Vinh Tuan Dat Nguyen (Percy)

CART 451

Professor Sabine Rosenberg

Project Proposal: The Human Algorithm

1. **What is your theme/topic/goal/issue to be tackled - why is it important to you?**

As social media are integrated more closely with our daily lives, the fine line between convenience and privacy becomes thinner and thinner. The problem tackled is the ambiguous methods behind personal data collection by big corporations. While some social media specify which types of data they collect though fail to explain the algorithms and methods of collection. An example of data transparency is Instagram's option for users to download their data on a json file which includes a post of your account activity as well as an archive of images/videos that the user posted or shared to other users. Although this transparency explains which data is collected, it does not clarify how it is utilized to accommodate each user's feed.

The algorithms that they employ to keep users engaged have never been publicly disclosed by the technology conglomerates. For now, the only valuable insights into these algorithms are given by former employers and leakers. Otherwise, content creators are pushing through waves of trials and errors until fortune, the god of algorithm, smiles on them.

The project, in the form of a website, will showcase a record of my interaction with instagram on a given day. The showcase will display the major features that I usually engage

with. My interactions will be video recorded. These data will be posted on my website as a series of video recordings for users to navigate through.

At the end of the virtual exhibition, participants will be asked a few demands that will reflect their current understanding of me.

2. **What form will your project evolve into - who is your audience?**

Hypothesis

This project aims to pitch human intelligence against AI algorithms in the context of popular social network service. In doing so, one might decipher and understand the methods by which AI functions and the intention of its human designers.

I would like to think that this user testing experiment has previously been used by social networks to fine tune their AI to perfection.

Audience

To stimulate an algorithm, strangers are ideal candidates as their feedback will be uninfluenced by an existing perception of me. The project also includes three trials to allow the human algorithms to receive feedback and re-attempt with a growing knowledge about my digital preferences. A total of ten participants are ideal for this project.

I will ask participants to name three Instagram posts that I would engage with. Users must also justify their suggestion with three keywords for each post. These demands are repeatedly asked over the course of three trials to see if there has been an improvement after every feedback. My feedback can take on a combination of the following forms : like, dislike, comment, share, time engaged. These are also the only feedback that Instagram AI algorithms receive.

However, as this project touches on personal matters, only close friends will be invited to participate in the first set of the experiment. If all goes well, I will make appropriate changes to the project then entrust it to credible acquaintances. They could be individuals recommended by my friends and family from a wide range of age and background.

For users who value entertainment rather than the moral implications in play, the collection of private data benefits and enriches their personal experience. However, for paranoid and curious individuals like myself, I would like to learn the specific procedures that operate these algorithms. At the moment, I vaguely understand that Instagram AI loves sharing similar content that I engage with.

3. **Discuss how each of the two readings listed above have inspired/motivated your current choices with regards to the project.**

The first reading warned me about the possibility of apophenia and pattern misrecognition during the data analysis process.

In order to prevent apophenia and misrecognition of the data given, I create multiple trials to allow participants to change and/or reinforce their conclusions overtime. Therefore, participants may produce more accurate results and compete with AI algorithms more adequately. I must equally present the data in a neutral and objective manner without any accidental favoritism during the collection and exhibition processes. My usage with Instagram will be presented in chronological order to not favor the placement of any particular content.

The Point of Collection by Mimi Onouha raises the importance of the means and methods by which data are collected which in turn dictate its meanings. If the participants

come from similar backgrounds and power hierarchy, the resulting data can both be true though lacking diversity and therefore misrepresenting the parties in question.

4. **What is your data? How will you collect it?**

There are two set of data in this project:

Dataset 1: My user experience on Instagram:

The first set of data is a collection of my behavior, my preference and my digital engagement. It involves not only what I consume, but also the manner and speed at which I consume them. This collective data, carefully detailed and annotated will give participants insights into my life.

To ensure that all interactions will be documented, I will make a new password for my phone for the respective day as a reminder to actively record whenever I engage with the app.

For search, feed and ads, I will screen record my engagement with each content. Any content by a private user will be censored, while public users will still be recorded.

Regarding privacy of the parties involved, I will only show the content of public accounts (including friends and family). In Instagram, any content shared by public accounts is visible to anyone on the net (including users without an account) and thus as per Instagram's policy I should be able to share them freely.

Dataset 2: The human algorithm's suggestions:

The second data set is the suggested posts and their respective keywords from the ten participants. They will be collected from the three trials.

If time allows, more tests will be added but this is a good base to start the project.

Since friends are the participants of the project, my demand should be challenging given that they already have a considerable advantage over strangers.

My original idea was demanding a one sentence feedback description about me, although that seems far too human for an AI algorithm. Thus by having participants categorizing me into short, concise key words would likely be the way AI algorithms perceive humans.

5. **What medium(s) do you intend to use and why?**

The project will be a highly responsive website made with HTML, CSS, Javascript, JQuery and Node.js.

Front end wise, I am confident in my ability to create a pleasant interface for the data exhibition as I am comfortable with HTML and CSS. The challenge lies in implementing user feedback with Node.js and hosting options that I will need to research into. Every guess and their respective reasonings are stored into an array in a json file.If I cannot complete this within the deadline, an emergency alternative is to link a google survey to my website where users may still upload their responses and receive feedback through email.

I chose the web because its flexibility allows me to display datas in different mediums while simultaneously allowing users to have a more interactive experience.

6. **At a very high level : what are the algorithm(s) that will be used and implemented to achieve your intentions?**

For suggestions made by the users, I will rank them on a scale of ten out of satisfaction. Each participant's individual score accumulates over the course of the

experiment. The secret scoreboard ranking is hidden from the participants and may only be revealed after the project.

This ranking system serves to concretize any improvements or diminishment of the participants' effectiveness at algorithms. I expect that most participants will improve their score with time. I hope that I may be able to draw patterns from the top and bottom players.

Another approach to detect improvement is to start out with simple demands and slowly ramp up the precision to see if participants can still satisfy more specific requests. However this method would be far too time consuming for it to be effective.

Inspirations:

1. **Lauren by Lauren-McCarthy (2017)**

   Project link: https://lauren-mccarthy.com/LAUREN

The strongest inspiration for this project is the Lauren project by Chinese-American artist Lauren Mc-Carthy.

Lauren takes on the role of a virtual assistant to different people's homes, similarly to Amazon Alexa. This project explores and encompasses the following themes:"the tensions between intimacy vs privacy, convenience vs agency they present, and the role of human labor in the future of automation" (). By installing custom designed networked smart devices, Lauren is able to control those devices as to the habitant's wishes. Her goal is to outperform the AI by understanding the individual instead of catering to each demand. She admits to forgetting that these projects are not about connecting with someone but rather imitating the role of technology in an humane approach.

I do find her project a slight inaccuracy as Amazon Alexa does not have camera recording while Lauren does. This gives Laurent a lot more data and unfair advantage over Alexa. Perhaps this was intended as Lauren's unique approach to her project.

While my project examines human-machine relationship by having human participants imitate the AI, its ultimate objective is to understand the AI rather than surpassing its capability. For this specific and repeatable task of sorting out content, the AI has a clear advantage over humans as they are unbounded by the mortal's survival needs. I hope that upon the completion of the project, I may be able to draw some patterns that will reflect the method that AI algorithms function.

2. **The Social Dilemma (2020)**

The Netflix docudrama directed by Jeff Olowski highlights major tech companies' objective of engaging users and netting profit despite their manipulation on humanity.

While social media encourages freedom of speech and advocacy, ill intended users and political parties may take advantage of these tools to spread propaganda. As social media construct a particular echo chamber for each user which then reaffirms their truth, this causes political polarization and radicalization amongst youth.

One may only address these issues by understanding the core of the issue - what information is being fed to these algorithms and what procedures do they follow to push out the most relevant content for users. How are these algorithms so accurate? Are my personal data revealing far more than I intend them to? How can these algorithms have such prowess and be able to shape the newer generation?

Combating the addiction and radicalization caused by social networks is one of the final goals of this project. But to be able to do so, one must first comprehend these black box algorithms.

3. Counter Narrative by Alexandre Bell (2017)

Learn more about the artwork:

https://artnowla.com/2019/04/21/alexandra-bell-counternarratives/

Counter narratives are a series of six New York Times news pages rewritten in an objective light. As the original news were racially biased and misleading, Alexandre Bell re-edits, redacts and highlights them with markers to point out untrue prejudice and assumptions regarding race. Both the corrected original and its revised counterpart were printed in large scale and posted around New York city.

This series of artworks encourage me to exercise caution and refrain from applying my own biases throughout the project. It invites the data collection and dissemination processes to be as raw as feasible.

Mockup of final project:

Phone Perspective

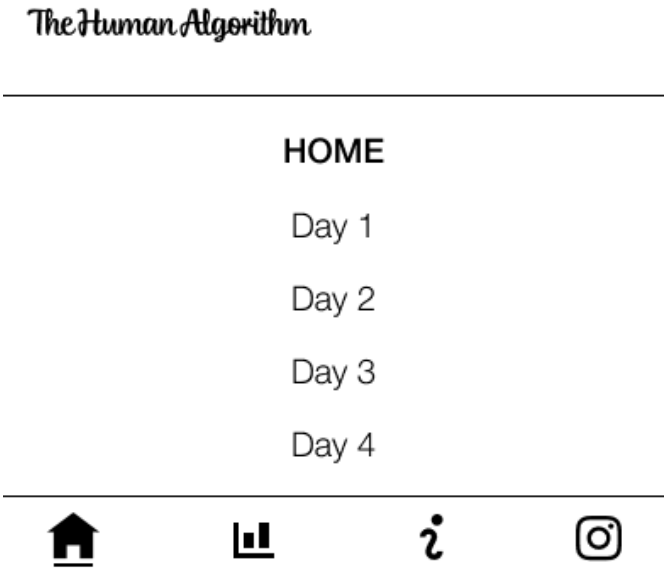

Fig 1 (above). Home button prompts participants to choose a week.

Fig 2 (right). Once a user chooses a day, Percy's activity throughout the day is displayed.

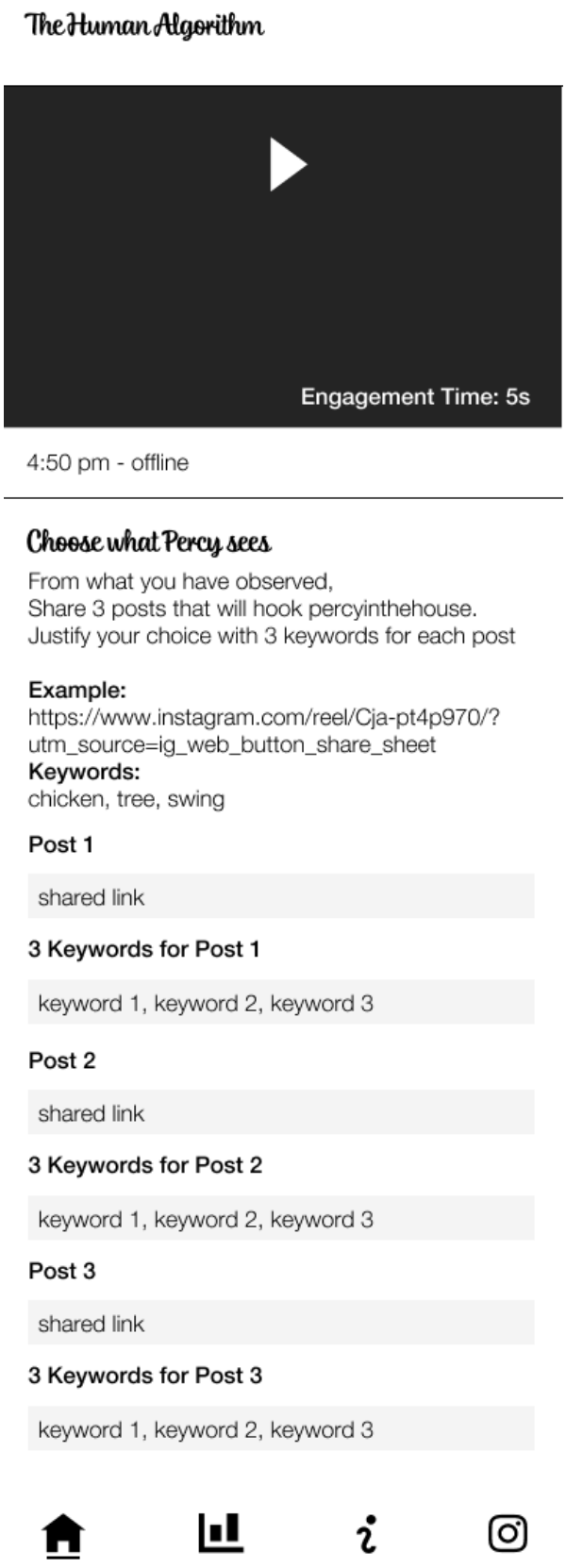

Fig 3 (left), After the activities section, text boxes and instructions are offered for participants to suggest 3 posts and keywords.

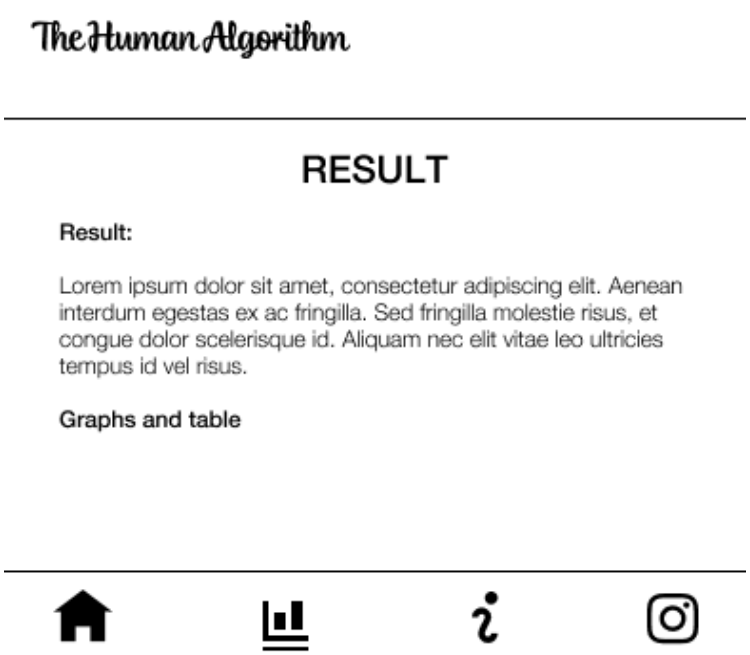

Fig 4. Result section where I display all of the gathered answers and their reasonings as well as the ranking board.

Any conclusions for this project will also be included.

Fig 5. The About section serves to explain the purpose of the project.

Desktop View

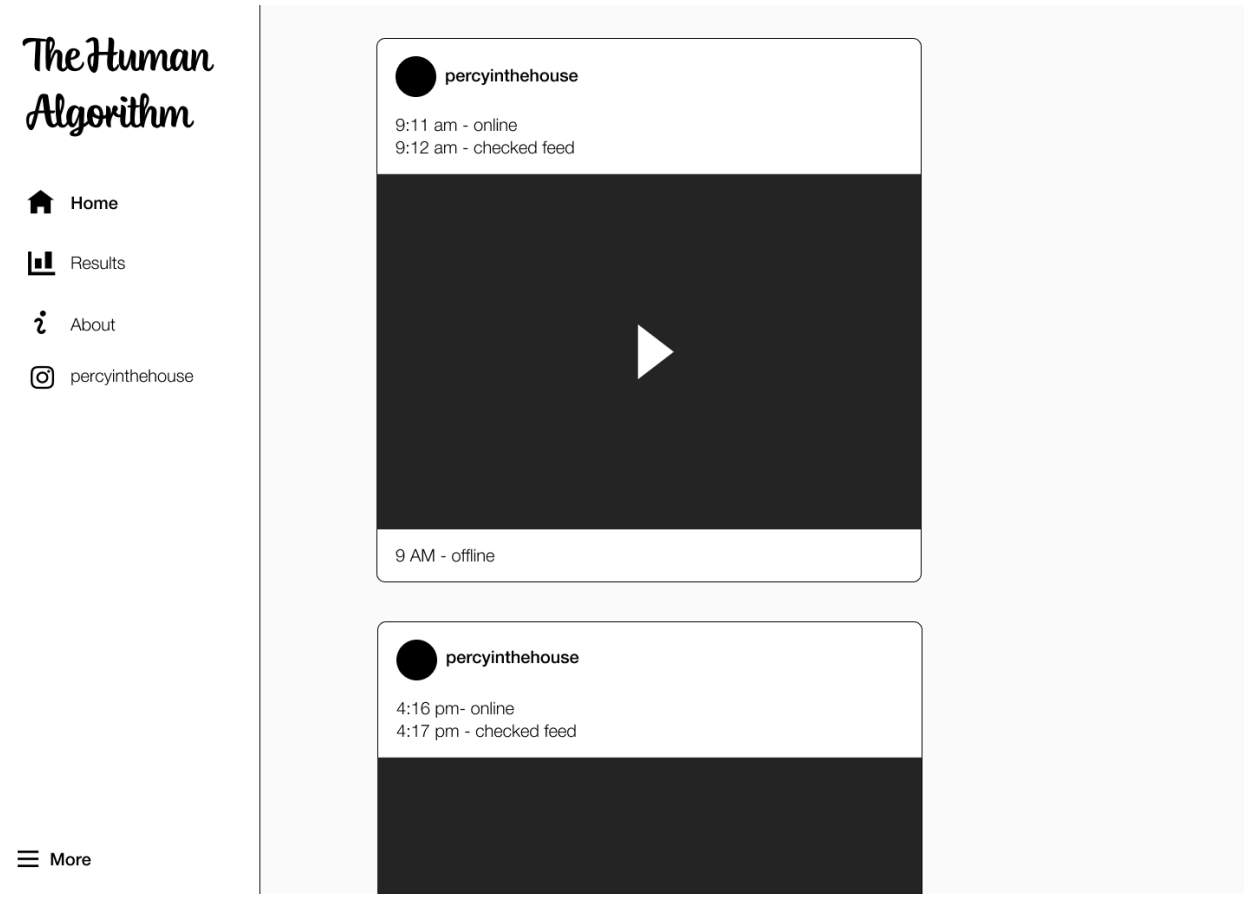

Figure 6. How the layout will be designed in desktop view.