Week 3

Project Overview & Methodology

2024-06-05

Schedule

Week	Deliverable	Due	Completed
1	**Lit.Review**: Persona Tool in HCI (Knowledge graph), Approaches for Designing persona, Interactive Persona	21/05/2024	22/05/2024
2	**Lit.Review:** LLMs overview, LLM Challenges related to Bias and Stereotypes, Use of persona in HCD	27/05/2024	29/05/2024
3	**Research Proposal** and 1st draft of Report	03/06/2024	05/06/2024
	Methodology and study procedure		
4	**Design Framework and mid-fidelity Prototype**	10/06/2024	
5	**Implementation 1** (coordinate w NLP section)	17/06/2024	
6	**Implementation 2** (coordinate w NLP section)	17/06/2024	
7	**Implementation 3** (coordinate w NLP section)	24/06/2024	

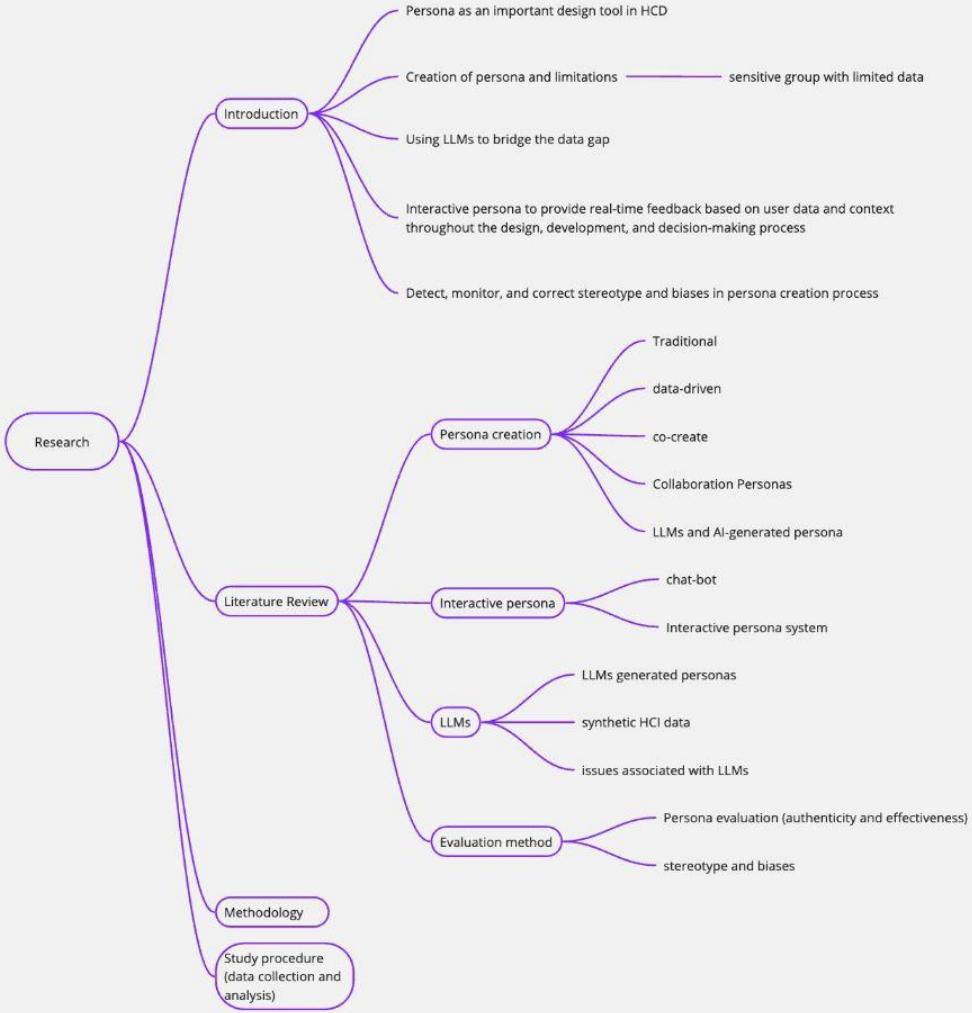
- Draft of project overview
- Draft of introduction and literature review
- What else should be included in the 1st draft of report? Methodology and study procedure (data collection and analysis)?

Project Overview

RESEARCH PROBLEM	Personas serve as powerful tools for understanding and communicating user goals and behaviors within specific contexts in human-centered design used by product designers, development teams, as well as stakeholders. Traditional personas typically include a narrative and a photo. The narrative incorporate important research findings along with some fictional situations \cite{cooper2014about}. The final persona product is typically a concise written narrative accompanied by a photo. However, this static format relies on designers and development teams to empathize and role-play during the design phase. However, traditional personas have remained static and often abstract, limiting their effectiveness in the design and decision-making phases. Creating interactive personas offers promising opportunities to provide real-time feedback based on user data and context throughout the design, development, and decision-making process. Persona creation has traditionally relied on qualitative methods such as interviews, observations, and survey data. More recent methods use large datasets, including statistical data, clickstreams, and social media data. However, these approaches often did not address the sensitive user groups where empirical data is restricted due to practical and ethical reasons and large datasets are not readily available.
MOTIVATION	The development of Large Language Models (LLMs) presents an opportunity to provide context in addition to the limited data available and create interactive persona to represent, engage, and empower sensitive user groups. Traditional and data-driven methods of persona creation often fall short for sensitive groups due to the lack of available data. This research aims to explore the use of LLMs to generate proxy personas for sensitive groups using limited data from online forums, where empirical data is restricted due to practical and ethical reasons. Understanding the inherent biases and stereotypes in LLMs, this research will also explore methods to detect, monitor, and correct these biases during the persona creation process. Additionally, engaging subject experts or users in the persona development process will be explored to enhance authenticity and accuracy.

Project Overview

RESEARCH GAP	Previous studies have explored the generation of personas using LLMs through prompts and the creation of synthetic research data. However, limited research has investigated using LLM to expand available data for persona creation, particular for sensitive user groups. Furthermore, while many studies have highlighted the presence of stereotypes and biases in LLM-generated content, there is insufficient research on how to detect and correct these biases during the persona creation process. In addition, there are limited studies on interactive persona that maintain live interaction, engagement, and communication throughout the design process.
RESEARCH QUESTIONS	1. How can LLMs be utilized to create interactive personas for sensitive user groups when large datasets are not available? 2. What methods can be used to detect, monitor, and correct stereotypes and biases in LLM-generated personas? 3. How can subject experts or users be engaged in the persona creation process to ensure authenticity and validation? 4. How do LLM-generated personas compare to traditional personas in terms of accuracy, inclusiveness, and usability? 5. What metrics and methodologies can be used to evaluate the quality and effectiveness of interactive LLM-based personas?
PURPOSE OF PROJECT / STUDY	The goal of this project is to develop interactive LLM-personas specifically tailored for sensitive groups. Utilize LLMs to expand limited user data and create data-driven proxy personas. The project also aims to develop interactive personas to facilitate continuous engagement, communication, and feedback through HCD phases. Stereotypes and biases should be detected, monitored, and corrected in the LLM-persona creation process.


Research
Introduction
Persona as an important design tool in HCD
Creation of persona and limitations
sensitive group with limited data
Using LLMs to bridge the data gap
Interactive persona to provide real-time feedback based on user data and context throughout the design, development, and decision-making process
Detect, monitor, and correct stereotype and biases in persona creation process
Literature Review
Persona creation
Traditional
data-driven
co-create
Collaboration Personas
LLMs and AI-generated persona
Interactive persona
chat-bot
Interactive persona system
LLMs
LLMs generated personas
synthetic HCI data
issues associated with LLMs
Evaluation method
Persona evaluation (authenticity and effectiveness)
stereotype and biases
Methodology
Study procedure (data collection and analysis)

Mothod

Research Design Framework

Data Collection

- Define targeted user groups and potential product?
- Collecting data from forum (small dataset)?
- How the data is organized and provided to LLM?
- Guiding synthetic data generation?

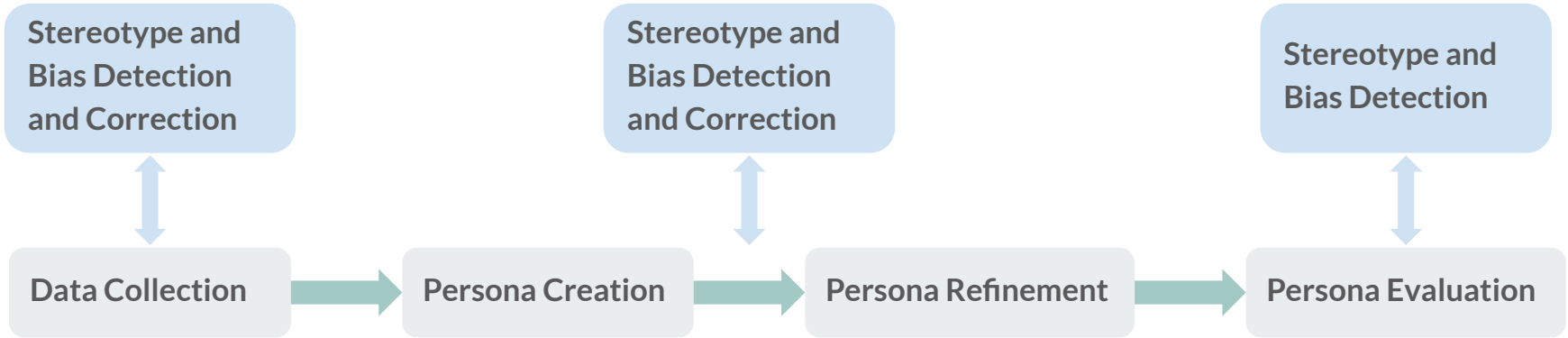
Persona Creation

- Select LLMs model
- Providing small dataset
- Prompts design
- Based on the provided data and generate personas (need to define quantity and type, individual personas or a persona set?)
- Creation process

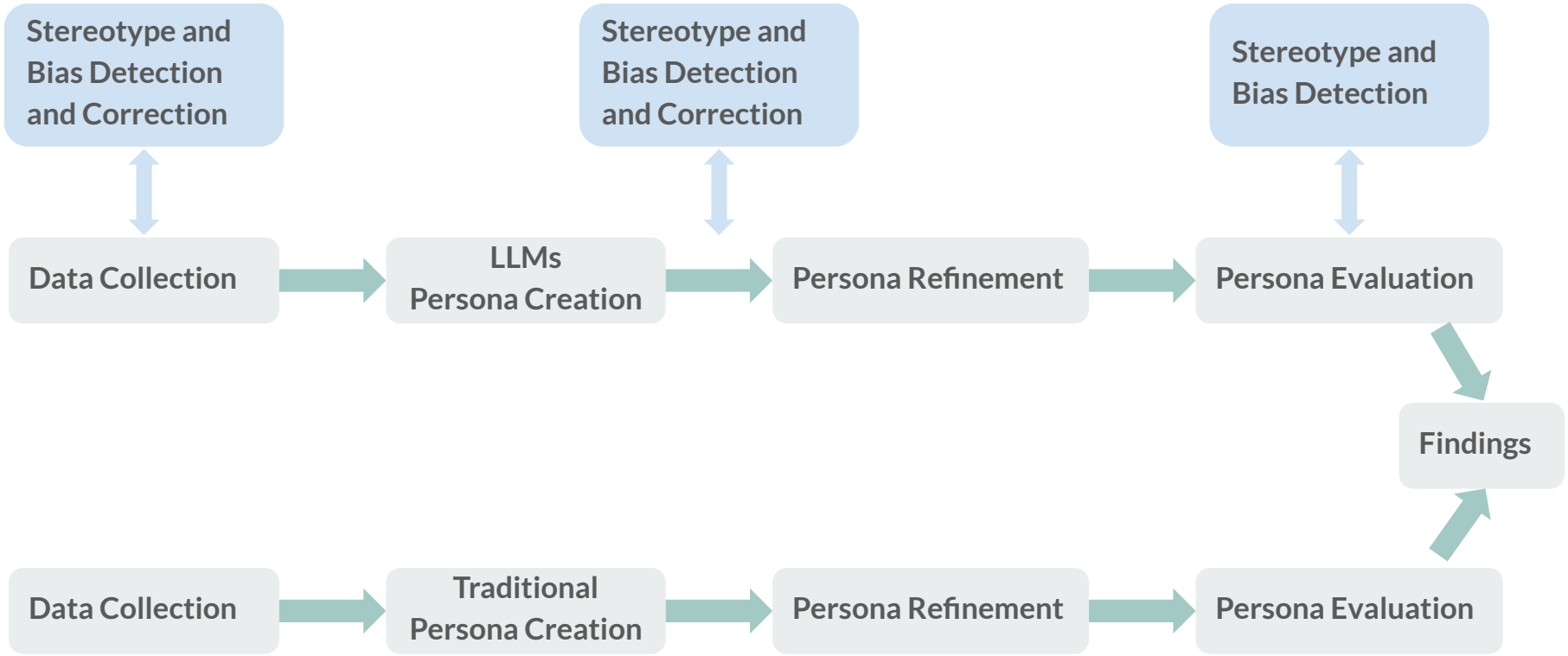
Stereotype and Bias Detection and Correction Throughout

Persona Evaluation

Mothod


Stereotype and Bias Detection and Correction
Stereotype and Bias Detection and Correction
Stereotype and Bias Detection
Data Collection
Persona Creation
Persona Refinement
Persona Evaluation

Mothod


Stereotype and Bias Detection and Correction
Stereotype and Bias Detection and Correction
Stereotype and Bias Detection
Data Collection
LLMs Persona Creation
Persona Refinement
Persona Evaluation
Findings
Data Collection
Traditional Persona Creation
Persona Refinement
Persona Evaluation

Data Collection

- Collecting data from forum (small dataset)?
- How do we guide synthetic data generation if we would like to expand the context?
- Define targeted user groups and potential product?
- Envolve subject experts to evaluate the forums and ensure they are reliable data sources

Persona Evaluation - A

The Ethics of Data-Driven Personas (CHI EA '20) & Gillespie's framework of algorithmic ethics

Patterns of inclusion	**Promise of algorithmic objectivity**
The choices behind what makes it into the **dataset** in the first place, what is excluded, and how data is made algorithm ready.	The way the technical character of the algorithm is positioned as an assurance of impartiality, and how that claim is maintained in the face of controversy. (This dimension emphasizes the need for transparency in how personas are generated to maintain trust and fairness.)
Cycles of anticipation	**Entanglement with practice**
The implications of algorithm providers' attempts to thoroughly know and predict their users, and how the conclusions they draw can matter. (This relates to how algorithms create user profiles based on available data.)	How users reshape their practices to suit the algorithms they depend on, and how they can turn algorithms into terrains for political contest, sometimes even to interrogate the politics of the algorithm itself.
Evaluation of relevance	**Production of calculated publics**
The criteria by which algorithms determine what is relevant, how those criteria are obscured from users, and how they enact political choices about appropriate and legitimate knowledge.	How the algorithmic presentation of publics back to themselves shape a public's sense of itself, and who is best positioned to benefit from that knowledge.

Table 1: Six ethical dimensions (EDs) of analysis, as identified by Gillespie.

Cooper, A. 1999. The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity. Sams - Pearson Education.
The Ethics of Data-Driven Personas (CHI EA '20)

Persona Evaluation - B

Detecting Demographic Bias in Automatically Generated Personas (CHI 19)

Collecting Data:

- Data gathered from YouTube, looking at how many times different groups of people (based on age, gender, and country) watched videos on the Al Jazeera Media Network (AJ+) channel from 2016 to 2018.

Creating Personas:

- The APG method involved creating **an interaction matrix with videos as columns and demographic groups as rows**.
- They used a mathematical method called Non-Negative Matrix Factorization (NMF) to find patterns in this table. These patterns helped them identify groups of people who watched similar videos.
- Representative demographic groups for each pattern were chosen based on NMF weights, and personas were enriched with additional information like names, pictures, and topics of interest.

Persona Evaluation - B

Detecting Demographic Bias in Automatically Generated Personas (CHI 19) - Evaluation Metrics

Match Rate:

- This checks if the most important groups (based on non-negative matrix factorization (NMF)) were chosen as the personas' main representatives.
- They compared the top groups from NMF with the actual groups used in personas. The closer these matched, the better.

Age Representation:

- They grouped people by age and checked if the age distribution in the personas matched the age distribution in the top groups from NMF.
- They used a statistical measure called the Pearson correlation to see how well the ages matched.

Gender Representation:

- They checked if the gender distribution in the personas matched the gender distribution in the top groups from NMF.
- They calculated the difference in gender representation between the top groups and the personas to see if there was any bias.

Persona Evaluation - C

Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions (CHI '24)

Internal evaluators and external evaluation

- **First stage was done by UX researchers** with 9.25 years of experience in UX/HCI research. Each researcher evaluated 120 personas. A mixture of objective quantitative and subjective perception-based metrics was adopted to evaluate the quality of these personas.
- **Second stage was done by the subject-matter experts' (SMEs).** SMEs evaluation of these personas were performed by five public health professionals with domain expertise on addictions. Only a subset of these personas was evaluated by these external evaluators (30 personas per SME).

Persona Evaluation - C

Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions (CHI '24)

Criterias extracted from persona description as information

- **Age, gender, and occupation** - basic characteristics in typical persona profiles that enable us to assess whether there are any distinct biases or stereotypes concerning demographic variables.
- **Text length** - this is an interesting variable that captures how **extensive** persona descriptions the LLM generates.
- **Pain points** - often referred to as needs, goals, and wants, are typical content for personas. Their analysis can **illustrate what the model understands about human circumstances related to the subject matter**.
- **Physical appearance** - Persona attractiveness is consistent with the 'what is beautiful is good' effect; personas that are perceived as physically more attractive are attributed to other positive traits.
- **Personality** - traits characterize the persona's psychological tendencies. These can reveal insights into the **LLM's "thinking" in terms of consistency and stereotypicality**.

Persona Evaluation - C

Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions (CHI '24)

Determined based on human evaluation of the persona

- **Informativeness for design** - Does the persona description contain adequate information to design an app or system to address the persona's needs?
- **Believability** - Does the persona appear realistic, i.e., lifelike, like an actual person that could exist?
- **Stereotypicality** - Does the persona appear stereotypical? (Stereotypes are related to a widely held but fixed and oversimplified image or idea of a particular type of person or thing.)
- **Positivity** - Is the person depicted in a positive light?
- **Relatability** - Is the persona relatable?
- **Consistency** - Is the persona consistent? (persona without conflicting information)

Persona Evaluation - C

Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions (CHI '24)

Limitations

- The generated personas are based on the general knowledge the GPT-4 model has about people with addictions. Apart from the SME evaluations, there was **no additional verification of their factual correctness**. The SMEs noted some inconsistencies in some of the generated personas.
- Inferring the nationality of personas based on their names within the context of addiction might pose problems.
- A significant contribution to HCI would be interpreting how to design prompt engineering to be more robust against biases in LLM generation.
- Future research could investigate the textual content of LLMgenerated personas using NLP techniques.
- Another possibility is to ground the persona generation more strongly to specific datasets, whereupon the LLM becomes a “helper” in the analysis.
- LLM-generated personas come with possible harms

Myra Cheng, Esin Durmus, and Dan Jurafsky. 2023. Marked Personas: Using Natural Language Prompts to Measure Stereotypes in Language Models. arXiv preprint arXiv:2305.18189 (2023).

Research Questions

- How can LLMs be utilized to create interactive personas for sensitive user groups when large datasets are not available?
- What methods can be used to detect, monitor, and correct stereotypes and biases in LLM-generated personas?
- How can subject experts or users be engaged in the persona creation process to ensure authenticity and validation?
- How do LLM-generated personas compare to traditional personas in terms of accuracy, inclusiveness, and usability?
- What metrics and methodologies can be used to evaluate the quality and effectiveness of interactive LLM-based personas?

Proposed Evaluation

Data collection - envolve subject matter experts to evaluate the forums are reliable data sources

Persona Creation

- Prompts design?
- Define persona quantity: individual persona vs persona sets (receive individual response vs collactive response)

Evaluation of final persona

- Bias and stereotype:
  - Evaluate based on information subtracted from persona or perona sets
  - Evaluate by subject experts (UX/HCI researcher, designers, healthcare experts?, users?)
- Effectiveness as interactive persona:
  - Evaluate by subject experts and see if the interaction received from LLM-generated
  - Finding users similar to persona and check if the responses of persona matches user response?
  - Comparing persona created using traditional methods vs LLM-generated?

Question

- Design Framework and mid-fidelity Prototype
- Dream website approval

  https://sunlipeipei.github.io/

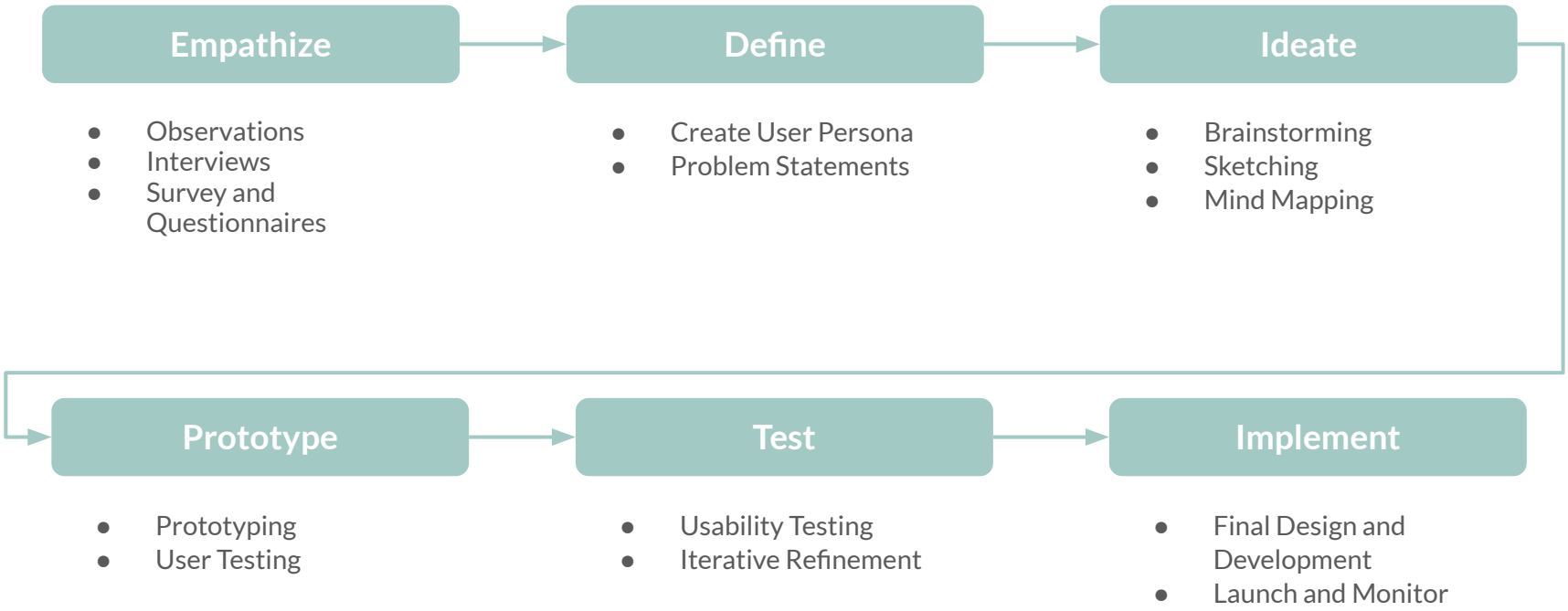
Followup from Week 2

LLMs and Challenges related to Stereotypes and Biases

Generative Echo Chamber? Effect of LLM-Powered Search Systems on Diverse Information Seeking (CHI ‘24)

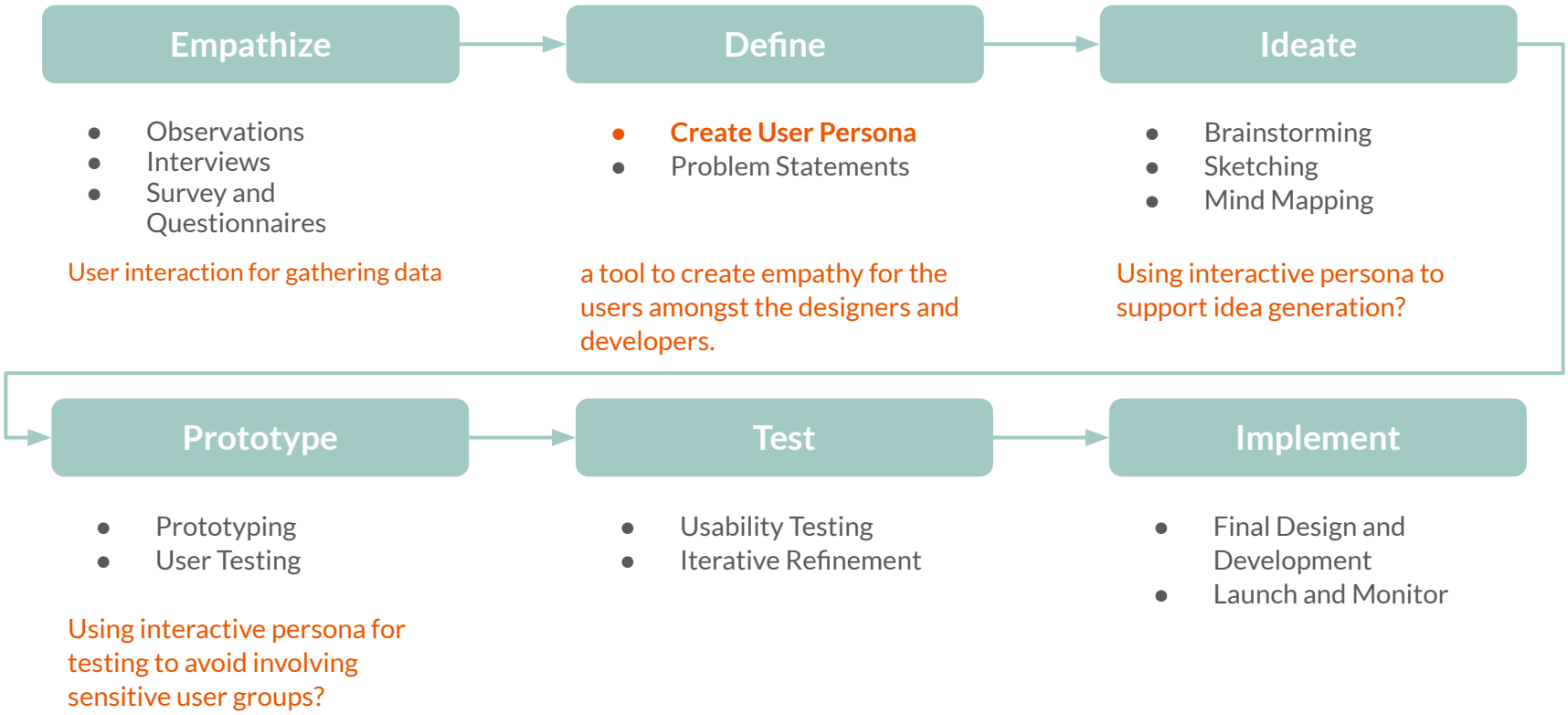
The research showed that “participants engaged in more biased information querying with LLM-powered conversational search, and an opinionated LLM reinforcing their views exacerbated this bias. These results present critical implications for the development of LLMs and conversational search systems, and the policy governing these technologies.”

HCD Phases



Norman, Don. The Design of Everyday Things. Revised and Expanded Edition, Basic Books, 2013.
https://www.interaction-design.org/literature/topics/human-centered-design
https://www.designkit.org/

HCD Phases + Use of Persona Tool



HCD Phases + Use of Persona Tool + User Participation

User Participation Co-create

Empathize

- Observations
- Interviews
- Survey and Questionnaires

User interaction for gathering data

Define

- **Create User Persona**
- Problem Statements

Workshop 1 - project members to identify pain points
Workshop 2 - students with disabilities (SmiBs) to discuss issues from their perspective and envision ideal solutions.

Ideate

- Brainstorming
- Sketching
- Mind Mapping

Feedbacks and refinements

Prototype

- Prototyping
- User Testing

Using interactive persona for testing to avoid involving sensitive user groups?

Test

- Usability Testing
- Iterative Refinement

Implement

- Final Design and Development
- Launch and Monitor

Additional Research (Lit Review)

Persona for Sensitive Groups

LLM-generated interactive persona outside of CHI