Black Book:
Drawing and Sketching
BBDS Scientific Journal

Vol.2 , Issue 1, Inside - Ouside, 2021


Article

# Globule: Control data in order to manipulate the algorithm.

Bruno, M. R., Silva; Helena, M., dos Anjos; Lola, M. F., Meintjes; Pedro, D. P., Tavares

Faculdade de Belas-Artes da Universidade de Lisboa

Abstract:

When people are connected to the internet, their personal data is being shared with companies that use it for their own benefit, almost as a factory farm for human beings. Hence, by tracking and profiling everyone, those companies make profits by manipulating our behavior. In fact, they are not trying to sell their products; each one of the internet users becomes “the product.”

In addition, there is the “filter bubble”, a term that refers to the results of the algorithm that constrains what one finds online, creating a unique universe of information for each of the users, altering the search results.

Globule is a free software that allows the users to scan all the information available about themselves across the internet, listing the data shared and creating new profiles of data to manipulate the algorithm and increase the level of control of internet search results. It offers a complete customization of the alternative profiles, creating a random profile.

Thus, this app interferes with the algorithm used to collect information, according with the user’s profile: controls the search results, preventing the filters and reducing restriction, debugging and redirecting them, to provide access to others; to boycott, deceive and confuse the system, or those using the data collected, in order to avoid any kind of harassment; or to protect the privacy users by managing the data that becomes available online.

Keywords: data, surveillance, control, manipulate, algorithm, interaction design.

## 1 INTRODUCTION

Every time we use a device connected to the internet, we expose our data to the algorithm, remaining available online in an uncontrolled way (fingerprints, facial identification, location, routes, behaviors, habits, preferences, or even data about our health).

This data is later shared with companies that use it for their own benefit, allowing them to track and profile customers to “evolve the services they provide”, manipulating others. But that makes the customers (we) “the product” being sold.

In addition, there is the “filter bubble”, which conditions and changes the results that the algorithm returns to us when we search online, creating a unique universe of information for each user.

The danger of accepting constant data surveillance is the self-censorship of behaviors and the loss of individual freedom.

## 2 RESEARCH

### 2.1 Problem statement

Common search engines like Bing and Google limit our search results, depending on our interests and private data that has previously been collected about us. This issue can lead to an inequality of opportunity, since the results for a certain topic are not universal, due to the fact that the same research could lead to different results.

### 2.2 Research

Our research began with looking for and analyzing some projects with features and mechanisms to identify and block data surveillance.

'Gobo' is an alternative social network with a focus on user privacy. This social network allows a certain control of the algorithms, or a set of "rules", so the user can decide what gets shown on his feed and know why.

"Social media companies use algorithms to control what we see on our feeds, but we don't know how these algorithms work. As a result, we're often unaware why certain posts show up in our feed while others don't (…) Social media companies make assumptions about what we want to see based on what we read and click on. They tend to show us content we're already engaging with, reinforcing our echo chambers. Instead of assuming what you want to see, Gobo allows you to add unfamiliar perspectives into your feed, so you can better understand the range of opinions that are shared online." [1]

Another project that we used as a reference was 'Smell of Data', a device that lets off a smell when a data leak happens on personal devices.

"The sense of smell helped early humans to survive. But now that our hunting and gathering has moved to the digital environment, our noses can no longer warn us of the lurking dangers in the online wilderness. (…) The Smell of Data is a new scent created to instinctively alert internet users of data leaks on personal devices." [2]

The 'Spy-proof smartphone case' is a modified iPhone 6 case, also a device designed to warn users if their smartphone is transmitting information that could potentially be monitored.

Snowden presented the design for the small surveillance device, dubbed the "introspection engine", conceived together with American hardware hacker Andrew Huang, stating that it "could prove useful for journalists reporting in hostile regions." [3]

### 2.3 User research

After an inquiry, we understood that our main target were people who were already aware of data surveillance and were looking for a way to protect themselves against it. Individuals who use the Internet regularly, between 18 and 44 years old (the equivalent to 69% of internet users in 2019). [4]

### 2.4 Benchmarking

Our research continued with the analysis of articles of projects such as 'Maximizing the Diversity of Exposure in a Social Network', published in IEEE Transactions on Knowledge and Data Engineering, which describes an algorithm that makes it possible to a vary the results of an online search. [5]

In order to provide citizens with new ways to stay informed and participate in public debates, in an information sharing and opinion forming environment, social-media platforms have to expose citizens to a sufficiently diverse viewpoints that challenge their assumptions, rather than being trapped inside filter bubbles.

"The resulting problem can be cast as maximizing a monotone and submodular function subject to a matroid constraint on the allocation of articles to users. It is a challenging generalization of the influence-maximization problem. Yet, we are able to devise scalable approximation algorithms by introducing a novel extension to the notion of random reverse-reachable sets. We experimentally demonstrate the efficiency and scalability of our algorithm on several real-world datasets."

On the other hand, 'Meta Search', published in Royal College of Art by Lisa Mandemaker, is a research project that explores the filter bubble in an interactive way. [6]

The filter bubble - a term coined by Eli Pariser - refers to the universe created by the algorithm for each user, selecting what information the user sees based on the information collected about the user. In other words, the

algorithm shows us what it thinks we want to see, and not necessarily what we want to see, interfering in the way we found ideas and information.

“Meta Search is a series of tests of different spatial contexts for performing a personal search on the Internet. The aim is to find new ways of searching and to test the limits of the filter bubble. I do this by creating a distance between the search engines we use on a daily basis, and reshape the way we can input our search queries. The designed context will influence the type of search, and by changing the input we could perceive the output in a different way.

In this way, Meta Search functions as a mental and/or social filter between the user(s) and the search engine, and as an attempt to break away from the filter bubble and its negative connotations.”

### 2.5 Goals

At this point our goals were to a) raise awareness about data surveillance, b) create a free way to let people protect themselves based on alternative data, c) create a program that would confuse and/or scramble the system, d) have access to the difference in results we get when combining our data with fictional data.

### 2.6 Formulation of design hypotheses

In order to achieve our objectives, we thought about a bot that filters all the data needed to enter a site, an application or a browser extension that tells us what data is being collected on a site and lets us change it. The application would allow the users to create fake data to hide their identity.. It could have been some kind of a “game” to see what results we get combining our real data with fictional data. A platform of fake parallel profile creation associated with the real one with fictional interests and searches that interfere with the AI system tapered search results.

Thus, our concept was a platform, linked to an already pre-existing account created in a browser (chrome), that enabled the opportunity to create new profiles, adjacent to the one that already exists. These profiles would be completely free, and controlled by the user. Characteristics such as age, gender, interests and employment can be imagined and conceptualized by the user, enabling him to explore new search results

So, we conceptualized a platform, linked to an existing account already created in a browser (chrome), that would provide the opportunity to create a new profile, adjacent to the one that already exists. This profile would be completely free, and controlled by the user. Characteristics such as age, gender, interests and academic and professional paths, could be imagined and conceptualized by the user, allowing the exploration of new search results.

## 3 UX Design

Based on these realities, our research proceeded with inquiries, through which we concluded that issues of privacy, security and search results affect most internet users.

Most, despite a more or less carefree use, believe that their devices capture information without consent and that their previous searches condition future searches.

### 3.1 Inquiries results

#### 3.1.1 User profile

From the total of 88 responses in the sample the main conclusions we can draw are that around 76% of the inquired are between the age of 18 and 25 years old and 24% over 30 years old. 68% are female and 32% male.

Relative to education level, 90% of the inquired are university students or had an higher level. About 70% are students and 25% employees.

#### 3.1.2 Internet usage

In terms of internet usage, 83% of the sample spends more than 5 hours a day online and is reasonably or very satisfied with search results.

#### 3.1.3 Surveillance and security

The frequency of rejecting cookies is balanced between “whenever possible” and rarely or never. About 49% of the sample never delete their cookies and only 11% do so daily or weekly.

87% consider that they rarely or occasionally provide their personal data and the other 13% are split between always or never.

Despite this more or less carefree use of the Internet, 97% believe that their devices capture information about them without consent. And more than 90% believe that their previous searches condition their future searches.

3.1.4 The APP

Just over 35% of the sample say they would use an application that added fictitious information to their personal information in order to change online search results. And suggesting functionalities such as:

- Selecting and/or intensifying search preferences (likes, interests, activities).
- Block sharing or changing personal data (location, age, nationality, religion, politics, sexuality).
- Control or change profile as a consumer.
- Know what the data is used for.
- Periodic random change of all possible data provided by the device you are searching on.

Note: After studying the answers of our inquiry, we understood that it should have had some adjustments and is perhaps incomplete.

3.2 Personas and user journeys

According to these conclusions and needs, we developed personas and their respective user journeys that are reflected in our strategy that has the purpose of:

1. CONTROL search results by debugging and targeting them, so that they are less filtered and influenced by the algorithm, in order to overcome previous activity records and access a different spectrum of each user;

2. BOICOT, deceiving and confusing the system, or whoever uses the collected data, in order to avoid any kind of harassment by companies;

3. PROTECT privacy, managing the data that is available online and making it difficult to access it.

Joana Almeida is a 20 years old student that lives with her parents in Coimbra, Portugal. She likes listening to music, shopping, going for walks, being with friends, and traveling. Her technological profile shows that she spends a lot of time online, mainly on social networks. Even though she has a high internet proficiency she has an average level of tech-savvy. Her frustrations are always having to create an account or provide email for a new service and then having her email box full of undesired messages. And his need is to search without constraints and interference.

Scenario: Joana, who has an active online life, does not intend to spend much time using the application and is curious to use it and wonders if it will work successfully. When logging in to create an account on the app she wonders if “this” will work and thinks “I hope this is safe and reliable”. With apprehension, she registers and logs into the app. After waiting for the profile to be analyzed, she checks the personal data that the app has collected and thinks “How do they know so much about me?” Surprised, she checks the collected information, corrects what is incorrect and her expectation increases. She proceeds to profile editing and select and hide the information she wants not to be shared. As she thinks “Will the ads for that brand stop appearing all the time?” she feels anxious and her anticipation diminishes. To test the app’s functionality, she opens a recurring-use website to see if the same ads appear as they did before using the app. “Wow, this really worked! She enthusiastically checks the impact of the app and the data on a search. His expectations rise again.

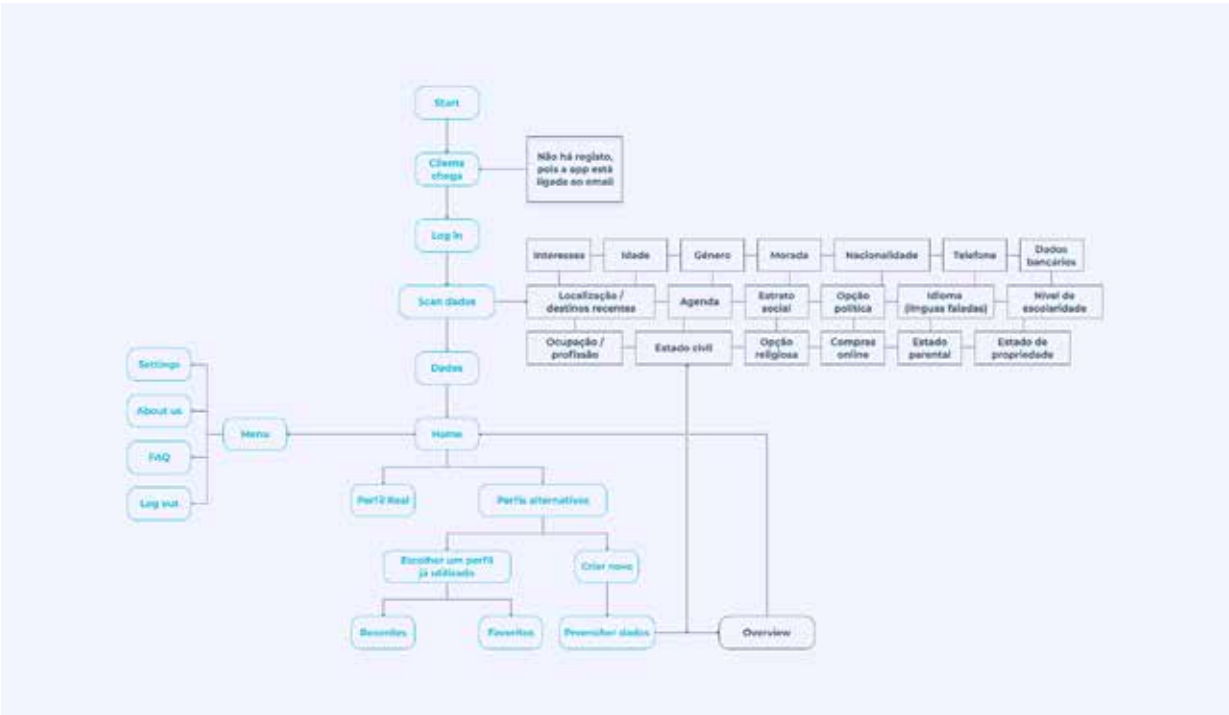
Nuno Cabral is a 27 years old software developer and lives with his boyfriend between Lisbon and Berlin. He enjoys reading, being with family, gaming and programming conventions. His profession and high technological knowledge mean he spends a lot of time online, less on social networks and always with high internet domination. His frustrations are related to data leaks and undeclared research. And he has a need for protected and undisclosed research and to protect his personal data.

Scenario: Nuno has read a review prior to the release of the application. After finishing installing the software, he is quite excited to test the program. As he logs in to create an account, he makes his appreciation of the design. “Finally, I will be able to test this app” he thinks excitedly. He checks the profile analysis and what personal data the app has collected, concluding with strange satisfaction that “Everything seems to be correct.” His expectation is still there. Next, he tests the functionality of the app in regards

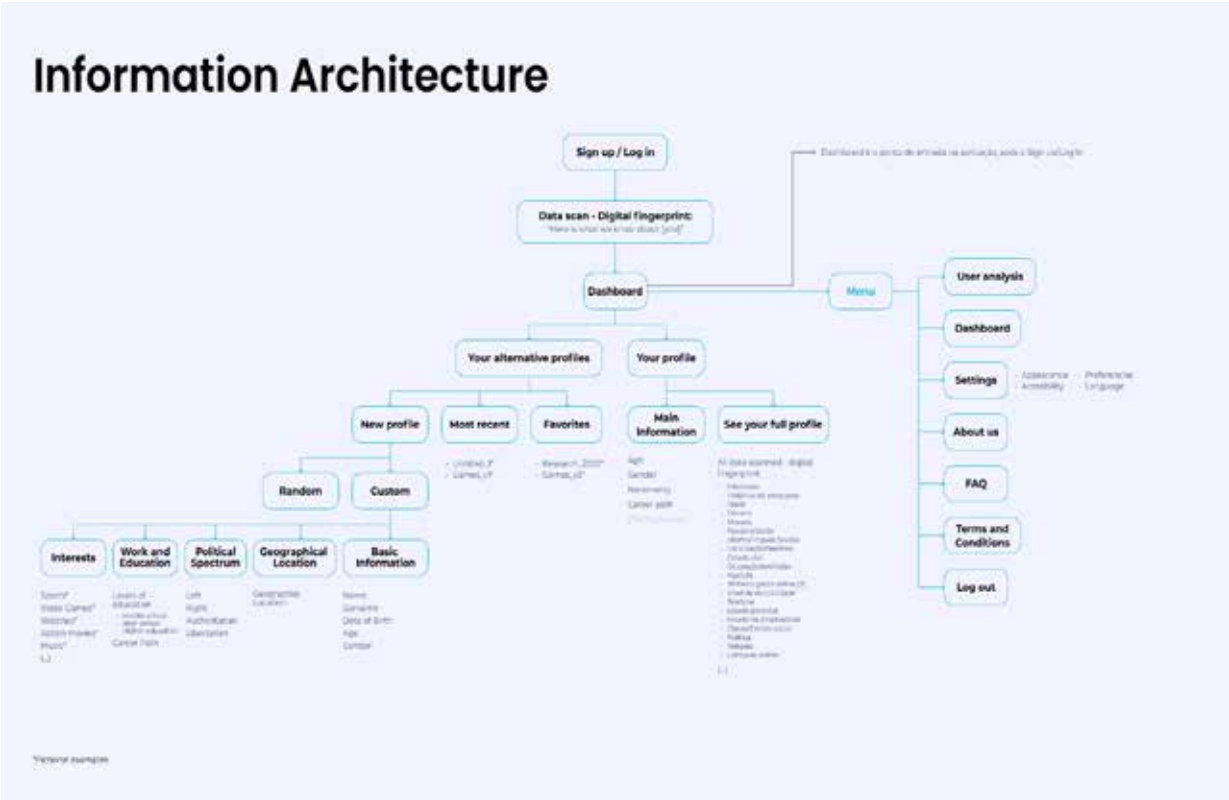
to navigation and analyzes its features. He concludes with approval that "It has a simple navigation, with all the promised features" and his expectation increases. He moves on to creating a new profile and choosing features for a new avatar. Excited, he thinks about the possibility of "making someone completely opposite!", and develops a customized (fake) profile to explore the internet. His anticipation rises again.

Duarte Dubeux is a 41 years old journalist and lives with his wife and daughter in Oporto, Portugal. He likes to read, watch movies, play tennis, go deep-sea fishing and watch soccer. He doesn't spend much time on social networks and despite some internet domain and some time spent online he is not very technologically savvy. His frustrations are not getting the information he wants and the circulation and creation of fake news. He feels the need to filter his search results and have access to diversified sources of information.

Scenario: Through a recommendation from a co-worker, Duarte begins using the application with hopes of achieving other results. As he logs in to create an account on the app he thinks suspiciously "I hope this isn't just a waste of time". He checks the profile analysis and what personal data the app has collected, thinking about the possibility of these influencing his reporting. His expectation increases as he feels interest in knowing the data used. He proceeds to create a new profile and choose characteristics for a new avatar. "Do I create a profile with a different location than mine?". With curiosity, he creates a personalized avatar to explore the Internet. His expectation rises again. While choosing features for a new avatar decides "I will change it according to the report!". He is amazed by the possibility of getting results from different demographic groups. His anticipation continues to rise.




Flow Chart




Information architecture


Wireframes / Quick Prototyping

### 3.3 User testing

Since the application is addressed to anyone, we tried to test users of various genders and age groups (between 20 and 60).

«(Click login)

Activity one: Locates and opens the menu.

Activity two: Locate and choose the new profile option.

Activity three: Locate and choose the option to set the location.

Activity four: Locate and choose the about us page.

Activity five: According to the profile presented, how old is Peter McAdam?»

When analyzing the results of our user testing, made to 8 persons, we conclude that in general users did not seem to have any major adversities when testing the wireframes, however, we did observe some patterns of difficulty.

Thus, in the "User Analysis", in which users did not know exactly where to find it, possibly because it is repeated on the dashboard and on their own page as well. On this same page we also noticed that the explore "new realities" button was not interpreted as a way to advance to the next page, partially because of its appearance, but also because of the language used. There was also difficulty when using the menu to go back to some previous screen.

We concluded that it is necessary to reformulate both expressions ("User Analysis" and "explore new realities"), and define only one path to "User Analysis", as well as add "back" arrows on each page, to avoid using the menu as the only way to navigate through the application.

## 4 PRODUCTION DESIGN

### 4.1 Available features

GLOBULE is free software that allows you to do a scan that captures, cross-references and reveals all the information deposited on online platforms where the user, logged into this application, has had activity and that is linked to it.

Each user can create one or more alternative profiles, in order to recover security and increase the level of control of internet search results, as well as the management, balance or intensification of the relevance of the information conveyed to the real profile, directing the navigation and respective results, according to the needs it reveals through the information entered in each alternative profile.

According to the user's needs, in 'Basic Information' he can manage, balance or intensify the relevance of the information that is verified to be conveyed to the real profile of the user logged into this application. He can also select, point or specify the 'Geographical Location', 'Political Spectrum' or information related to 'Work and Education' he wants to associate with the alternative profile that is to be developed.

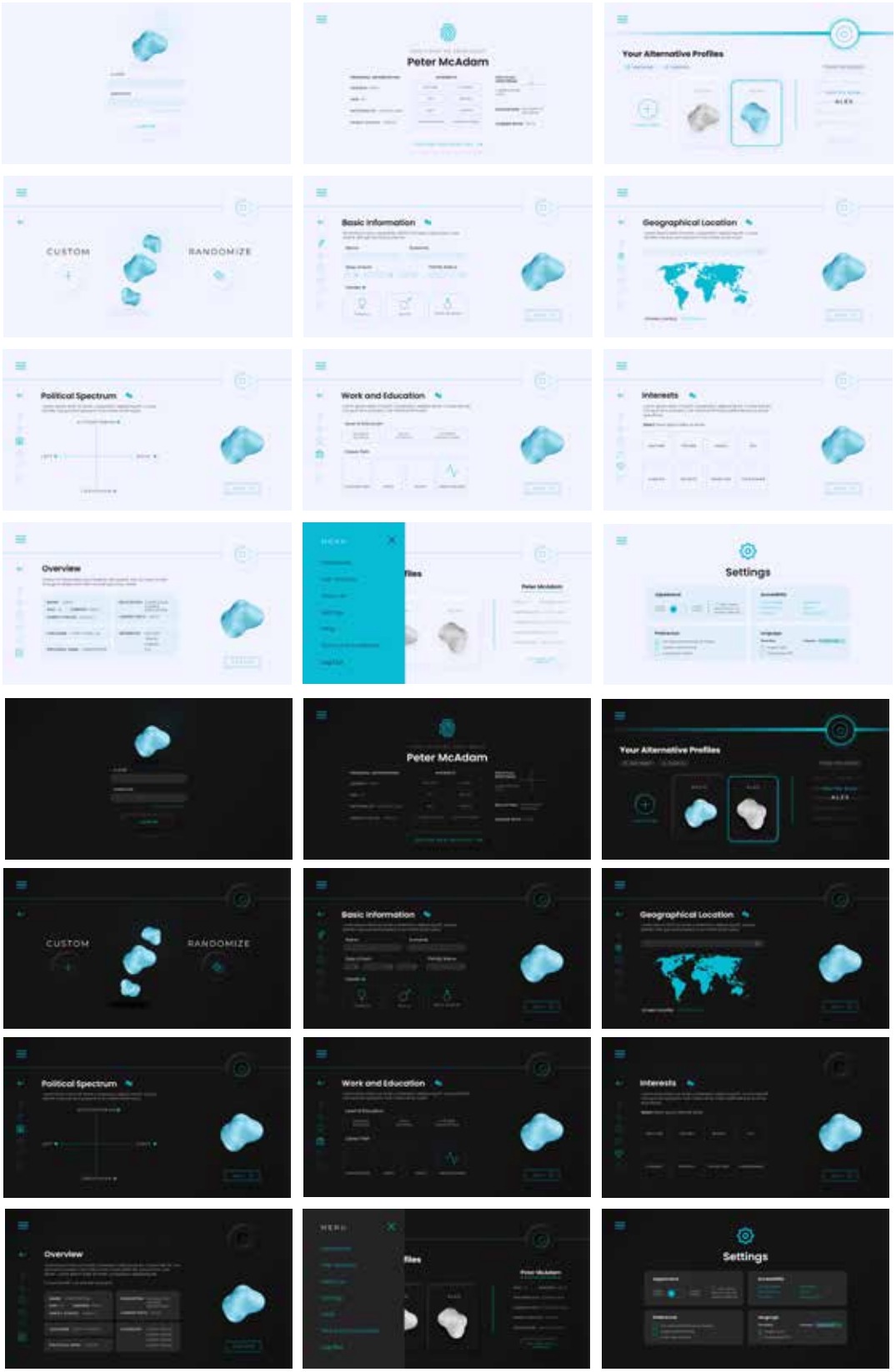
The user can increase or decrease his 'Interests' levels, or point out other items of interest that he wants to associate with the alternative profile under development. And at any point check, correct, change or delete each item of personal information associated with each alternative profile in 'Overview'. Or even enter random information by clicking on the 'data' icon available in each section, and/or activate the 'random' option in the periodic automatic switching of alternative profiles created in GLOBULE following the options 'Explore New Realities' and 'Randomize'.

GLOBULE allows the creation of alternative profiles, in order to check or correct data, interfering with the information collected by the algorithm according to the user's needs:

CONTROL search results, so that they are less filtered and influenced, debugging and redirecting them, surpassing previous activity records and providing access to a spectrum other than your own;

BOYCOTT, deceive and confuse the system, or those using the data collected, in order to avoid any kind of harassment; or

PROTECT your privacy by managing the data that becomes available online and making it difficult to access it.




Visual Image / UI Design

### 4.2 Functional prototype

https://framer.com/share/Prototype-GLOBULE--srkzRyWO9Y0ciW6hVtvl/zPWFZiOrkF0qQ9qPyySVox-1%3A11110u3nveraru5g

### 4.3 Video (Teaser)

https://www.youtube.com/watch?v=ssMVxPuhEL8

## 5 CONCLUSION

To sum up, throughout the project, our goal was to create a means to battle data surveillance and, consequently, protect our data from companies like Google or Facebook, that are continuously trying to access their customers' private information.

As a result of that intention, we designed Globule, an application that allows the user to control which data is shared on-line, therefore managing his/her search results. Each user can create as many profiles as needed, or simply choose a random profile (just to not share private information).

It is also import to point that one of our priorities was to combat the prejudice and stereotypes such as the frequent use of the male body to represent the human race, that is why we chose a blob as our main symbol - a genderless mass that is in constant change (a metaphor that relates to the fact that the user is able to keep changing his data in the applicaton). There are also two appearances - light and dark modes - to accommodate the user to every environment (day and night, for example).

## REFERENCES

[1] https://www.media.mit.edu/projects/gobo/overview/

[2] https://smellofdata.com/

[3] https://www.ibtimes.co.uk/edward-snowden-designs-spy-proof-smartphone-case-warn-if-youre-being-monitored-1571928

[4] https://www.statista.com/statistics/272365/age-distribution-of-internet-users-worldwide/#:~:text=As%20of%202019%2C%20a%20third,aged%2018%20to%2024%20years

[5] https://ieeexplore.ieee.org/document/9262054

[6] https://www.rca.ac.uk/showcase/show-rca/lisa-mandemaker/