with_the_syrian_revolution_dl	Ended	TheSyrianRevolution	Oct 10 2017 1:22am	0	0	0	⋮
http://www.aljazeera.com/indepth/opinion/2016/12/facebook-hurt-syrian-revolution-161203125951577.html	Ended	TheSyrianRevolution	Oct 10 2017 1:22am	0	0	0	⋮
http://www.bbc.com/news/world-middle-east-26116868	Ended	TheSyrianRevolution	Oct 10 2017 1:22am	0	0	0	⋮
https://en.wikipedia.org/wiki/Syrian_Revolution_Network	Ended	TheSyrianRevolution	Oct 10 2017 1:21am	0	0	0	⋮

RESCAN JOB DIRECTORY ≡✓   LAUNCH WEB UI ⬆

Capture quality is much less than Archive-it and WebRecorder. The page only displayed text and images. All formatting is lost.
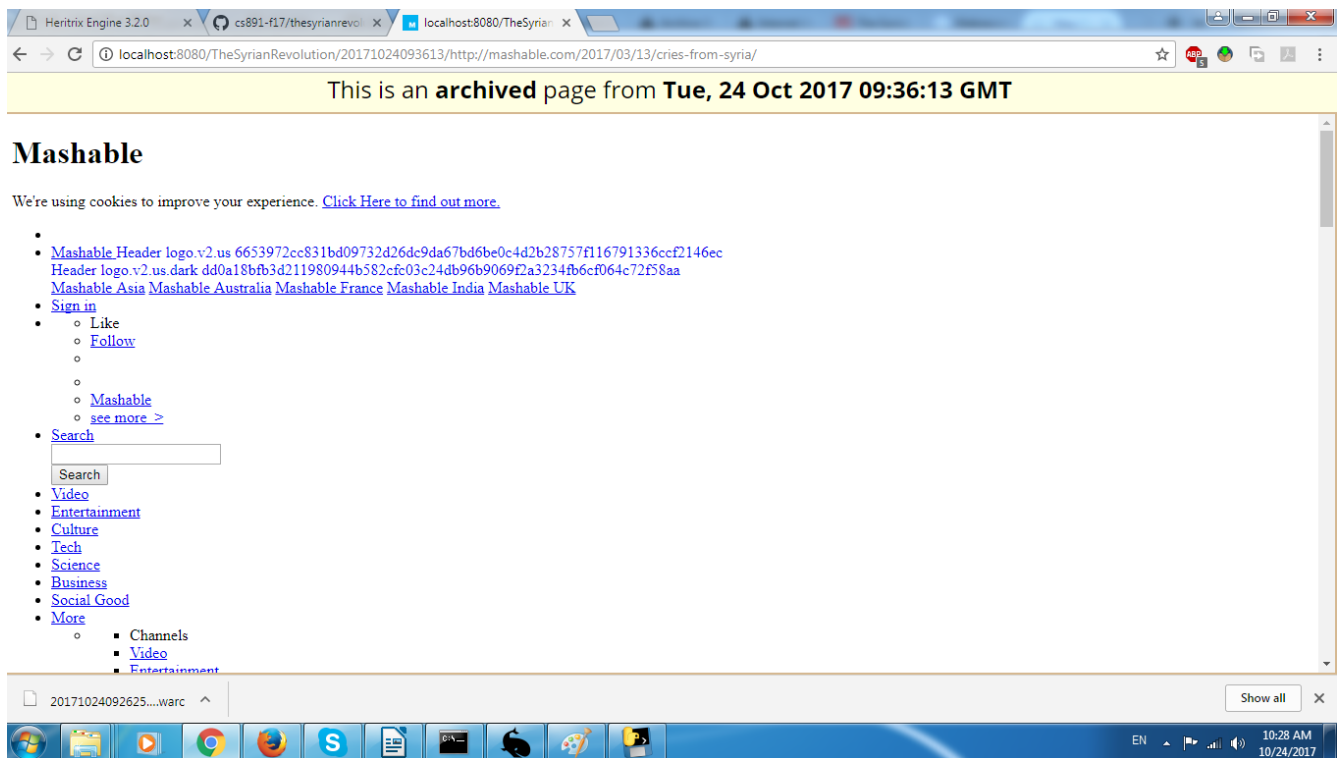
Question 3:

For each tool list

- advantages
- disadvantages
    - include at least one screenshot where the tool did not capture all of a page's resources (if applicable)
- situations or types of pages where you'd recommend this tool over the others
    - include at least one screenshot of where the tool captured a page better than the other tools (if applicable)

Answer:

Archive-it:

Advantages:

Easy to use once the settings and configurations are understood

No need to manually navigate pages to archive them (auto crawling)

Disadvantages:

Settings and configurations for crawling must be set before crawling to get the best results

Some images were missing on some archived web pages, but these images were not part of the main subject of the page (from related articles, videos, ...etc).

In the screen shot above, few images were missing, but the main image in the article is not.
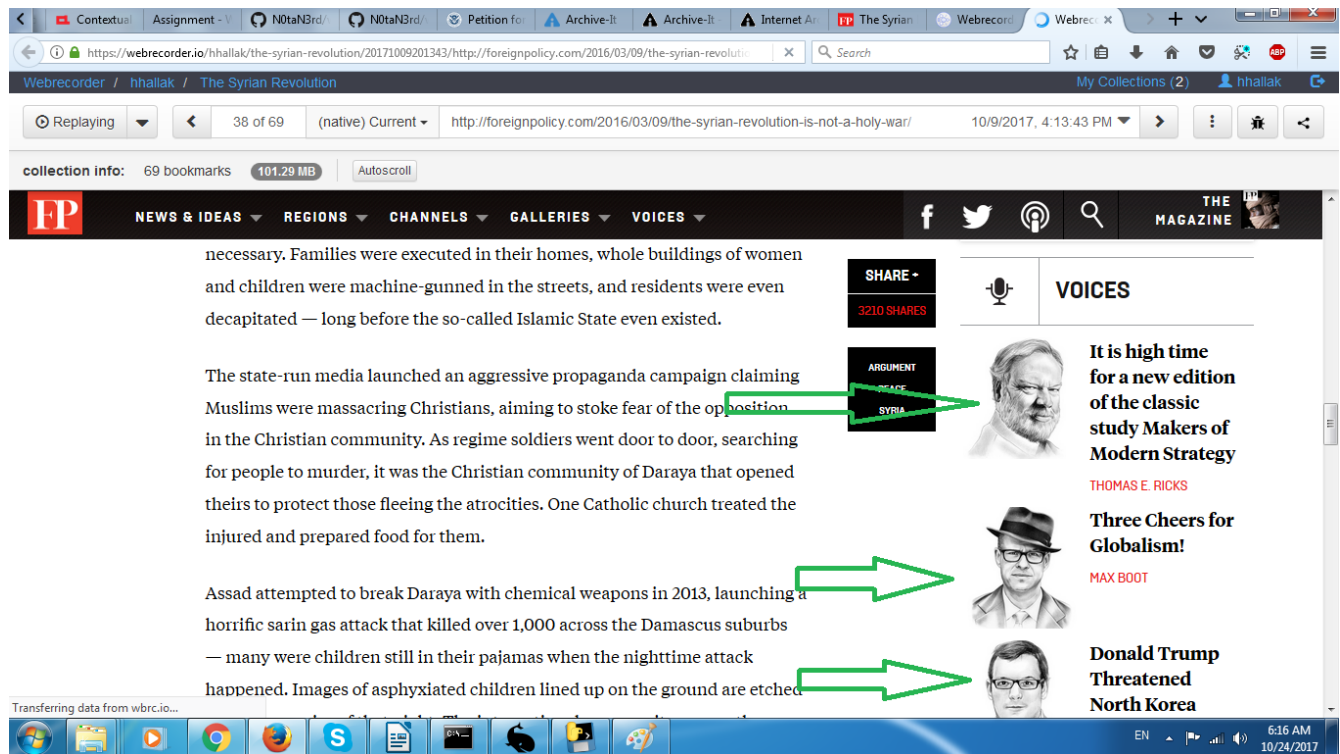
WebRecorder:

Advantages:

Easy to use

User friendly

High quality captures

In comparison to Archive-it, WebRecorder was able to capture all images in the page.

Disadvantages:

The page must be completely downloaded to be archived

The page must be manually browsed to be recorded

WARCreate:

Advantages:

Easy to install on Google Chrome

Easy to use

Produced warc file that can be saved and replayed

Disadvantages:

The page must be completely downloaded to be archived

Slow response

The page must be manually browsed to be archived

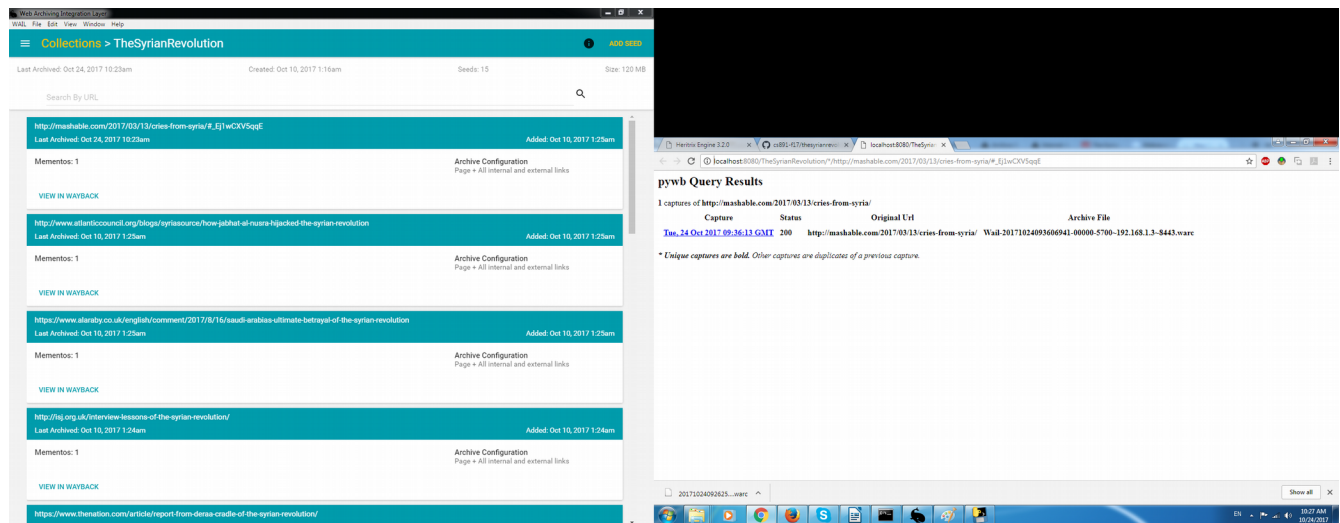The need for another software to replay saved warc files

WAIL:

Advantages:

Easy to use

User friendly

Gives good feedback at runtime (error and success messages)

Disadvantages:

Inability to add multiple seeds at a time. Seeds must be added one by one.

Inability to import multiple WARC files. The files must be imported one by one.

Formatting was not archived. It could be CSS issue? Or may be WAIL does not save CSS by design?

Annoying terminal windows open and close during runtime

Note: I found WebRecorder good for pages that the user navigated and found that they are important to him/her. WebRecorder captures the page as it is without missing hardly any content.

On the other hand, Archive-it is better if the user wants to capture lots of pages he/she did not navigate yet, or does not have the time to navigate and archive at this moment. Just plug the seeds in, and let the crawler do the work!

Question 4:

Reflect upon your collection. Were 15 seeds enough to characterize the topic you wanted to cover?

Answer:

I have chosen this topic very carefully. I am interested in it, and have a good background about it. I picked my seeds from different sources, so the topic is covered from all different angles. The 15 seeds were enough if someone wants to get a general idea about the Syrian revolution. Of course, more seeds are needed and more analysis and research are required if someone wants to know every detail about it, and may be put together a time line for the Syrian revolution and how it turned into a civil war.

Question 5:

How could someone use your collection in the future?

Answer:

This collection is good if someone is interested in conducting a research about the revolution and educating others. A time line for the revolution from 2011 to 2017 would be something useful.

Question 6:

Identify (include screenshots and links) at least one resource that changed over the time you were collecting.

Answer:

The section "Trending" in this URI was completely changed over the time I was collecting:

http://www.aljazeera.com/indepth/opinion/2016/12/facebook-hurt-syrian-revolution-161203125951577.html
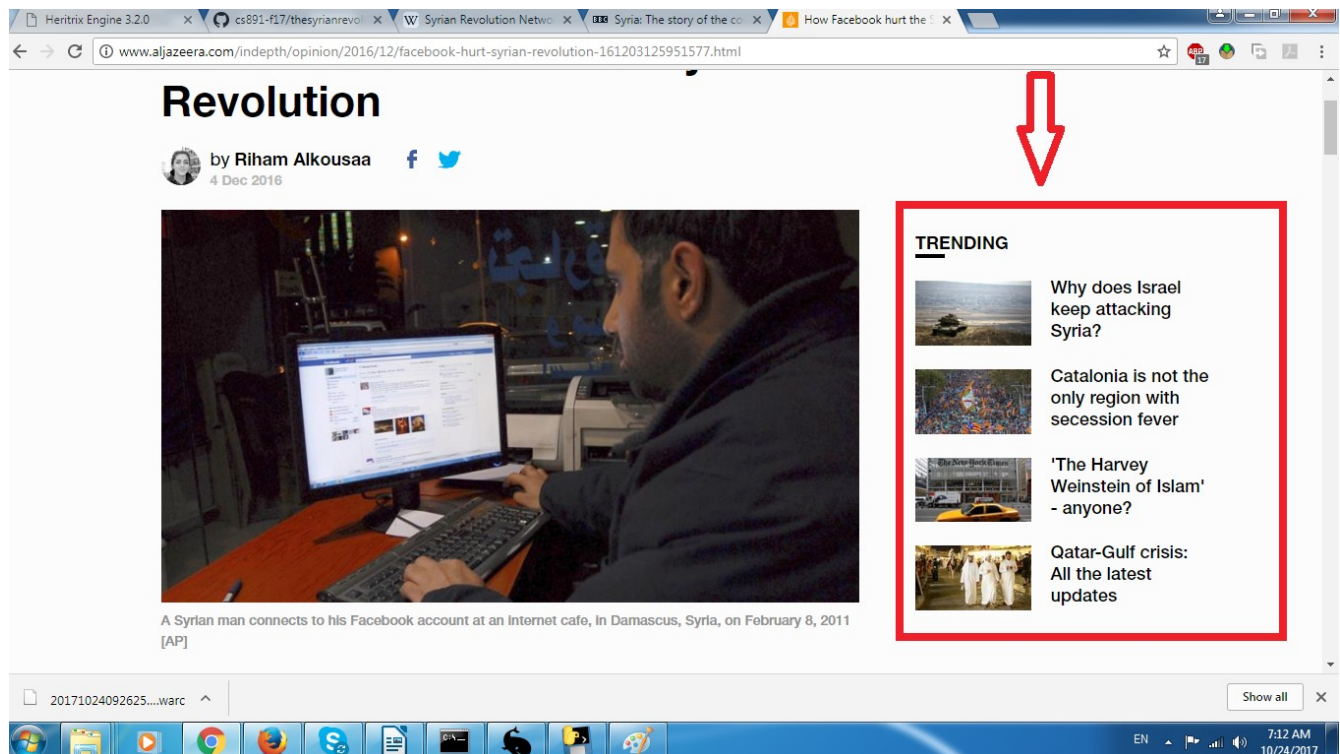
Here is the page I recorded:

https://webrecorder.io/hhallak/the-syrian-revolution/20171009180312/http://www.aljazeera.com/indepth/opinion/2016/12/facebook-hurt-syrian-revolution-161203125951577.html