

Designing for Data Donation

ALgorithms EXposed Workshop Report

Simply Secure

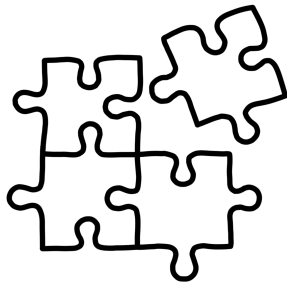
February 2020

Contents

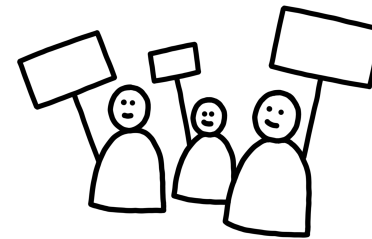
INTRODUCTION /	Problem Space	3
	Design Challenge	4
	Approach	5
OUTCOMES /	My Data	6
	Data Donation Ideation	9
	User Testing	12
	Outreach Concepts	14
SYNTHESIS /	Barriers	18
	Guidelines	20
	Credits	25

Problem Space

As users of social media and e-commerce platforms,
we are targeted by **invisible algorithms**.



ALgorithms EXposed (ALEX) is a browser extension that gathers algorithmic behaviors to reconstruct the proprietary algorithms that intrude into our lives. The extension ALEX is not a final solution to the problem; it is only part of a chain of mechanisms working to solve this growing issue. Its outputs will serve users, researchers and policymakers by mapping the extent of personalization practices, the evidence of algorithmic targeting, and the associated damages.



In order to achieve its goal, ALEX needs to collect and analyze large amounts of data. While the technology is highly complex and abstract, ALEX strives to be accessible by a large and diverse user group to spur collective activism and present viable findings.

Design Challenge

Prospective volunteers face many barriers to sign on as research participants due to **unawareness** of the problem space, **fear** of potential harm, or **usability issues** of the interface in browser-based user studies.



How do we communicate the **significance of the situation** while avoiding alarmism and sensationalism in a time of co-existing crises?



How can we prove **trustworthiness**?

- ▶ Why would users, who know the risks, volunteer to donate their data?
- ▶ Why would users, who just learned about the risks from ALEX, volunteer to donate their data?



What does an **onboarding** process look like with these constraints in mind?

Approach

Design Thinking Workshop

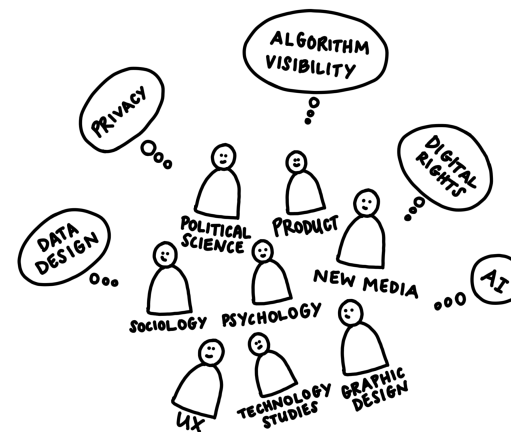
We developed and facilitated a one-day design thinking workshop, structured to gather diverse input on aspects of the ALEX tools.

GOALS

- ▶ Design a “data donation” onboarding process
- ▶ Test and collect feedback on the current ALEX system
- ▶ Design sample messaging, outreach, and educational campaigns

Participants

Our participants were experts in design, digital rights, data sovereignty, and algorithm ethics, with an interest in developing ethical design patterns. Backgrounds ranged from UX to Political Science.



Participants shared concerns of the intersection of data governance and other circumstances.



OUTCOMES /

My Data

The foundation of the workshop was an activity for participants to consider their relationship with data. Our focus was to explore potential research questions that could be supported by ALEX. Individually, each participant completed the worksheet entitled “My Data”. Then, the self-assessment led to a partner discussion. With themes in mind, the group reconvened to solidify their synthesis.

The worksheet is titled "MY DATA WORKSHEET" and contains the following sections:

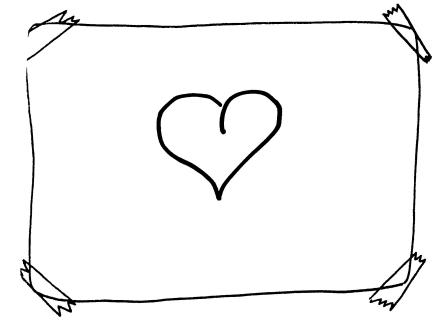
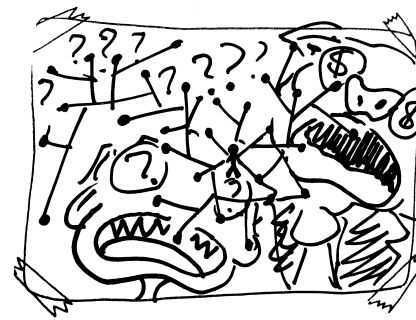
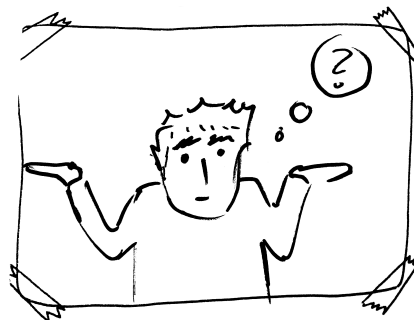
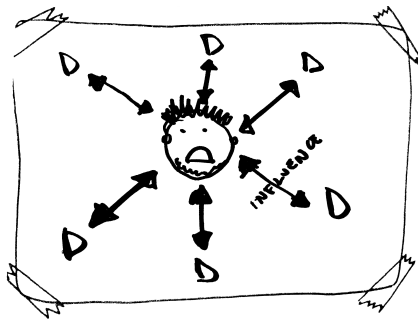
- 1. A DRAWING OF MY RELATIONSHIP WITH DATA:** A large empty box for drawing.
- 2. I AM COMFORTABLE SHARING MY DATA WITH:** Three categories with "YES", "NO", and "MAYBE" options:
 - THE GOVERNMENT? (with a drawing of a government building)
 - COMPANIES? (with a drawing of a server rack)
 - MY FAMILY? (with a drawing of three people)
- 3. HOW SENSITIVE IS THIS DATA?:** Three categories with "LOW", "MEDIUM", and "HIGH" options:
 - EXAMPLE: BIRTHDAY (with a drawing of a birthday cake)
 - EXAMPLE: FACEBOOK (with a drawing of a Facebook logo)
 - EXAMPLE: LOCATION (with a drawing of a location pin)
- 4. WHO COULD BENEFIT FROM MY DATA?:** A large empty box for writing.
- 5. WHO COULD BE HARMED BY MY DATA?:** A large empty box for writing.
- 6. WHEN IT COMES TO ONLINE TRACKING WITH REGARD TO MY PERSONAL DATA, I WOULD BE CURIOUS ABOUT:** A large empty box for writing.

The “My Data” worksheet helped participants consider their relationship with data.

OUTCOMES /

My Data ^(1/2)

① A DRAWING OF MY RELATIONSHIP WITH DATA ②



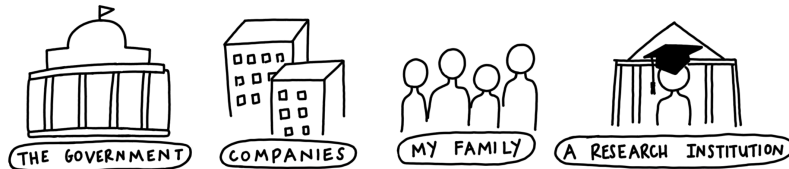
Participants' first task was to draw their relationship with data.
Four examples serve to represent the group.

Based on the open-ended worksheet prompt, participants independently made a drawing of their relationship with data. Although participants' drawings are open for interpretation, we connected themes of confusion, anxiety, stress, but also love (there were some researchers and data scientists in the room of course).

OUTCOMES /

My Data (2/2)

② I AM COMFORTABLE SHARING MY DATA WITH:



Overall, there was little consensus amongst participants about their openness to donating data. Participants asked:

☞ Are the companies ethical?

(Participants were wary of personalized marketing campaigns)

☞ Is the government providing me with services?

(Data could help improve services and maintain rights)

☞ What is my relationship with my family?

(Families are close, but perhaps don't share the most intimate details)

I am comfortable sharing my data with **research institutions**:

Yes, I am.

57%

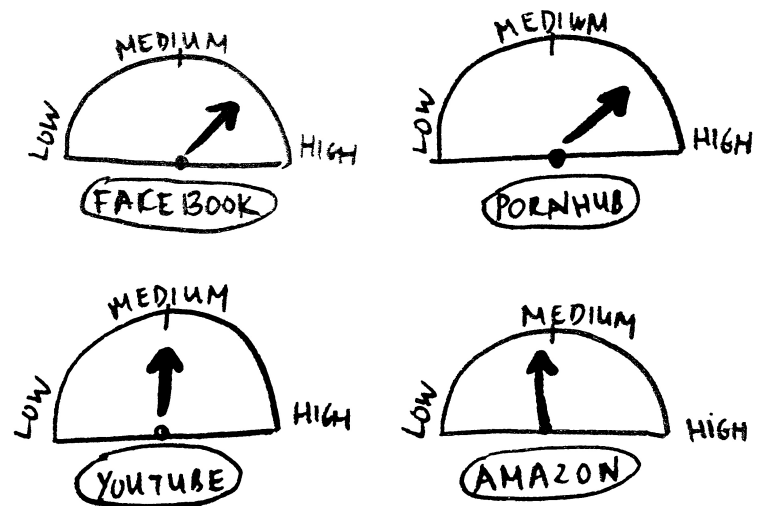
Maybe, it depends.

43%

No, I'm not.

0%

③ HOW SENSITIVE IS THIS DATA?

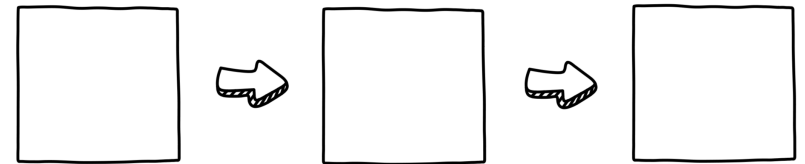


The aggregated responses for participants evaluating the sensitivity of the types of data according to platform.

ALEX aims to expose algorithms on various platforms. On a scale of low to high, participants rated how sensitive the data is generated from each. The aggregated results revealed that all data is considered sensitive. However, the most sensitive data is housed on Facebook and Pornhub.

OUTCOMES /

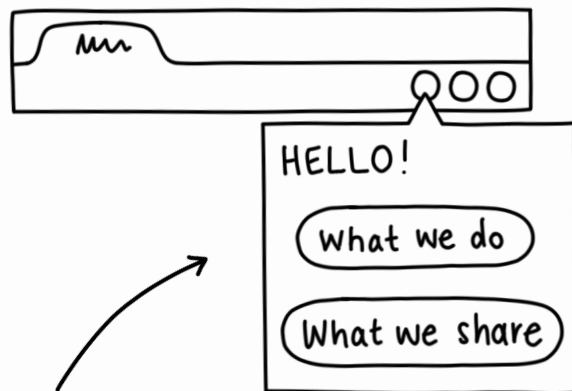
Data Donation Ideation



To establish best practices for collecting user data, participants collaborated in designing a workflow of data donation. Participants sketched mockups, and through analysis agreed on guiding themes and refined recommendations.

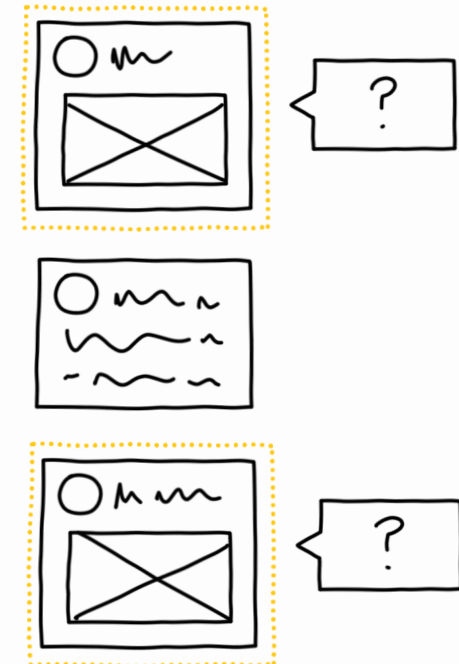
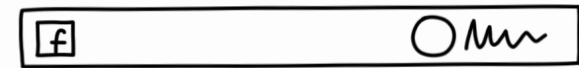
Data Donation Ideation ^(1/2)

Discussions revolved around four themes, which served to guide sketching.



Transparency

Overall, participants desired to know why the data was collected and how it will be handled. Any system that handles users' data needs to be transparent from the moment of first encounter. This requires clear communication on the website and within the extension toolbar. This is especially important during the onboarding. Clarity contributes to informed consent. Therefore, ALEX should to always give the option for more information and deeper explanations.



Interaction

To conceptualize ALEX's interaction with the interface of its target platforms, participants considered Facebook's news feed. The extension could highlight news stories on the feed and upon hover gives a detailed explanation of the pushed content, including why it was shared and the aggregated data that ALEX has collected so far.

Data Donation Ideation (2/2)

Consent

Explicit consent means clear, simple questions with an array of choices. Users should be given the opportunity to refer to or revoke their consent at any time, easily and quickly. Clear opt-ins simplify the task for the user. Only allowing for all-or-nothing consent will force some users to opt out. Straightforward, simple language helps users understand the types of data they could donate.

Customization

Providing a robust set of customization features is necessary to respect the privacy of the user. A proposal of a central registry for data with lists of who has access would put users at ease. Users may want to choose to vary settings for different situations.

OUTCOMES /

User Testing



Using the “think-aloud” method, testers uncovered issues of the current ALEX browser extension and website. Participants guided their partner tester through a series of predetermined tasks, noting critical incidents, errors, remarks, and questions. These recommendations may be generalizable to other ways of donating data.

User Testing

Testing scope:

- ▶ Onboarding: installation, comprehension of ALEX & donated data
- ▶ Usability: user expectations, system status, data presentation
- ▶ Exporting: downloading and viewing CSV data
- ▶ Conclusions: confusions, improvements, learnings

Discovery

Participants had often learned about the ALEX project through a recommendation. To install the extension, they Googled the ALEX website or searched in their browser's extension store. With differing names (e.g. ALEX, fbTREX, yTREX), participants struggled to find the ALEX extension.

Onboarding

From the beginning, users wanted to know why they should contribute to ALEX. After exploring the ALEX extension for almost an hour, some users were unsure how to describe it and wondered which data exactly was being collected.

Usability Issues

Call-to-action buttons are hidden in drop down menus or at the bottom of the pages, causing uncertainty and frustration for the user. Some elements that seem to be active buttons are inactive.

Terminology & Language

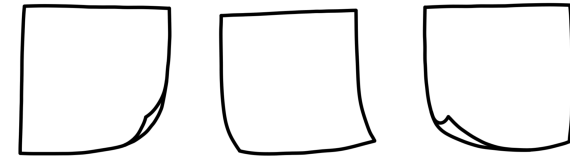
Inconsistent language confuses the user. ALEX uses terminology that can be standardized, such as: organic, without ads, sponsored. Users objected to some language was condescending or harsh. ALEX shows its bias against Facebook, yet all ALEX users are also users of Facebook. How can ALEX empower people to understand Facebook's algorithms without shaming them for using Facebook?

Data Presentation

Many users did not understand the benefits of a CSV file. Without prompting, they said they would not have downloaded their data. They preferred to see their data within the extension.

OUTCOMES /

Outreach Concepts



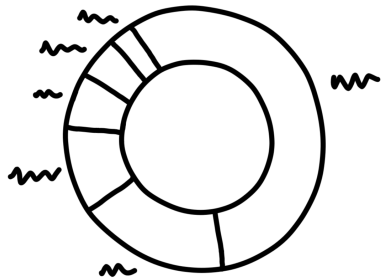
Learning Models

To kickoff the design of educational materials and position ALEX as an learning tool, we utilized a Design Thinking Sprint to craft profiles that would help us to understand how to reach various target users. In interview pairs, participants assumed the identity of a friend or relative to develop each profile. Participants identified needs, constructed problem statements, and generated methods for outreach.

Future Headlines

“What would be an ideal future campaign for ALEX?” Participants wrote their ideas on post-its while walking to encourage rapid ideation, stimulate creative thinking, and set a physical destination for their creative goal. The group used an affinity diagram to uncover emerging themes.

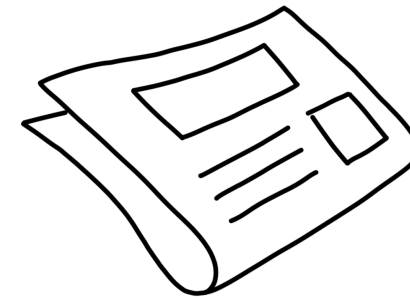
Learning Models (1/2)



Feature visualizations of data

Profile inspiration: Robin, 27-year-old PhD student in statistics

Robin wants to dig into the data. He wants to know what ALEX offers to its users. The ALEX website can provide the tools that allow learners to create visualizations of the findings. Statistical analysis will help users like Robin to understand the impact of the algorithms that they can't see.



Vary media and learning approach

Profile inspiration: Regina, 83-year-old Facebook user

Regina's mental models sometimes don't align with the platforms she uses. Therefore, participants concluded that ALEX could be more accessible by adapting the format of media and learning styles. An onboarding video might help one learner, while a written explanation helps another. Regina might have more clarity if data were presented in a traditional, familiar media, like a newspaper.

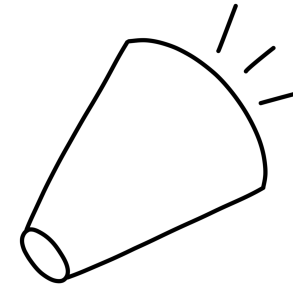
Learning Models (2/2)



Leverage credible channels

Profile inspiration: Elena, 31-year-old biologist

Elena wants to see changes to society that are based on scientific research methodology. She trusts credible sources, especially the scientific journals to which she subscribes. She is skeptical of unreviewed sources, like a blog. A trustworthy source would lend credibility to ALEX's findings and Elena would save time reviewing their validity.

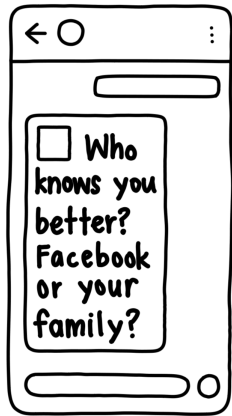


Consider tone and voice

Profile inspiration: Armando, 73-year-old retiree

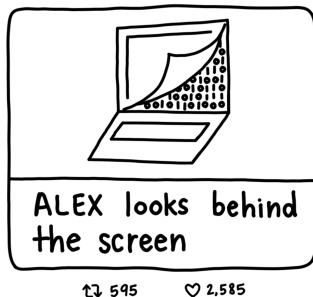
Armando has a keen interest to learn more about day-to-day technologies because he wants to connect with his grandson. Yet, he is weary of breaking news and disillusioned by politics. His goal is to be informed and have topical knowledge so that he can converse and connect with others. Therefore, the presentation of the issues can use a variety of tones to reach the most amount of audiences.

Future Headlines



Control

ALEX could give insights to the ads and recommendations to help the user understand what Facebook or YouTube know about them. Users would like to see what ALEX has learned about the algorithm and the reasoning behind specific posts. There could be more interaction in the ALEX badges on posts.



Transparency

A feature of ALEX would allow the user to see what's behind their news feed and understand its workings. Participants imagined if they could physically peel open their news feeds to see what's behind it.



Polarization

ALEX can answer the question: Does polarization exist and to what extent? Sophisticated algorithms can lead to exploitation, allowing for discrimination and polarization to gain traction. Working with this in mind, ALEX can start to consider how to build better algorithms.

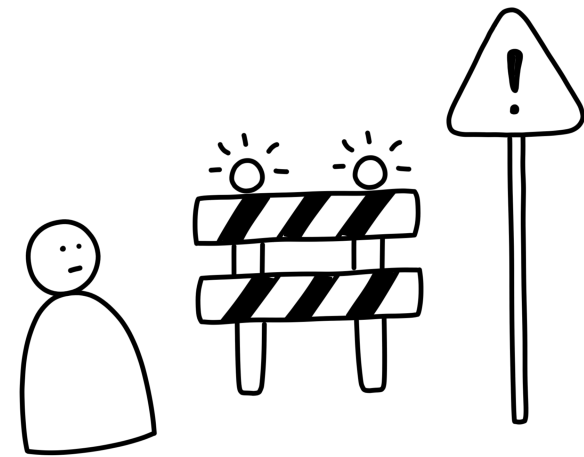


Extremism

ALEX can work quantify extremist posts and bring that data into light. Understanding the policies of these platforms, how to identify the post, and when it's been removed can help users be more vigilant.

SYNTHESIS /

Barriers to Data Donation



Barriers

Sharing data is intimidating. Participants are sensitive to sharing data in the era of data breaches, manipulation and exploitation. Participants discussed the types of harm that could arise from breaches on those platforms. Noted were concerns related to career, finances, reputation, and more.

High-risk scenarios

For some users, the risks of data exposure are too high for themselves and those they care about. Donors have the privilege to contribute when others do not. Because sharing is linked to personal risks, data donation tools must not shame any user who is unable or unwilling to donate.

Context

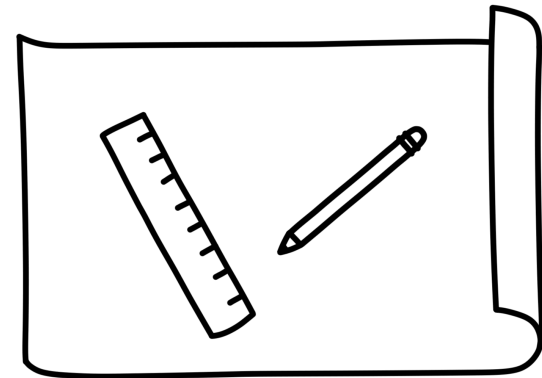
The propensity to share personal data is linked to subjective perception of safety; those who belong to marginalized groups are especially sensitive to their data. Data sensitivity is culturally contextual. When considering trust in government, participants considered not just their governments, but also repressive regimes that might use collected data for persecution. Participants were particularly vocal about data permanence: once data is collected by a trusted institution, that data will still remain if the institution becomes repressive in the future.

Community

In regard to sharing, participants voiced concern for their family and community: “My data impacts those around me.” Sharing could put friends and family in danger, depending on the type of data. Data on social networks can be easily linked to one’s social network. The risk is extended past the individual.

SYNTHESIS /

Guidelines of Data Donation



Transparency

Complete transparency is key to giving users the information they need to make an educated choice. Participants' comfort with sharing data depends not just on the goal, but also on the approach, such as the clarity and honesty of the communication around data sharing.

Inform the user of...

- ▶ The research goal
- ▶ Why the research necessitates data donation
- ▶ The risks of exposing personal data
- ▶ The measures taken to reduce risks
- ▶ Who is performing, supporting, and leading the research
- ▶ Who has access to users' data
- ▶ What information will be collected, with examples
- ▶ How long the data will be stored
- ▶ What happens to the raw data once research is finalized

And use various methods:

- ▶ Informative onboarding is an opportunity to demonstrate the value of investigating the problem space
- ▶ A tutorial could help users acquire the knowledge of algorithms and the ALEX tool
- ▶ Empty states serve to show ALEX's potential, with sample data or detailed explanations of featured topics
- ▶ Varying the communication style, explanations, and resources helps diverse sets of users learn
- ▶ An updated milestone timeline of achievements or goal-tracking involves the user and provides clarity.

Consent

Informed consent is complicated. Participants need to understand what their data is and where it's going. How can ALEX collect informed consent without becoming a burden to the user? By offering a variety of controls, users can choose the consent model that works for them.

Remind and Repeat

ALEX can't assume opt-in to their collection tool. Data is collected over time - it's not a one-time donation. Therefore, users might want to set reminders. ALEX can employ pop-up reminders or other types of status indication to ensure users are aware of their consent to donate data. A trade-off is that pop-ups can be distracting and contribute to constant calls for attention that burden users.

Allow for customization

Users should be able to choose the data they share. Participants were more willing to share their data given customizable control. This includes sensible defaults which help the user choose the safest options. For example, these ideas were visualized a slider with ranges or a checklist of specific examples of the type of data.

Set an expiration date

ALEX's options must allow users to change their mind. Consent can be limited to a specified data range or set to expire automatically at any time in the future. At that time, ALEX can ask users if they still want to participate.

Consistency

Standardization is vital in establishing user confidence in the system. Most users are unable to evaluate the trustworthiness of a system's code, so they rely on other means, often the visual design. The trust that comes from the visual identity translates to trust in the technology and organization.

Unify visual identity and branding

Design is a communication tool that helps the user understand the priorities of the organization. Iconography, graphics, color scheme, voice, and patterns all need to be carefully considered and documented for regular reference. The patterns should align with ALEX's values and convey ALEX's commitment to security and privacy. Furthermore, users will be able to search for the tool and recommend it to friends if they recognize the brand.

Employ clear and approachable language

Prompts need to be friendly and sympathetic, without shaming the user. Neutral and supportive language will comfort the user instead of shaming them for their internet behavior. Additionally, the tool requires frequent explanations of the meanings of terminology, processes, and labels. Terminology should be standardized so that the user is able to make connections.

Adhere to established standards

Not only must the ALEX design standards align internally, ALEX must adhere to existing browser extension interface standards. These conventions offer familiarity with a new system, providing comfort to the user.

Deliver

Users are looking for a tool that gives back. When users volunteer as donors, ALEX can offer them something in return, whether it be tools, resources, updates, results, and recognition. User interaction with the tool can be an excellent motivator and serve as a means of education.

Analyze and interpret for users

Users expected a tool that would provide personal insight into their online behavior. With analysis and interpretation, ALEX increases the value in donating data. Aggregation of data will help users understand algorithms, manipulation, and targeting. This dynamic resource could be the incentive that keeps users engaged and excited to donate.

Spur engagement with campaigns

Users may have a desire to have more involvement in the project, therefore ALEX can take inspiration from crowdfunding platforms. ALEX can give like-minded individuals a way to form a group and spread their message. This idea contributes to organizational and goal transparency. When people are part of the campaign they are more engaged and more in control.

Credits



ALgorithms EXposed (ALEX)

ALEX is a project based at the University of Amsterdam and funded by the European Research Council. The project aims to investigate automated personalization and filtering for research and activism. Read more at tracking.exposed.

ALEX Team

Claudio Agosti	Giulia Corona
Stefania Milan	Davide Beraldo
Umberto Boschi	Salvatore Romano

Special thanks to

[CityLAB Berlin](#) for providing the event space.



Simply Secure

Simply Secure exists to advance trustworthy technology that protects vulnerable populations. Their work spans capacity building, design support, open research facilitation, and community convening. Read more at simplysecure.org.

Simply Secure Project Team

Georgia Bullen
Eileen Wagner
Ngọc Triệu
Kelsey Smith

For more information about this report,

please email us at contact@simplysecure.org.