

Blacklight

archiver

sponsored by **The Markup**

work done by team ultraviolet beam

Meet the Team!



Aidan Bartlett
Backend / Flex



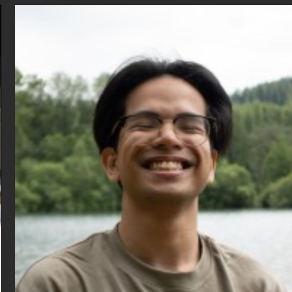
Jordan Lee
Front/Backend, Flex



See Park
Front/Back-end



Suraj Gangaram
PM / Flex



Addison Mercado
UX Design/ Front-End

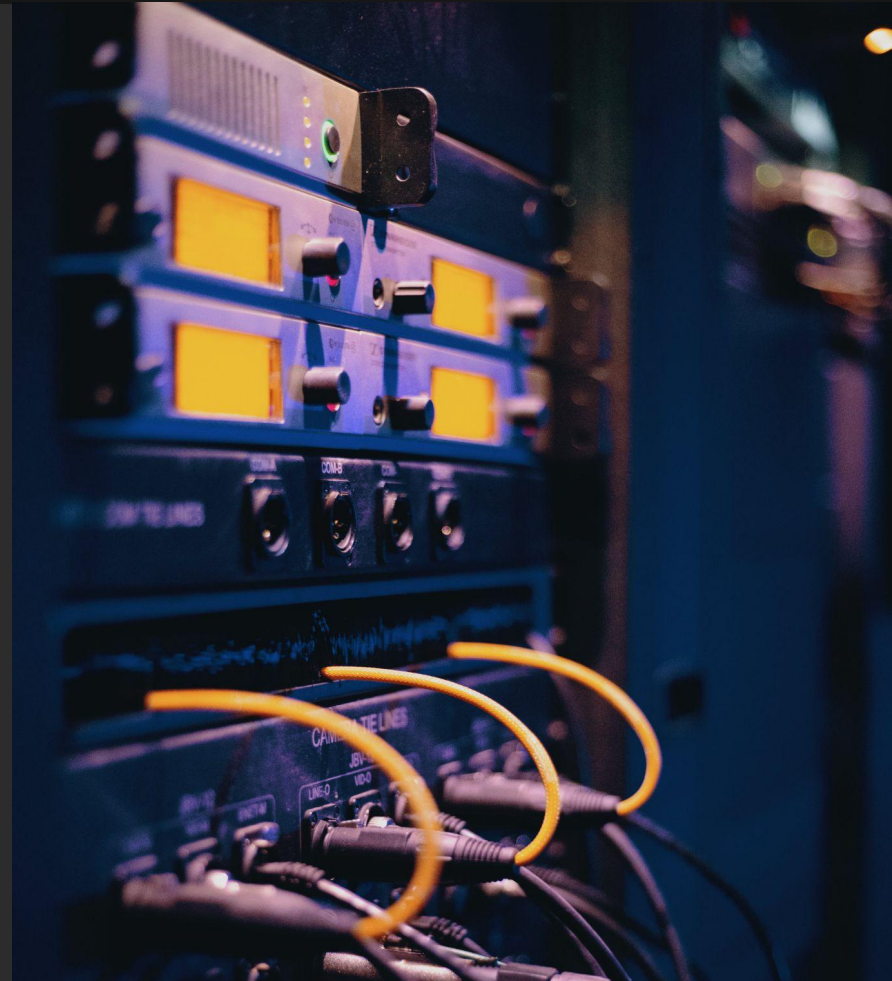
Sponsor + Project Overview

What is Blacklight?

- A custom-built tool created by **The Markup**, a nonprofit news organization challenging technology 'to serve the public good', to analyze websites for **privacy violations** and **data leakage**
- Scans websites for third-party **trackers, cookies**, and other **tracking technologies** that may **compromise user privacy**
- Provides users with insights into how websites **collect** and **share** data, aiming to raise awareness about online privacy issues

Problem Context

- Previous user research showed that users wanted **the ability to compare previous scans**
- **Internet privacy** is of growing importance as data-selling becomes **more rampant**
- Opportunity to **improve** and **expand** on Blacklight's functionality



Problem Statement

How might we **enable** internet users who are curious/concerned about their privacy to measure the **extent** of how certain websites **track** people's data over time so that they can make **better-informed decisions** about their digital privacy?

Key Research Insights



Literature Research

99.5% chance that a user will become tracked by **all top 10 trackers within 30 clicks** on search results



Market Research

There is a **variety of tools** out there to **combat** or **inform** about third-party tracking; however **each one does it unique ways**



User Research

From our survey, **83.3% care about online privacy**

Personas

Letty | Investigative Journalist



- She writes for a non-profit news organization that does investigations into how corporations and government affect the public good. She wants to find data and uncover a big story that will make waves and lead to meaningful change.

Priya | PhD Student



- She is a graduate student at a large university studying cybersecurity with an emphasis on data privacy. She wants to find original research in her field for her PhD dissertation.

Solution Approach + Key Features

Solution Approach

Our project seeks to augment the existing Blacklight tool to enable users to compare current and previous 'scans' for a website, so that users are cognizant of how companies manage their data and are better equipped to take steps to secure their data

Goals:

- *Develop a website-scanning tool to empower internet users to understand and monitor how websites track their data over time.*
- *Align with The Markup's mission of digital transparency and privacy advocacy by exposing big tech companies' data practices.*

Scan Comparison

Website Address Options

Visited on [website] on _____ Learn More ▾

Blacklight Inspection Results

Blacklight works by visiting each website with a headless browser running custom software built by The Markup. To learn more, read our [methodology](#).

[Compare Previous Scans](#)

× Choose two scans to compare

< February >							< March >							
S	M	T	W	T	F	S	S	M	T	W	T	F	S	
28	29	30	31	1	2	3	25	26	27	28	29	30	1	2
4	5	6	7	8	9	10	3	4	5	6	7	8	9	
11	12	13	14	15	16	17	10	11	12	13	14	15	16	
18	19	20	21	22	23	24	17	18	19	20	21	22	23	
25	26	27	28	29			24	25	26	27	28	29	30	
							31	1	2	3	4	5	6	



Website Address Options

Visited on [website] on _____ Learn More ▾

Blacklight Inspection Results Between 01/17/24 and 02/08/24

Blacklight works by visiting each website with a headless browser running custom software built by The Markup. To learn more, read our [methodology](#).

01/17/24 02/08/24

2	Ad trackers found on this site. ▾ This is less than half the average of seven that we found on popular sites	1
2	Third-party cookies found. ▾ This is less than half the average of three that we found on popular sites	2
#	Tracking that evades cookie blockers wasn't found.	▾
#	Session recording services not found on this website.	▾
#	We did not find this website capturing keystrokes.	▾
#	Facebook Pixel not found on this website	▾
#	Google Analytics' "remarketing audiences" feature not	▾

Comparison Button & Calendar Picker Interface

Side-by-Side Comparison

Scan of the Day

Website Address Options ▼ Scan Site

Scans of the Day
chormusical.com - 05:26:23
youtube.com - 21:01:32
espn.com - 03:54:42
amazon.com - 07:12:12
scan time, additional info
scan time, additional info



Website Address Options ▼ Scan Site

Scans of the Day
chormusical.com - 05:26:23
youtube.com - 21:01:32
espn.com - 03:54:42
amazon.com - 07:12:12
scan time, additional info
scan time, additional info

Scan of the Day Landing Page

View Scan of the Day Option

User Testing & Validation

Concepts to Validate:

- Ease of use of the current layout
- What kind of data do we need to display for our users?
- Overall aesthetics of the front end

Lessons Learned:

- Average user may have trouble navigating to "Compare Previous Scans" button
- Show which days have scans available on the date picker
- Refine and experiment with the data visualizations

Product Demo



Ethical Considerations

Concern:

Blacklight users (what websites they scan and how they interact with the site) could be mishandled or compromised

Decision:

We do not plan on saving/storing/sharing or handling user data for the Blacklight tool in any regard

Positive Impact:

Blacklight Archiver enables users to be aware of how their data is being handled online without compromising their own data

Next Steps Beyond Capstone

1

Refine the design of the front-end and add more UI/UX features

2

Continue refinement of the data pipeline to ensure optimal operation

3

Dictate the levels of data visualization available to the user based on data scrapes

Thank You!