

Berk Emre Mert
bmert24@student.oulu.fi

Week 1: Regulation and privacy on the web

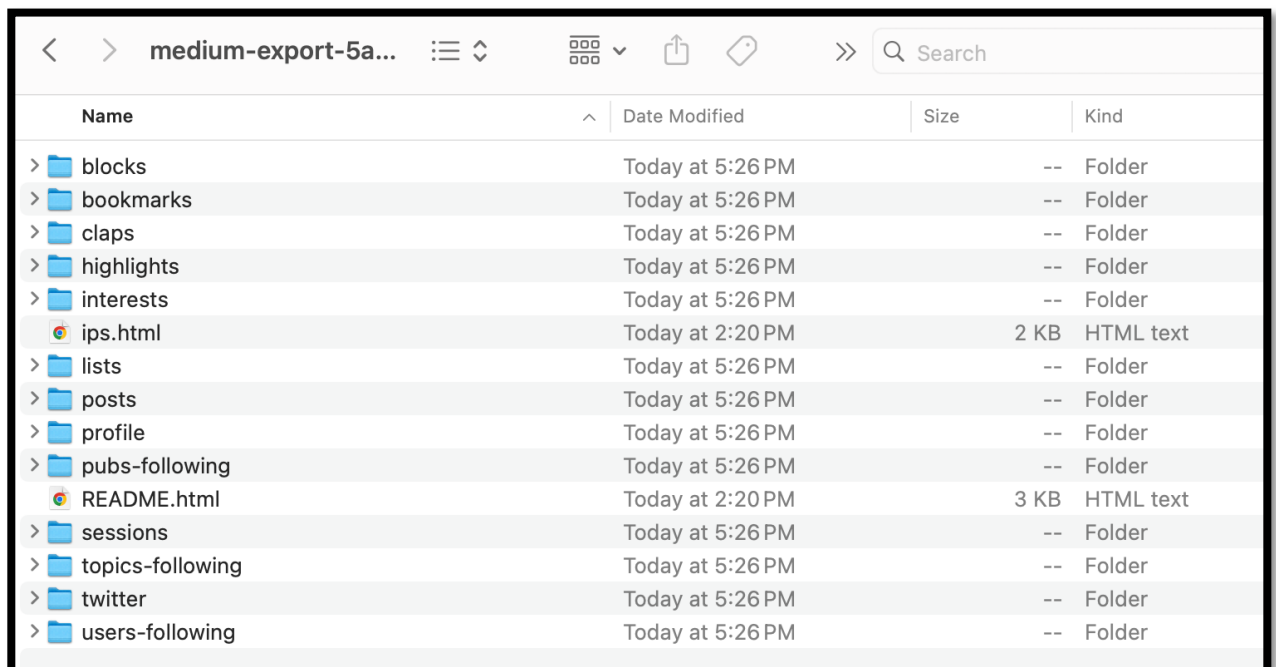
Task 1: GDPR Data Request

1. Which company did you choose to request your data from and how long did it take to submit that information?

I chose Medium.com, it's a blog website for both personal and profitable use. I emailed them and there was an automated response first, then I created my request again via the link they provided with that automated e-mail. After my request, they responded via e-mail within an hour.

2. In what format did you receive your data and how many pages/bytes of it were there?

They sent me .zip file including the following folders:



Name	Date Modified	Size	Kind
> blocks	Today at 5:26 PM	--	Folder
> bookmarks	Today at 5:26 PM	--	Folder
> claps	Today at 5:26 PM	--	Folder
> highlights	Today at 5:26 PM	--	Folder
> interests	Today at 5:26 PM	--	Folder
ips.html	Today at 2:20 PM	2 KB	HTML text
> lists	Today at 5:26 PM	--	Folder
> posts	Today at 5:26 PM	--	Folder
> profile	Today at 5:26 PM	--	Folder
> pubs-following	Today at 5:26 PM	--	Folder
README.html	Today at 2:20 PM	3 KB	HTML text
> sessions	Today at 5:26 PM	--	Folder
> topics-following	Today at 5:26 PM	--	Folder
> twitter	Today at 5:26 PM	--	Folder
> users-following	Today at 5:26 PM	--	Folder

The files in those folders are in .html format.

3. What kind of information was there that you had not directly given yourself?

- IP addresses I used

- My sessions (My active and inactive sessions)

- My X friends who are also on Medium

4. Did you understand the meaning of the information?

Yes it was very simple and well-organized. They also provided me a roadmap to explain which file includes which informations.

Archive for berkemremert

Folder Structure

Medium gives you the ability to export your personal data and stories as HTML files in a .zip archive.

Each folder will contain one or more HTML files with available data.

/blocks

- Blocked users

/bookmarks

- Bookmarked Posts

/claps

- Posts you've clapped for

/highlights

- Posts you've highlighted

/interests

5. Do you think that the company collected the least amount of data required to provide the service you needed?

I think they are very open and transparent when it comes to keep the data. I read their privacy page on the website as well and I can say that they don't keep any unnecessary information that bothered me.

6. Were you surprised about something, what?

I was surprised about the IP addresses because I didn't know that they were keeping that information however I understood why they do that afterwards. I looked up for it and found that it's to protect themselves against syber attacks, also they understand if I am logging in from a new IP and asks an authentication code, so they protect my account as well.

Task 2: Cookie Raid

- How many necessary cookies are there?

71

- What happens if you try to disable necessary cookies?

It gives the following error:

Mandatory - can not be deselected. Necessary cookies help make a website usable by enabling basic functions like page navigation and access to secure areas of the website. The website cannot function properly without these cookies.

- List at least 6 different expiry times from the necessary cookies

CookieConsentBulkSetting-# [x5]
Enables cookie consent across multiple websites
Maximum Storage Duration: Persistent
Type: HTML Local Storage
1.gif
Used to count the number of sessions to the website, necessary for optimizing CMP product delivery.
Maximum Storage Duration: Session
Type: Pixel Tracker
test_cookie
Used to check if the user's browser supports cookies.

Maximum Storage Duration: 1 day
Type: HTTP Cookie
datadome
Used in context with the website's BotManager. The BotManager detects, categorizes and compiles reports on potential bots trying to access the website.
Maximum Storage Duration: 1 year
Type: HTTP Cookie
JSESSIONID [x2]
Preserves users states across page requests.
Maximum Storage Duration: Session
Type: HTTP Cookie
jagex-support-data
Used to locally store a copy of current announcement articles to reduce the necessary rate of requests
Maximum Storage Duration: Persistent
Type: HTML Local Storage

- List 3 different cookie types from necessary cookies
 - HTTP Cookie
 - HTML Local Storage
 - Pixel Tracker
- In Statistics, what are the expiry time and cookie type of **JXWEBUID**

JXTRACKING
An identifier we use to track the pages you visit on our websites. We use the information to make improvements to the website.
Maximum Storage Duration: Session

Type: HTTP Cookie

- In Statistics and Marketing, what other Apps or Companies are listed, other than Jagex
 - Google
 - Framer
 - Taboola
 - Tiktok
 - Reddit
 - Meta

Task 3: Browser Fingerprinting

Task 3 A) EFF Cover Your Tracks

- Return two categories and results where you had the *highest* and *lowest* amount of identifying bits of information.
 - Hash of canvas fingerprint (9,77)
 - Hash of WebGL fingerprint (10,08)
- Why does identifiability matter?

Identifiability matters because it can compromise privacy and security, making it easier for attackers to target individuals or misuse data.
- What pros and cons can you think of to being unique vs common?

Pros of being unique:

 - Stands out and attracts attention.

- Offers a sense of individuality and originality.
- Can lead to innovation and creative solutions.

Cons of being unique:

- May face resistance or rejection from others.
- Could be isolating or misunderstood.
- Might require more effort to maintain or explain.

Pros of being common:

- Easier to fit in and gain acceptance.
- Less likely to face criticism or resistance.
- Can offer stability and familiarity.

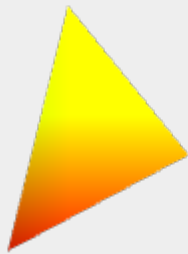
Cons of being common:

- May feel unremarkable or overlooked.
- Limited opportunities for standing out.
- Can lead to a lack of differentiation or uniqueness.

Task 3 B) AmlUnique

- Canvas + WebGL Data

Canvas	0.05 %	Cwm fjordbank glyphs vext quiz, 😊 Cwm fjordbank glyphs vext quiz, 😊
--------	--------	--

WebGL Data	0.21 %	
------------	--------	--

- Screen width, Screen height

Screen width	6.36 %
Screen height	5.65 %

- Screen available width, Screen available height, Can you explain your result?

Screen available width 6.49 % 1440

-Returns the amount of horizontal space in pixels available to the window.

Screen available Height 0.07 % 825

-Specifies the height of the screen, in pixels, minus permanent or semipermanent user interface features displayed by the operating system, such as the Taskbar on Windows.

Task 4: TikTok Challenge (bonus)

Task 4 A) Analysing the HAR file

-**Browser:** Safari 16.3

-Operating System: macOS 10.15.7

-Screen Size: The screen size is not directly available in the logs. This information is typically accessed via JavaScript using `window.screen.width` and `window.screen.height`, which is not provided here.

-Local Language: The local language is not explicitly mentioned, but the presence of URLs like `default.esm.js` and `captcha.js` suggests the user might be using English or another common language supported by TikTok.

-Browser Language: English (Great Britain)

-Region: The region can be inferred from the URLs and domains:

- The use of `tiktok-web-us` in the URLs suggests the user is accessing the US version of TikTok.
- Domains like `byteoversea.com` and `ttwstatic.com` are commonly associated with TikTok's global services, but the `us` in `tiktok-web-us` points to the US region.

-Timezone: The timezone is not explicitly mentioned in the logs. This information is typically accessed via JavaScript using `Intl.DateTimeFormat().resolvedOptions().timeZone`, which is not provided here.

-Search Keyword: pellesecurity

Task 4 B) TikTok data obfuscation

- 1) I opened TikTok on Firefox Developer Edition and from Developer Monitor / Network I copied the following cURL:

```
curl 'https://mcs-va.tiktokv.com/v1/list' -X POST -H 'User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10.15; rv:137.0) Gecko/20100101 Firefox/137.0' -H 'Accept: */*' -H 'Accept-Language: en-US,en;q=0.5' -H 'Accept-Encoding: gzip, deflate, br, zstd' -H 'Content-Type: application/json; charset=utf-8' -H 'Referer: https://www.tiktok.com/' -H 'Origin: https://www.tiktok.com' -H 'Sec-Fetch-Dest: empty' -H 'Sec-Fetch-Mode: cors' -H 'Sec-Fetch-Site: cross-site' -H 'Connection: keep-alive' -H 'TE: trailers' --data-raw '["events":[{"event":"window_size_change","params":{"page_name":"","others_homepage":"","enter_from":"","others_homepage":"","time_from_origin":92354,"is_landing_page":0,"previous_page":"","search_user":"","page_url":"","https://www.tiktok.com/@pellesecurity","userAgent":"Mozilla/5.0 (Macintosh; Intel Mac OS X 10.15;
```


After encoding, I got the following Headers:

Content-Type: application/json; charset=utf-8
Content-Length: 8
X-Tt-Logid: 20250311180126803529321A19A7E69409
Access-Control-Allow-Origin: https://www.tiktok.com
Access-Control-Max-Age: 1800
Access-Control-Allow-Credentials: true
Access-Control-Allow-Methods: GET, OPTIONS, HEAD, PUT, POST
Cross-Origin-Resource-Policy: cross-origin
Cache-Control: no-store, no-cache, must-revalidate
Pragma: no-cache
Expires: 0
x-tt-trace-host:
019aaaf2b44d18f0ce0f0299326d0f5b2e430bafa4d539a359f3b14fc89ac23e59367a
9403ef5a0b6cd082d4120586b837d86e07e566589696d5dabb821d36b4c69c591ba
6a710d376c0b3085b98fa4ccfa08454e04ed098322ef111102987d2f95d2c74050e2c
b0d08490024789b5f2f4
x-tt-trace-id: 00-250311180126803529321A19A7E69409-42658BC82FC22D75-00
Server: TLB
X-Origin-Response-Time: 5,23.48.249.166
Date: Tue, 11 Mar 2025 18:01:27 GMT
Connection: keep-alive
x-tt-trace-tag: id=16;cdn-cache=miss;type=dyn
Server-Timing: cdn-cache; desc=MISS, edge; dur=12, origin; dur=5
Server-Timing: inner; dur=2
X-Parent-Response-Time: 16,23.195.36.70