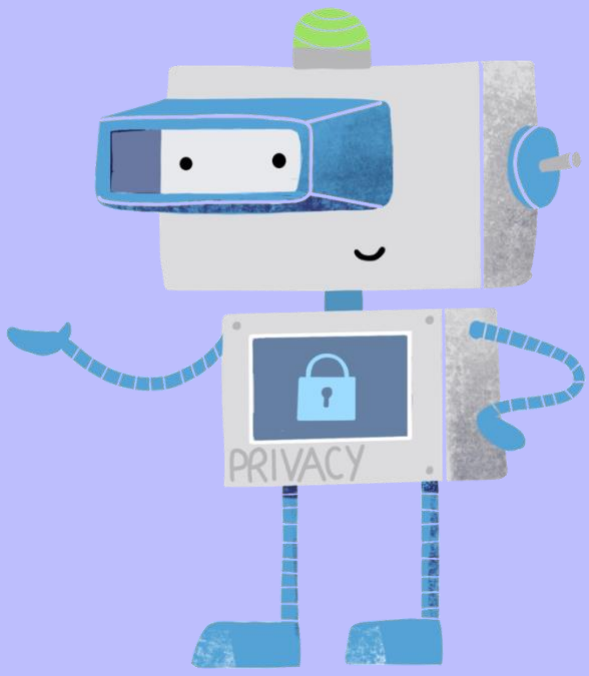


PrivacyBot

A simple way to start exercising your privacy rights.

James Carney | Archana Kulkarni | Joanne Jia | Cameron Lopez
Berkeley School of Information
MIMS 2021 Capstone Project

Acknowledgements

We couldn't have completed this project without the help of an amazing team of mentors and contributors. Thanks especially to Chris Hoofnagle for shepherding us through this project, and to Lesley Matheson, who aided in telemetry and additional project directionality. Thank you to Esther Jan, who designed our amazing logo. And a big thank you to Tobias Holl, Daniel Aranki, and Ashkan Soltani for technical guidance.

Executive Summary

From Spring 2020 to May 2021, our team of four MIMS students developed a product that streamlines the California Consumer Privacy Act's data-deletion request process and allows users to more easily protect their data. The tool, dubbed "PrivacyBot," is a production-quality API and accompanying user interface that acts as a router to allow users to remove their personal data from a mass number of data-hoarding companies simultaneously. We completed four main iterations of research and development, summarized below.

First iteration

- Opening rounds of research and development, and completion of a tool that was able to effectively send out formatted emails with user information from a single email address of our choosing.
- From our research, we found that people largely cared about their own privacy but took steps to protect themselves only when convenient. Additionally, submitting data deletion requests online is extremely tedious and daunting even for the experienced. And finally, while all participants said they cared about privacy, only those with sufficient knowledge about privacy took steps to actively prevent sharing passive information.

Second iteration

- Iteratively designing the first prototypes of our user interface, conducting a competitive analysis of our "competitors" which led to a change in project direction, and the creation and incorporation of what we believe to be the most exhaustive list of data brokers and people search sites in existence.
- We found that building enough trust to prompt users to enter their information would serve as a challenge.

Third iteration

- A round of thorough unit testing to get some semblance of response expectancy, another round of user interviews to uncover user trust levels, a shift to open-source and implementation of 0auth verification, and the creation and incorporation of a functional user interface.
- We found that shifting the project to open source and adding google email verification would increase trust levels amongst our user base.

Final iteration

- One final round of usability testing coupled with a diary study to understand how users respond to the installation instructions, the application UI, and the volume of email

correspondence. After making final tweaks, PrivacyBot released on our landing page at privacybot.io, and can be run locally using a few simple command line prompts.

- Findings included that users were overwhelmed with the volume of emails, prompting a change in our filtration and initial email scheme. Additionally, users still had questions as they walked through the flow, so we added informational drop-downs as a final touch.

Next steps

- Met with the office of the California Attorney General as well as with a reporter from Consumer Reports. Our hope is for PrivacyBot to spark a data democratization revolution and put an end (or at least increase the regulation of) the data brokerage industry.

Intro

In recent years, with the introduction of the California Consumer Protection Act (CCPA) in the United States and its associated data deletion clauses, it has become slightly more accessible for the privacy-concerned consumer to exercise their privacy rights and protect their personal data. However, to remove personal data from sites deemed malicious or non-essential, individuals must manually visit the websites of each corporation that contains their data and submit a formal data deletion request. As you can imagine, this process is tedious and requires knowledge, time, and resources, rendering it out of reach for many consumers. For this reason, our team of four MIMS students have developed a product that streamlines the data-deletion request process and allows users to more easily protect their data.

The tool, dubbed "PrivacyBot," is a production-quality API and accompanying user interface that acts as a router to allow users to remove their personal data from a mass number of data-hoarding companies simultaneously. Its production was heavily interdisciplinary, incorporating facets of both qualitative and quantitative user research to determine the target audience and product space while concurrently employing software engineering and system design to generate the back-end functionality of the product. Additionally, iterative design and design research were employed to create several usable interfaces to carry out multiple purposes. Together, nearly every major career track that can be furthered through an education at the Berkeley School of Information was represented in the production of this project. The following pages offer an in-depth analysis of that product journey, from inception to deployment.

Background and Definitions

Before we get into the product journey and descriptions, let's define some terms that will be critical in your understanding of our product space and of what PrivacyBot aims to accomplish.

CCPA

The California Consumer Privacy Act of 2018 (CCPA) gives consumers more control over the personal information that businesses collect about them, and CCPA regulations provide guidance on how to implement the law [1]. This landmark law secured new privacy rights for California consumers, including:

- The right to know about the personal information a business collects about them and how it is used and shared;
- The right to delete personal information collected from them (with some exceptions);
- The right to opt-out of the sale of their personal information; and
- The right to non-discrimination for exercising their CCPA rights.

The scope of this project is restricted to the subset of rights under CCPA privacy rights as they pertain to “Delete” and “Opt-Out” of personal information and particularly with respect to usage by data brokers.

Right to Delete

Businesses must designate at least two methods for consumers to submit requests—for example, a toll-free number, email address, website form, or hard copy form [2]. From what we have seen, almost all businesses offer at least some form of online methods such as their email address or a website form. Businesses’ privacy policy page must include instructions on how consumers can submit their requests. Businesses cannot make consumers create accounts just to submit a deletion request unless consumers already have accounts with them.

Right to Opt-Out Of Sale

Businesses are required to provide a clear and conspicuous “Do Not Sell My Personal Information” link on their website that allows consumers to submit an opt-out request [3]. Again, creating an account can not be necessary for this process.

Data Brokers

Under California law, a data broker is “a business that knowingly collects and sells to third parties the personal information of a consumer with whom the business does not have a direct relationship” [4]. This law exempts certain businesses that are regulated by other laws from this definition. Exempted businesses include consumer reporting agencies (commonly known as credit bureaus) and certain financial institutions and insurance companies. Data brokers collect information about consumers from many sources including websites, other businesses, and public records. The data broker then analyzes and packages the data for sale to other businesses.

People Search Service

People search services are companies such as White Pages that scrape public government records and sell your data to other people or companies who want information on you specifically.

Defensive OSINT

Open-source intelligence (OSINT) is a multi-factor methodology for collecting, analyzing and making decisions about publicly available data to be used in an intelligence context [5]. *Defensive* OSINT can therefore be thought of as using the “inverse” of these methodologies so that those who employ OSINT methods cannot find any of your information, such as removing your information from people search sites.

Ideation

Now that we've provided a bit more context behind our project's key terms, let's decompress the initial project ideation. Our idea for PrivacyBot stemmed partially from prior privacy knowledge of a few of our group members, and partially from a class taught within the School of Information — Information Law and Policy. This class contained a segment on the CCPA which introduced our group to the landmark legislation and sparked our idea for a way to mass-send data deletion requests, as completing the process individually seemed incredibly tedious. Our next instinct was to delve into several businesses that make a living from selling data and see how they managed the data deletion process, as well as look into forums and message boards within the privacy community.

From our investigations, we found that privacy enthusiasts as well as people who employ "defensive OSINT" have long utilized methods which seek to remove their data from people search sites. Usually these are in the form of worksheets which list the top people search sites and provide instructions on how to remove a person's data manually [6]. Examining the privacy policies of some of the listed companies, we observed that many of them used web forms as their preferred means of CCPA deletion request submission [7-10, to name a few]. So, in the Spring of 2020, we began exploring methods of automating this process so as to avoid the tedious nature of completing forms by hand. Thus, the first technical iteration of PrivacyBot aimed to produce scripts to automatically complete the data delete request forms, with the defensive OSINT community as a vague starting point for our target users.

First Iteration

Unfortunately, we quickly ran into issues with auto-completing forms. Our first proof of concept used the web interactivity module Selenium which allowed us to programmatically fill out and submit these CCPA forms, all done from reusable python functions [11]. However, there were a few limitations that prevented this from being used at scale. Firstly, Selenium is extremely brittle and in the case of even minor HTML changes between websites, it falls apart. The second major issue was with overcoming reCAPTCHA, which there really wasn't a feasible solution to. Understanding these problems, we quickly pivoted to another acceptable method for deletion requests under the CCPA: email [12]. We found that nearly all businesses have a privacy email that handled such requests, and we knew that these could be sent at scale. This became the basis for our project, and led to the first round of user research and our first official software milestone.

Our first round of user research was carried out prior to production of our newfound software objective, and consisted of unearthing a more concise target audience. So, we conducted a series of semi-structured interviews in order to better understand who would be using our tool [Appendix, Oct-Dec 2020]. For your consideration, this research, along with the extensive amount of additional qualitative research we conducted over the course of the project's development, was conducted through an interpretivist epistemological lens. This means we sought to uncover the *meaning* behind our user's opinions to give us deeper context and aid us in PrivacyBot's development. Therefore,

during these initial interviews we explored the answers to research questions like “What do people know about privacy?”, and “What needs do people have regarding privacy?”, in addition to “Who would be our target audience?”, and “How do they imagine it to work?”. Through our inaugural interviews, we identified the following patterns:

1) People largely cared about their own privacy but took steps to protect themselves only when convenient.
2) Submitting data deletion requests online is extremely tedious and daunting even for the experienced.
3) While all participants said they cared about privacy, only those with sufficient privacy knowledge took steps to actively prevent sharing passive information.

The second insight aligned fairly well with other studies conducted by Consumer Reports, and all three in combination prompted us to scope PrivacyBot production toward the *privacy conscious community* — those concerned about their online privacy who have interest in going out of their way to protect themselves, but that aren’t necessarily intensely technical [13]. Thus we had our first target audience direction: Our product’s intention became to help those who already understand the value of privacy to execute data deletion requests more efficiently.

While this critical research was taking place, our engineers were also working toward our new software objective (iteration 1.5 if you will) of creating a functional email routing service. To get technical, this was done using the Python SMTP module. The SMTP establishes a TLS connection to Gmail in order to read in email credentials from environment variables, log in to the provided email account, and ultimately send an email. This version had no user interface yet — all interactions were conducted through POST requests to the Gmail API via a user’s terminal or through services like Postman, so it required a lot of technical knowledge. At this point we hadn’t yet finished our first rounds of user research and were still envisioning a deployed web service, and so were developing our tool with that as the end goal. This would shift as more research offered new insights down the line.

To summarize: at the conclusion of our first iteration of research and development, the tool was scoped toward knowledgeable members of the privacy community and was able to effectively send out formatted emails from an email address of our choosing.

Second Iteration

After accomplishing our first software milestone of developing a functional email router and working to determine a direction for our target audience, we shifted some of our efforts into iteratively developing a user interface for iteration 2.0 [Appendix, Prototype Iterations, Iterations 1-3]. From our ongoing user research we knew that our target audience was not solely confined to the ultra-technical privacy community but also to slightly less technical privacy-concerned individuals, thus a functional interface would allow PrivacyBot to expand to this wider audience rather than

restricting the tool to the terminal command line. We defined two main goals for constructing the interface, the first being to make it as simple as possible for a user to enter in their information and select which companies to remove their data from, and secondly to identify any other features that may be useful through user research.

Through a series of usability tests coupled with the ongoing research into our target audience, we discovered several critical insights which helped in creating our first interface prototype, the most prominent of which related to user information input. We encountered significant difficulties with how to prompt users to enter their information without raising suspicion and causing them to abandon the tool. In our initial prototype we incorporated several pop-ups coupled with informational text that attempted to answer questions about the tool before the user had a chance to raise them, but this was a temporary fix and did not persist to our final iteration [Appendix, Prototype Iterations, Iterations 1-3]. These were just the beginnings of our struggle to build enough trust with users to allow them to enter their information into our tool, which we grappled with until the bitter end (aka our product drop).

While conducting these usability tests and working to design our interface prototypes, we were similarly hard at work on one of the most critical aspects of the project, i.e., our data broker and people searcher list. Firstly, to understand what the standard benchmark for something like PrivacyBot would be, we conducted a competitive analysis of other commercial tools that provide data deletion services. Our research informed us that our competitors charge exorbitant prices for the service and support a fairly limited number of data brokers and people search sites [14, 15]. Given this knowledge, we compiled what we believe to be the most exhaustive list of data brokers and people search sites that currently exists in order to create the best data deletion tool on the market. The completed list is made up of data brokers from several states' Attorney General websites and other miscellaneous online sources, and contains 5+ times the number of supported data brokers than the other leading tools [14, 15]. It also contains information pertaining to what exact information each data broker needs in order for a data deletion request to be processed, stored as individual booleans in one large csv file. In all, the list took more than 40 hours to compile and we see it as PrivacyBot's "secret sauce" that sets us apart from other similar paid services. Coupled with the fact that our tool is completely free, we believe it will be an immensely important contribution to the privacy community.

After we had compiled the list of data brokers and people searchers, the software aspect of iteration 2.0 consisted of building out the email router to be able to support sending mass emails to our extensive csv list. It was at this point that we learned from various testing methods that sending emails on users' behalf was problematic because it required PrivacyBot to be an "authorized agent," something we weren't qualified for nor something we wanted to act as [16]. Additionally, sending emails on users' behalf caused some of our emails to get flagged as spam, due to a single email account sending batches of identical emails. It seemed as if a deployed web service presented too much privacy and legal risk, so we made the decision in iteration 2.0 to pivot our project to fully locally run. This meant we would send the emails from the user's email account as opposed to a centralized PrivacyBot email client. This change provided several major benefits:

1) All user PII would now be handled within users' machines and with their email clients, so we don't risk processing any user PII.
2) It removed the issue of continually supporting a web service.
3) Later research will reveal that it increased user trust in PrivacyBot
4) We no longer had to use our own email server to send CCPA requests on users' behalf, which solved our authorized agent and spam issues.

In summary, iteration 2.0 consisted of iteratively designing the first prototypes of our user interface, conducting a competitive analysis of our "competitors" which led to some great insights and a change of course in our project direction, and the creation and incorporation of what we believe to be the most exhaustive list of data brokers and people search sites (and their deletion request requirements) in existence. This iteration gave PrivacyBot a new direction and certainly solved a lot of issues, but as with anything, it raised more questions than it answered.

Third Iteration

Now that we had some semblance of functional software and with our newfound technical changes (i.e. conversion to a locally run tool), our third iteration began with product testing. The goal of the testing was to collect data on how businesses reacted to our email requests and to work out some kinks in our tool. We distributed the testing of 100+ data brokers each amongst our team members and emailed our CCPA requests using PrivacyBot. We documented information such as: response time; whether an email bounced back; if the request was denied or completed; if the company asked for additional information; if they directed us to an online form. We then compiled this information and used it to create an infographic to be displayed on our website, to give users some idea of what they should expect with regard to company follow-up after using PrivacyBot.

While we tested this iteration of PrivacyBot's functionality, we conducted another round of interpretivist-style user interviews with participants who were recruited using a survey we deployed on several different subreddits [Appendix, Jan-Mar 2021]. The goal of these interviews was to understand how our target audience felt about our new technical changes, understand what information they were comfortable with sharing, and importantly to uncover whether users trusted our product. With regard to trust, we found that shifting the project to open source would increase trust levels amongst our user base. And as for the changes that came with our shift to a local tool rather than a hosted web service, we discovered that having to enter in an email address and password (which would now be required) would turn users away unless done through official Google services. Not only that, but it also posed a large security risk to store email credentials in plaintext within environment variables, as certain malware can gather environment variable content from tools like ours.

From these interview findings we made the decision to shift our project to be open source upon its release, and also to add 0auth verification through the Gmail API as a method of handling

email credentials [17]. This method was not only more secure, but it increased the likelihood that a user would be comfortable giving email permissions to PrivacyBot, which we drastically needed. There is a minor downside associated in that using the Gmail API limits our user base to Gmail users with vanilla Gmail accounts (no SSO etc.), but Gmail has an incredibly wide user base and it's fairly easy to set up a temporary account, so this small sacrifice bore little weight. And as for the open source aspect, we agreed that once PrivacyBot has been released to the public we will license it using an MIT open source license.

The final major step of our project's third iteration was implementing the user interface functionality in the form of a reactJS application [Appendix, Prototype Iterations, Iteration 4]. We determined this to be the most efficient way to develop our front end, as reactJS applications can run locally in a user's web browser and can easily pass data to the Flask API that initiates the actual requests.

As we worked through the back end implementation to connect the react app to the Flask API, we used various iterative usability tests along with the incorporation of several new technical limitations to reformulate our user interface. Because we had shifted from a web service to a locally run version, we now had separate user flows for our landing page and for our application in iteration 3.0. From the usability tests we found that this shift did not change much in terms of usability — the app itself still resonated best with users when it consisted of three simple pages: an opening hero page, a data entry page, and a summary page. However, importantly, users suggested branching away from the serious security tone from previous iterations and developing a more user friendly theme. That's right, iteration 3.0 is when the PrivacyBot mascot was born! We shifted our designs to be slightly more playful and easy-going, and centered the experience around our helpful friend the PrivacyBot.

To summarize, iteration 3.0 involved a round of thorough unit testing to get some semblance of response expectancy, another round of user interviews to uncover user trust levels, a shift to open-source and implementation of 0auth verification, and the creation and incorporation of a functional user interface. At this stage, PrivacyBot was beginning to come together and we were quite happy with how things were turning out.

Final Product

As we approached the completion of a production-ready product, we conducted one final round of usability testing coupled with a diary study to understand how users respond to the installation instructions, the application UI, and the volume of email correspondence over the period of one week [Appendix, Apr-May 2021]. From these final methods with the near-completed product, we found several critical insights that we couldn't have achieved without the fully functional API, each imperative to improving the final user flow of our product application.

The main issue reported was that the initial number of emails the user received upon using PrivacyBot was, to say the least, overwhelming. Our plan was to have the user cc'd on each sent email

to simplify company response tracking. However, we were unanimously informed that the email volume upon submission was too great. As a solution, we implemented several interface and technical changes, the first of which being to eliminate the feature that cc'd the user on each email. This drastically reduced the number of emails the user received upon submission of their delete requests and allowed company responses to arrive in a more staggered, organic manner. The second alteration was to create filters for the number of requests sent out to accommodate users that didn't want to deal with follow up responsibilities. The filters include an option with only the largest data brokers and people searchers, an option with only people searchers, and a final option to send to every business on our list [Appendix, Prototype Iterations, Iteration 4]. The goal of these changes are to allow more users to enjoy PrivacyBot, regardless of the amount of effort they wish to put in after using it.

Another issue we encountered is that users had an abundance of questions yet again about the information they were prompted to input into the application. Despite our previous efforts to mitigate this issue with a link to the FAQ and extensive explanations within the reactJS app, it appeared as if this didn't suffice as users did not want to be taken out of the application. Our solution was to add informational drop-down menus adjacent to each field input, informing users of why each piece of information may need to be utilized by a data broker and giving salient examples in attempts to placate any user queries [Appendix, Prototype Iterations, Iteration 4]. Our end goal was to ensure that any questions users may have are given as many opportunities to be answered as we could muster while still maintaining the sleek, easy feel of our tool.

After implementing the above changes, the end result of our extensive research and hard work is a production-quality API that acts as a router to allow users to remove their personal data from a mass number of data-hoarding companies simultaneously. As of now, PrivacyBot can be downloaded from our landing page at privacybot.io, and run locally using a few simple command line prompts. The application then opens and users are prompted to enter in their information, select a subset of companies to request their data be deleted from, and then submit their requests after a quick built-in Gmail authentication step. They are then sent a confirmation email and shown a summary page of the process that just took place, and any follow-ups from the companies they sent deletion requests to are directed to their own inboxes. It's a very simple process and can be completed in around 10 minutes depending on how long it takes the user's device to download some of the required python packages.

Impact and Next Steps

Our project has certainly garnered initial attention and has the potential to have a great deal of impact not only within the privacy community, but for the general public and especially for data brokers. For example, many of the responses to our deletion requests have been human-generated, likely from companies' small privacy teams. We can only imagine what will happen to these teams' workloads if our tool gains traction and they start receiving hundreds or even thousands of requests. We believe that if done correctly our tool could realistically spell the end for more than a few data

brokers. And we aren't the only ones who think so — a reporter from Consumer Reports reached out to us expressing interest in writing about our tool and was eager to try it herself. Others from Consumer Reports' Digital Lab also met with us and proposed partnering with PrivacyBot in the long term and building out a multi-year strategy to scale our tool.

Our team also met with several members from the Office of the California Attorney General. Specifically, they were excited about the prospect of having a large amount of telemetry data pertaining to data broker compliance. They recognized that PrivacyBot has the power to emphasize major flaws in the CCPA request pipeline and in the Act's enforcement, and could potentially lead to legislative changes in the future depending on our results at scale. The AG office additionally suggested partnering and having some way for users of our tool to report the compliance rates of given companies and were quite interested in the continuity and institutional support of the project long term.

As for next steps, we are currently working toward eliminating the need to do any sort of command line entry in order to use the product, so as to simplify the process and increase the appeal to those of less technical capability. We will release PrivacyBot as a completely free, open source tool and continue to work with any interested party in the short term with hopes of passing the project off to someone just as passionate about privacy rights as we are, whether that be Consumer Reports or someone else. Our hope is for PrivacyBot to spark a data democratization revolution and put an end to (or at least increase the regulation of) the data brokerage industry.

Works Cited

[1] TITLE 1.81.5. California Consumer Privacy Act of 2018 [1798.100 - 1798.199.100]

[2] Cal. Civ. Code § 1798.135(e)(1).

[3] Cal. Civ. Code § 1798.135(b)(1).

[4] Cal. Civ. Code § 1798.99.80

[5] "Open-Source Intelligence." Wikipedia, Wikimedia Foundation, 24 Apr. 2021, https://en.wikipedia.org/wiki/Open-source_intelligence.

[6] Bazzell, Michael. "Personal Data Removal Workbook & Credit Freeze Guide". *Extreme Privacy: What It Takes to Disappear – 3rd Edition*. 2021.

[7] "Privacy Policy". *Truthfinder*. 2021. https://www.truthfinder.com/opt-out/v2/

[8] "Privacy Policy". *Refinitiv*. 2021. https://privacyportalde-cdn.onetrust.com/dsarwebform/5f7a2da0-bed0-45e8-ac2c-c1f297e2efdc/cc7a82a6-f352-4dd4-afd9-00087cb9c97e.html

[9] "Privacy Policy". *Tracers*. 2021. https://www.tracers.com/opt-out/

[10] "Privacy Policy". *US Search*. 2021. https://www.ussearch.com/opt-out/submit/

[11] "Selenium Automates Browsers. That's It!" *SeleniumHQ Browser Automation*, www.selenium.dev/.

[12] Cal. Civ. Code § 1798.140(i).

[13] Mahoney, Maureen. "California Consumer Privacy Act: Are Consumers' Digital Rights Protected?" *Consumer Reports Digital Lab*. 1 Oct. 2020.

[14] *PrivacyDuck*, www.privacyduck.com/?gclid=Cj0KCQjwp86EBhD7ARIsAFkgakj8lu-6k4-PAa9lZotisDL-r2BuqvqbO82uqEQtFc__lbOH-tbDX74aAin-EALw_wcB.

[15] *Confidently.com*, https://confidently.com/index.html

[16] Cal. Civ. Code § 1798.135(c); §1798.140(y).

[17] Grimes, Roger A, and Josh Fruhlinger. "What Is OAuth? How the Open Authorization Framework Works." CSO Online, CSO, 20 Sept. 2019, www.csoonline.com/article/3216404/what-is-oauth-how-the-open-authorization-framework-works.html.

[18] "MIT License." Wikipedia, Wikimedia Foundation, 27 Apr. 2021. https://en.wikipedia.org/wiki/MIT_License

Appendix

Research Timeline and Methodology

October - December 2020

Method 1: User Interviews

Goal: Understanding privacy needs across experience levels, discovering a target audience

Participants:

Pseudonym	Occupation	Age	U.S. Region	Data Privacy Knowledge Level
Ann	Coding Bootcamp Student	26	Mid West	Low
Ben	Software Engineer Manager	29	West Coast	Medium
Cathy	Design Researcher	28	South	Low
Diana	Law School Student	22	West Coast	Expert
Esther	Software Engineer	26	West Coast	Medium
Freddie	Cybersecurity Leader	35	West Coast	Low
Kevin	Mechanical Engineer	22	West Coast	Low
Parth	Electrical Engineer	23	West Coast	Expert
Melanie	I-School Student	26	West Coast	Medium
Sloane	I-School Student	NA	West Coast	High
Ryder	I-School Student	28	West Coast	Medium

We underwent two rounds of user interviews with the above subjects, the first round consisting of 4 participants, the second round consisting of the remaining 6. The goal of these interviews was to uncover privacy needs across different knowledge levels in order to narrow our target audience.

Main findings:

1) People largely cared about their own privacy but took steps to protect themselves only when convenient. Similarly, they would willingly use not-as-safe tools if it was more convenient.

2) Submitting data deletion requests online is extremely tedious and daunting even for the experienced. Most users we talked to stopped at a handful of deletion requests.

3) While all participants said they cared about privacy, only those with sufficient knowledge about privacy took steps to actively prevent sharing passive information.

Method 2: Usability Tests

Goals: Make it as simple as possible for a user to enter in their information and select which companies to remove their data from, and identify any other features that may be useful through user research.

We ran a total of 9 usability tests (3 per iteration for the first 3 iterations) over the course of our first iterative design cycle, the results of which can be seen farther down in the appendix. Each usability test took between 20-30 minutes and involved the user completing the user flow in its entirety as well as giving feedback on the landing page layout and information.

Main Findings:

1) Some means of company selection was important

2) Users had lots of questions about why so much of their data was needed

3) The landing page should be used as an informational tool, it should answer a majority of their questions while not being too wordy

January - March 2021

Method 1: Online Survey

Goal: Quantifying needs and gauging interest among our target market, the privacy conscious community)

We designed and posted an online survey link to four separate privacy-related Subreddits: r/cybersecurity, r/cybersecurityadvice, r/bigdata, r/OSINT to gather initial feedback on our product idea and to assess the needs of our intended audience. *We also used this survey as a means of recruiting for our subsequent user interviews.*

Main Findings:

1) We found from our survey that there was a good amount of interest in PrivacyBot — we received 12 survey responses, 10 upvotes, and a few direct messages asking for a link to our tool.

2) We learned that many privacy conscious individuals have submitted data deletion requests in the past, but not through any free or paid services. All were at least "somewhat likely" to use a service to automatically submit data deletion requests. 86% of surveyed privacy enthusiasts find this service useful even if it only handles the initiation.

3) There was some unease with letting a service handle their data. 86% reported to trust this service to some degree with their data (name, email, state, etc). Only 1 person reported to be "somewhat unlikely" going to trust the service.

4) However, they placed more trust in reputable sources such as an academic institute and open-source projects. It was also important that tools transfer data in encrypted state or follow all PII compliance.

Method 2: User Interviews

Goals: Understand how our target audience feels about our newly found technical limitations, understand what information they're comfortable with sharing, and uncover whether users trust our product

Participants:

Pseudonym	Occupation	Age	U.S. Region	Data Privacy Knowledge Level
John	Coding Bootcamp Student	26	Mid West	Low
Jacob	IT Technician	NA	West Coast	Medium
Thomas	Software Engineer	NA	South	Expert
Nathan	Cybersecurity Researcher	28	West Coast	Expert

This round of interviews consisted of 5 total interviews, each taking ~45 minutes. We wanted to re-define our target audience because of some newly-realized technical limitations, and we also wanted to uncover whether or not users would trust our product. And if they didn't, we wanted to uncover what we could do to increase trust levels.

Main Findings:

1) Making the project open source would increase trust levels

2) Having to enter in email address and password would turn users away unless through official Google means

3) Including more reasoning for why certain data needed to be included increases the likelihood that users will enter it, e.g. "Adding this will get rid of x more companies".

April - May 2021

Method: Combination Usability Test and Diary Study

Goal: Testing the usability of our tool from end to end and perceived level of trust

After undergoing unit testing of our tool ourselves, we recruited two students from the I-School community (one novice and one more experienced) for a full usability test and diary study. The usability test consisted of a 30 minute user flow where the user walked through the set-up and deployment of PrivacyBot. These tests were followed by a week long diary study in which the user reacted to the volume of emails they received from data brokers. We concluded with a 30 min debrief at the conclusion of one week to understand broad opinions and suggestions of where we could improve.

Main Findings:

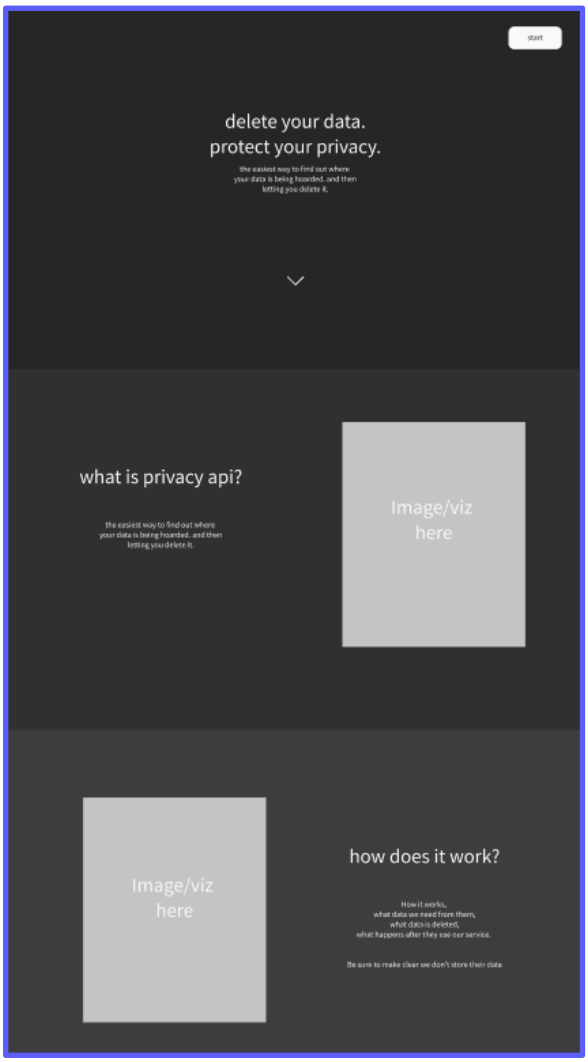
1) The initial number of emails the user received upon using PrivacyBot was overwhelming

2) Users had an abundance of questions yet again about the information they were prompted to input into the application

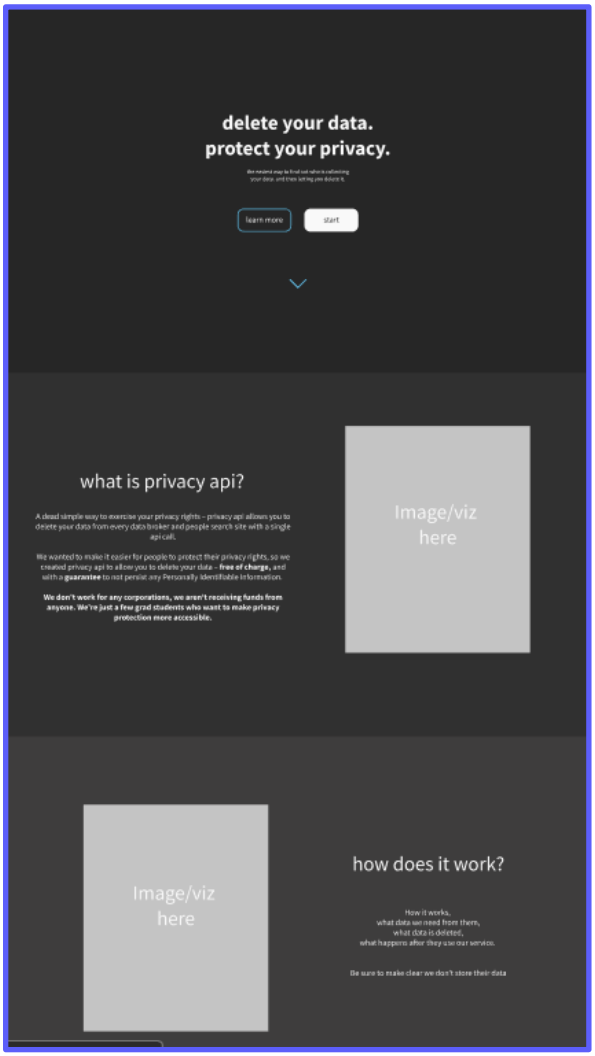
Prototype Iterations

Landing Page (Entire page not shown for space/interpretability reasons)

Iteration 1:

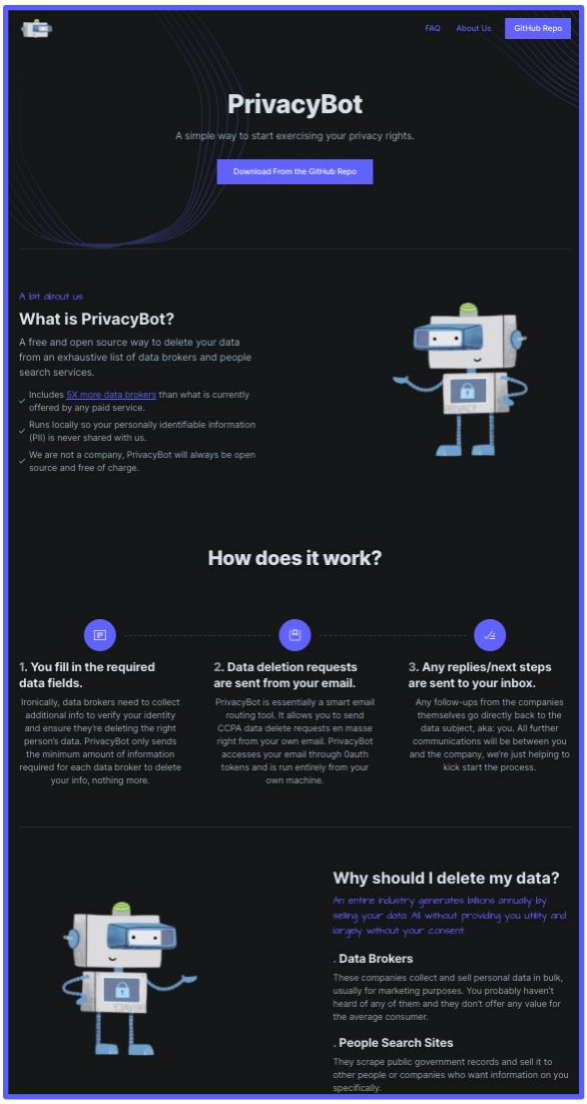


Iteration 2:



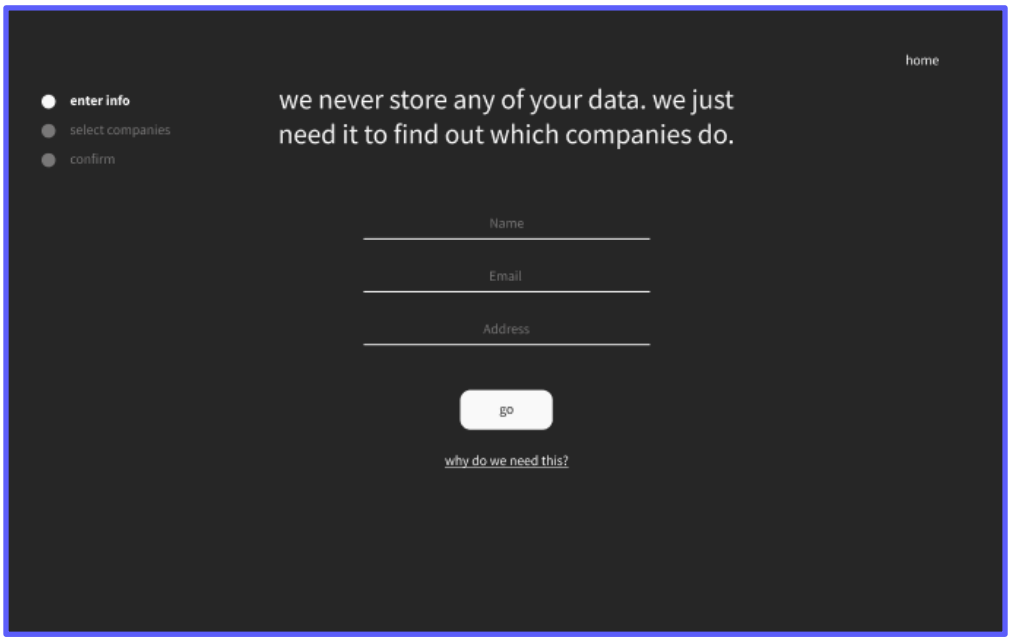
Iteration 3:

Iteration 4:

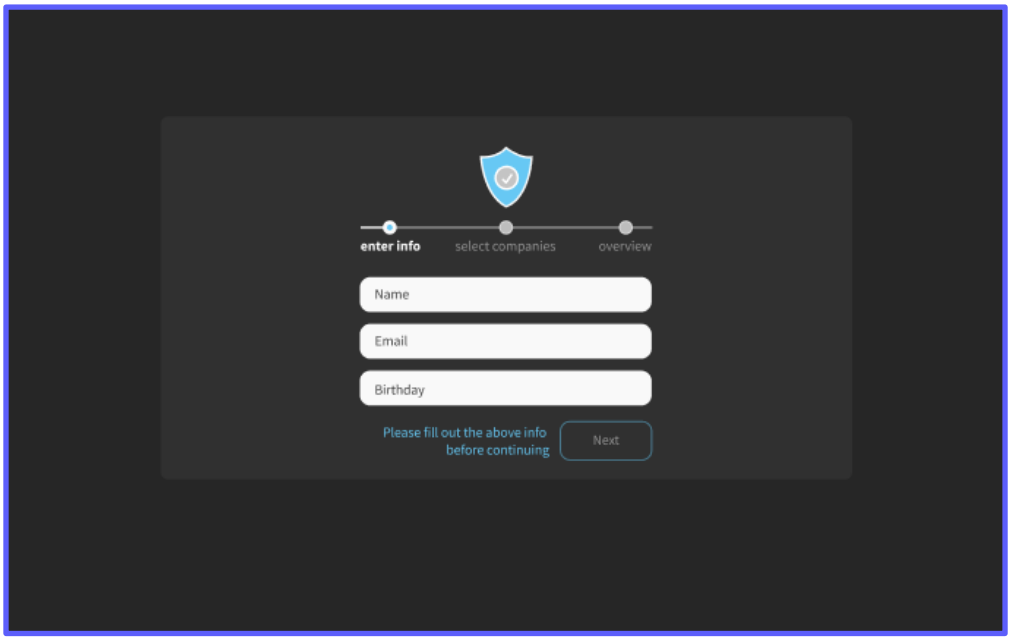


Data Entry Interface

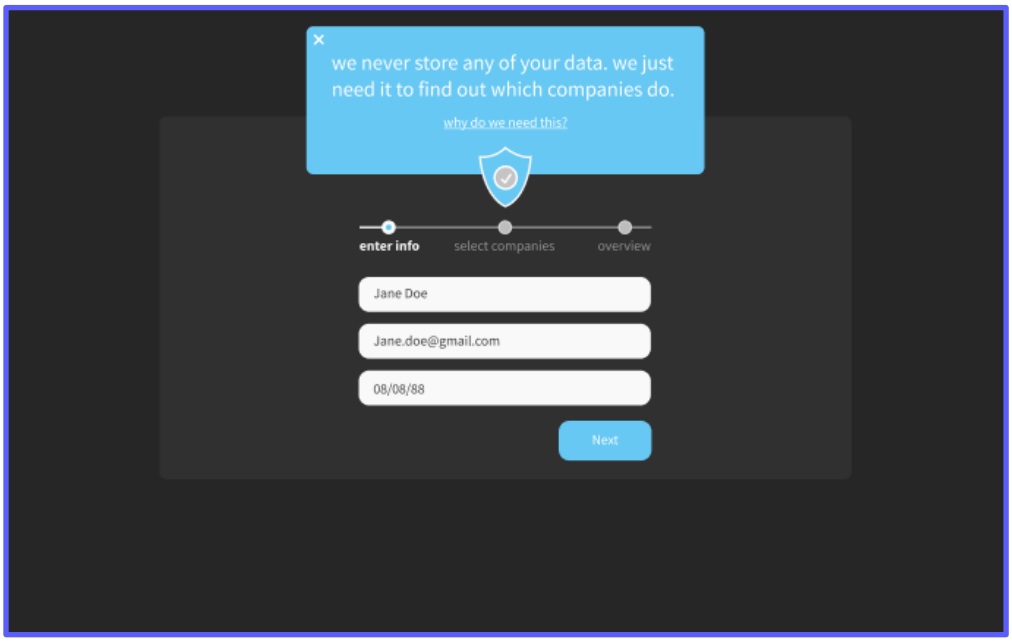
Iteration 1:



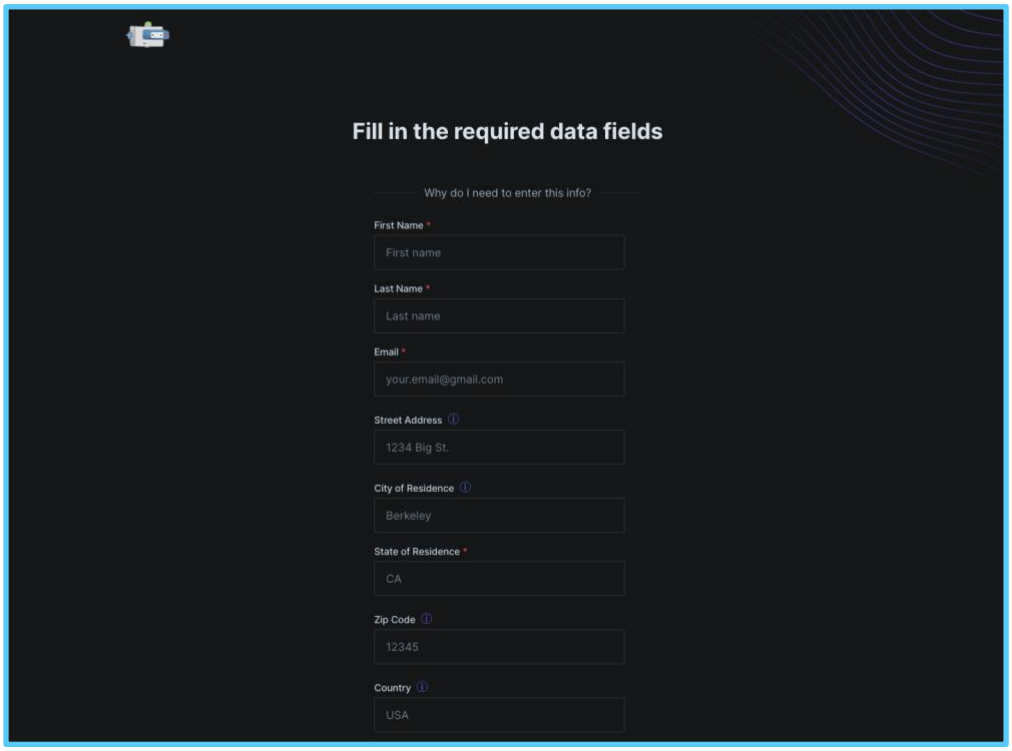
Iteration 2:



Iteration 3:

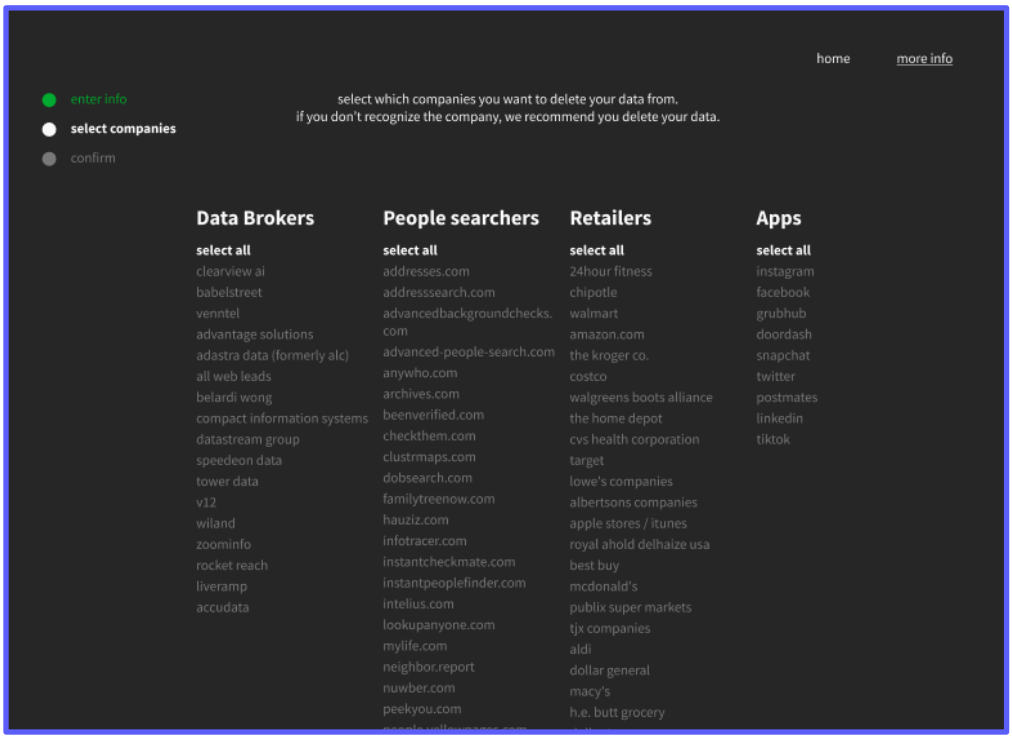


Iteration 4:

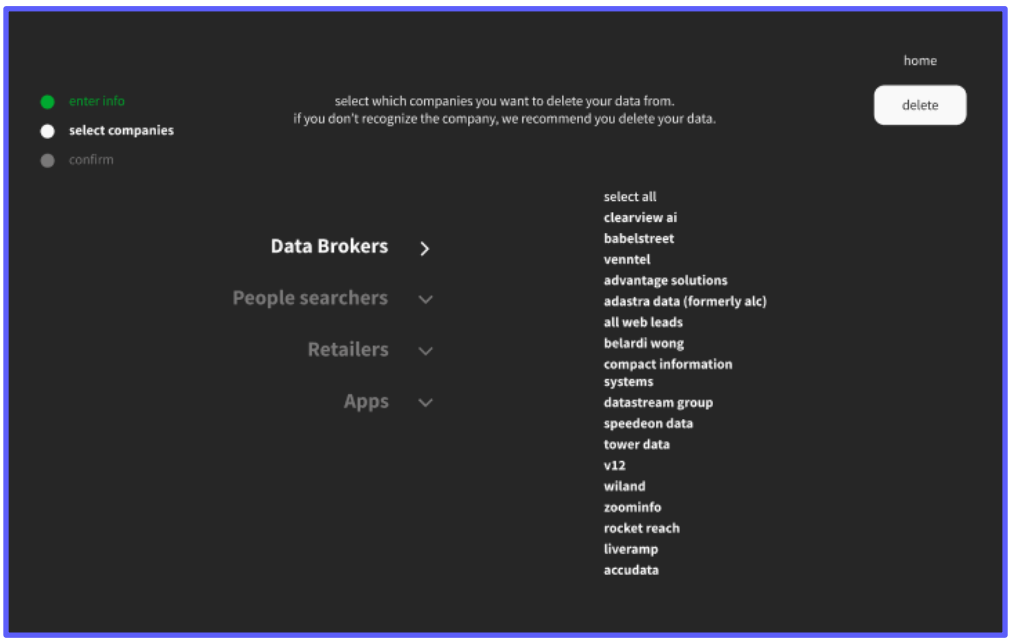


Company Selection Interface

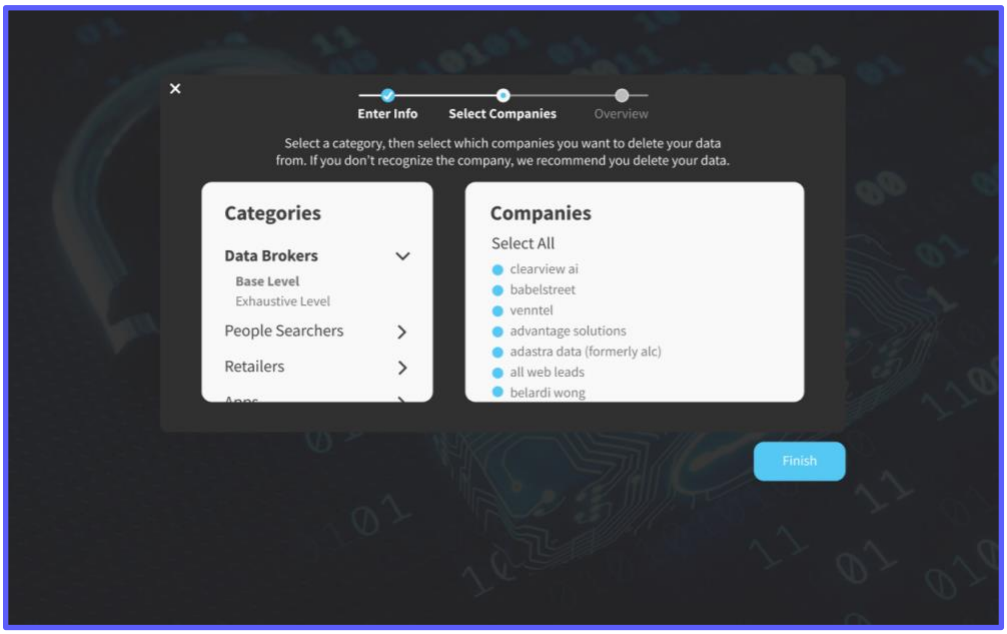
Iteration 1:



Iteration 2:



Iteration 3:



Iteration 4: