

User-Centric Design of an Exploratory Search System for Scholarly Entities in Natural Language Processing

Ferdy Dermawan Hadiwijaya

17.07.2023, Bachelor's Thesis Kickoff Presentation

Chair of Software Engineering for Business Information Systems (sebis)
Department of Computer Science
School of Computation, Information and Technology (CIT)
Technical University of Munich (TUM)
www.matthes.in.tum.de

1. Introduction and Motivation
3. Research Questions
4. Methodology and Current Status
5. Next Steps

User-Centric Design of an Exploratory Search System for Scholarly Entities in Natural Language Processing

User-Centric Design

The How

Exploratory Search System for Scholarly Entities

The What

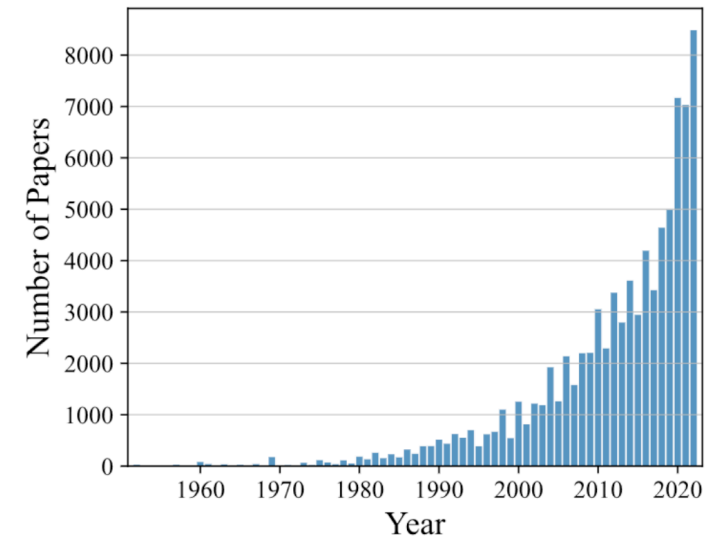
Natural Language Processing

The Why

Introduction - “Why” ?

Natural Language Processing

- The rate of publication for new research is increasing exponentially.
- These publications are available in large quantities as unstructured text.
- This makes it difficult to get an overview of new or unknown scientific fields.
- Furthermore, it is challenging to stay up-to-date with newly published research.



Number of papers released in NLP domain over the years

Introduction - “What”

Exploratory Search System for Scholarly Entities



An exploratory search system can highlight connections and trends in NLP research.



This thesis delves into an exploratory search solution through a **web application**.



Comparable existing search engines exist and will be discussed as part of initial findings.

- RQ1 What are the existing approaches of researchers to search, explore, and keep up with the NLP research?
- RQ2 How can we curate and present information in a web application to support user-friendly search and exploration of scholarly entities in NLP?
- RQ3 What approaches can we use to achieve a performant semantic search and exploration of relevant scholarly entities in Natural Language Processing?
- RQ4 How can we systematically evaluate the usefulness of our proposed approach to our target users?

RQ1 Existing Approaches

User interview with two NLP Researches at TUM

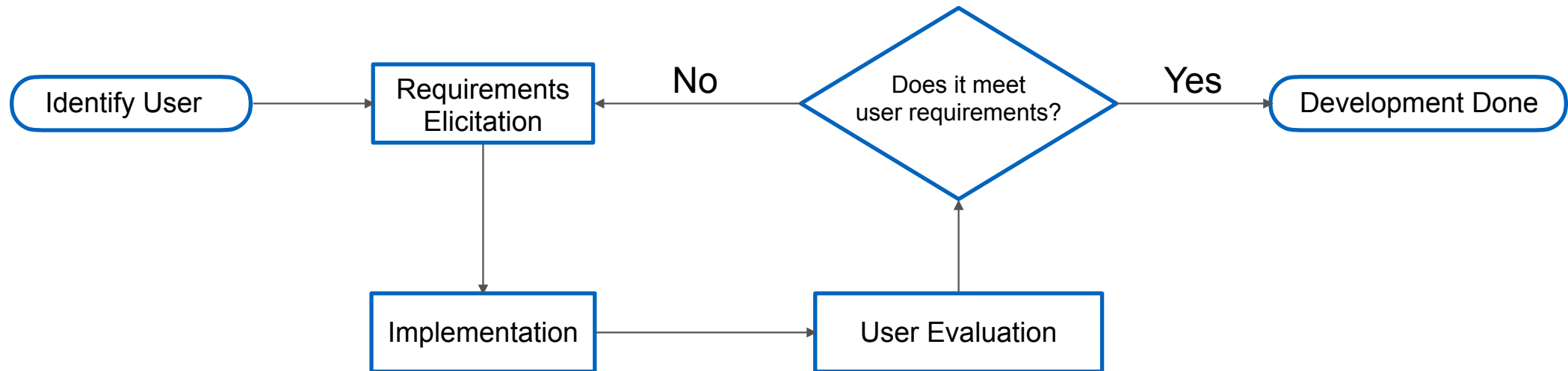
Result Discovery: Social Media, Individual venues, researchers, and institutions.
 Search: Search engines and online database.

Findings

- The information is spread across various tools making it challenging to identify connections within the research field.
- No NLP-specific solution exists.
- The search engine cannot fully understand the subtle meanings of the search queries.
- The identified pain points align with our initial hypothesis.

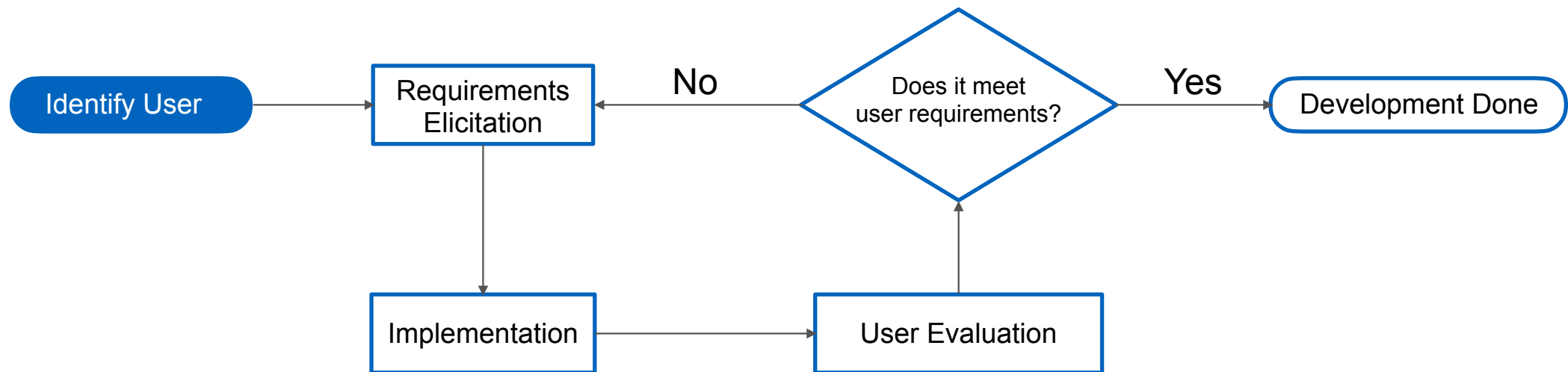
RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design



RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design



RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design

Identify User

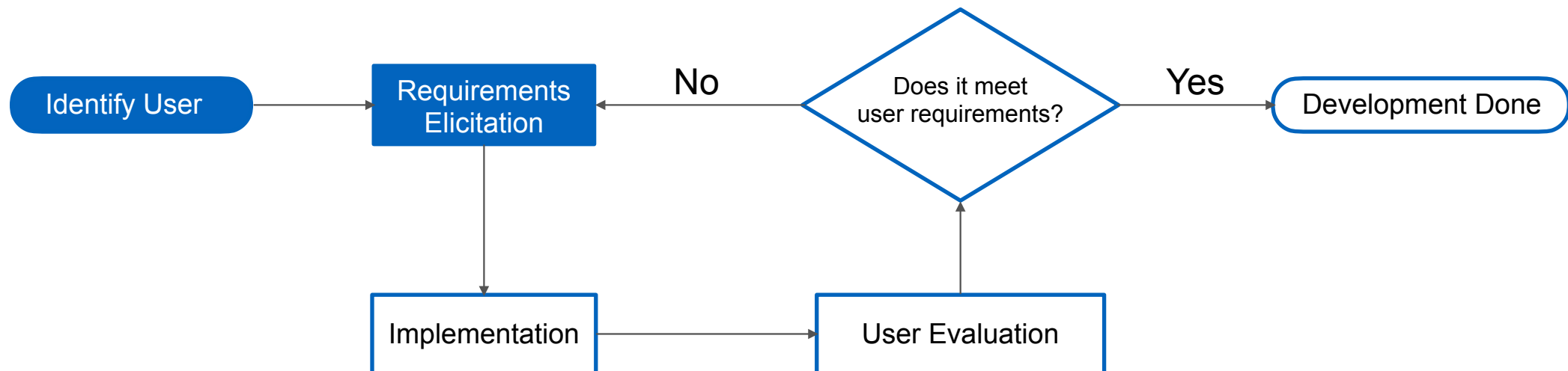
Persona



Researcher at different stages

RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design



RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

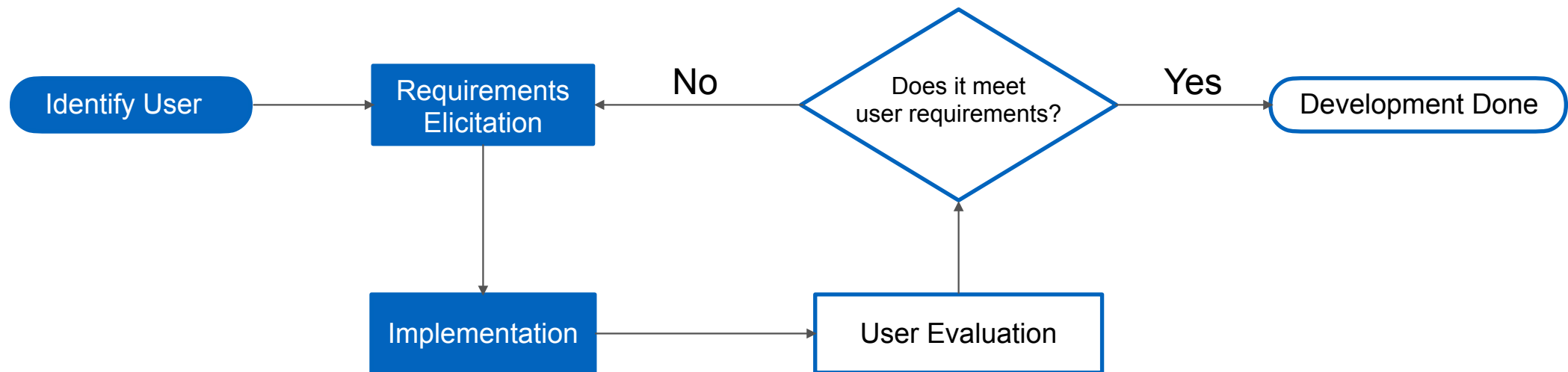
Iterative User-Centric Design

Requirements Elicitation

- Familiar design and low learning curve.
- Ability to navigate and discover connections between scholarly entities.
- Clear metrics and labels as indicators for significance and relevance.
- Comprehensive collection of relevant papers and entities.
- Relevant search results and options to sort and filter.

RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design



RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

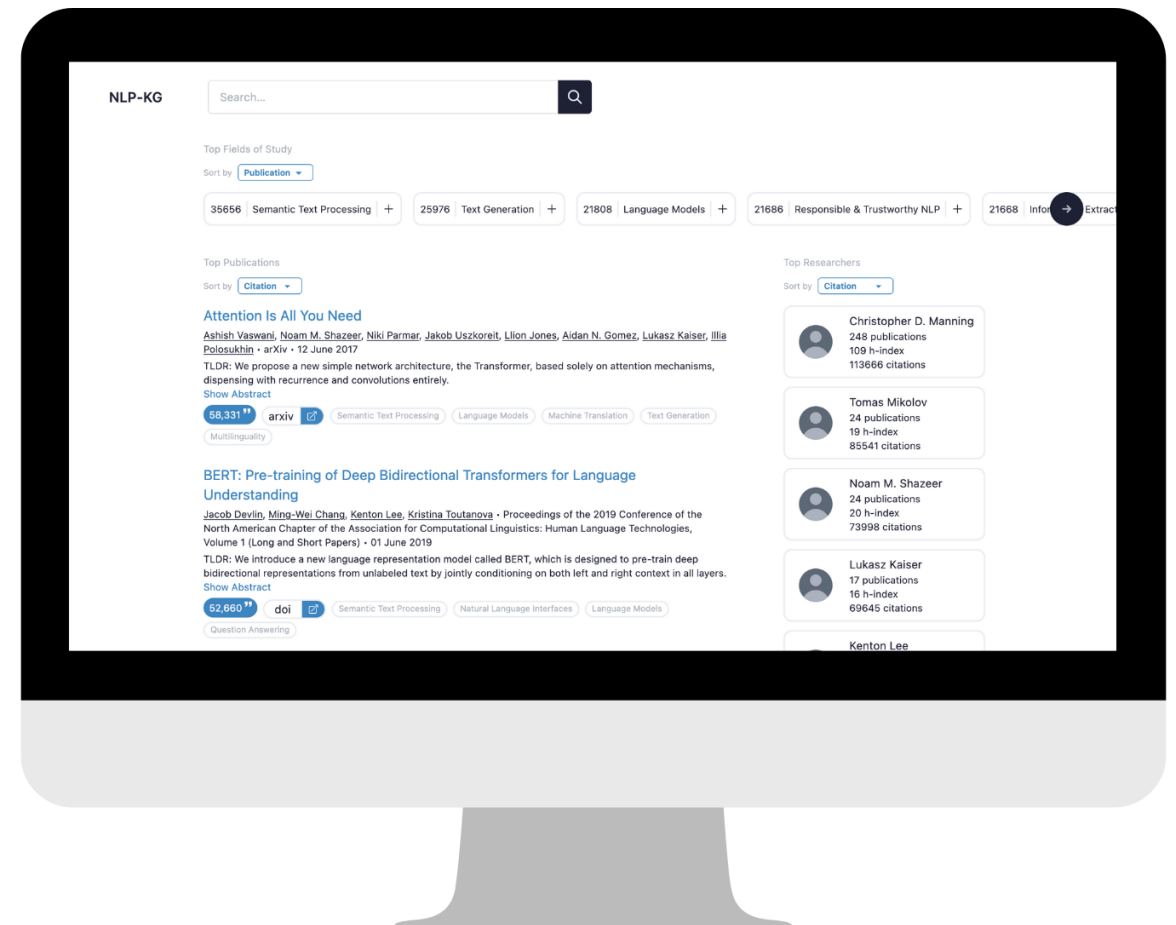
Iterative User-Centric Design

Implementation

Designed and implemented a **first iteration** following user's requirements.

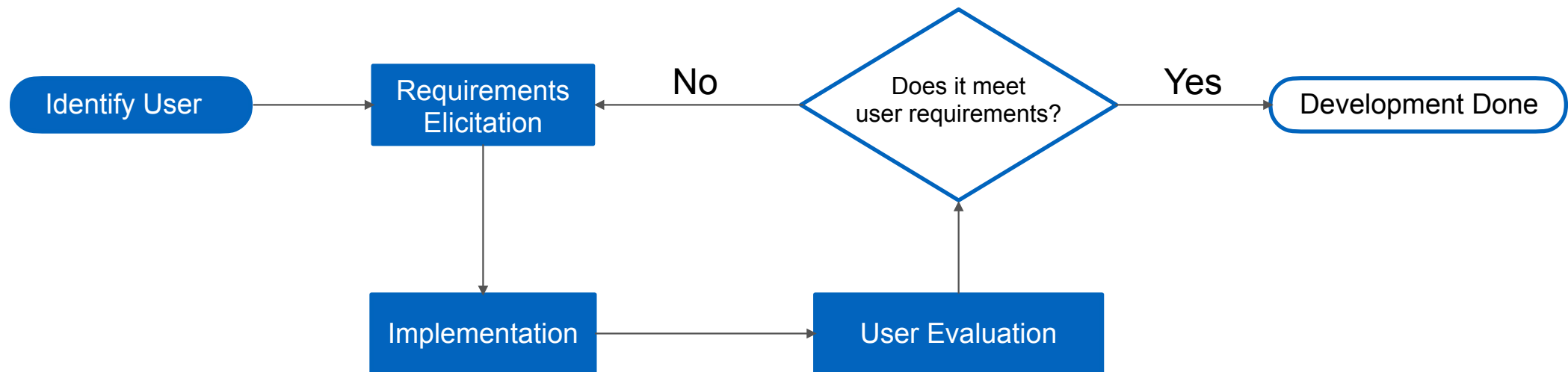
We started with a low fidelity prototype before implementing.

The system design will be elaborated further in RQ3.



RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design



The user evaluation will be elaborated in RQ4

RQ2 How to present information in a web app for user-friendly NLP entity search and exploration?

Iterative User-Centric Design

Result NLP-specific solution improves researchers search and exploration process with features more relevant to NLP researchers.

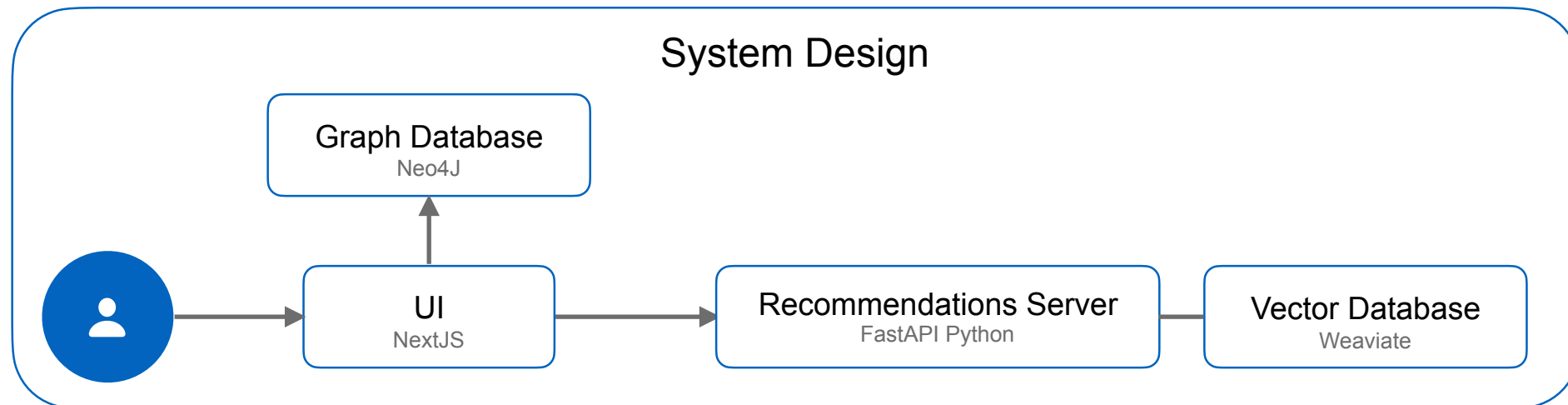
- Semantic understanding enables contextual connection for search and exploration.
- Trend analysis provides valuable insights into NLP.
- Research field hierarchy visualizes the structure and the state of NLP research.

RQ3 How to develop a performant and reliable system for semantic search and exploration?

Iterative Development

Steps Technical Literature Review, Implementation, and Evaluation

Findings We face a tradeoff between application speed, development speed, and recommendation accuracy.



RQ4 Web app evaluation

Usability testing

Plans We design interview questions based on success metrics using Likert Scale



Feature

Accurate and relevancy



Design

Intuitiveness and simplicity

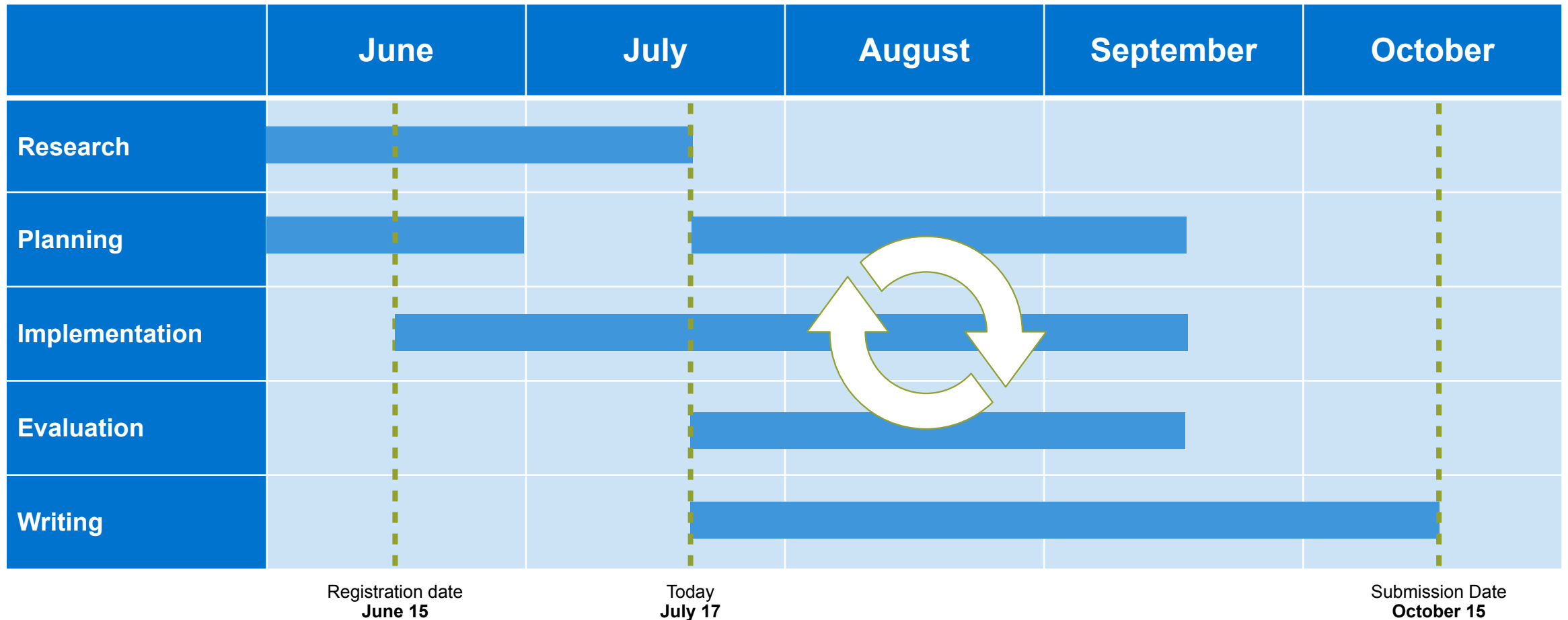


User Satisfaction

Experience and usefulness

Next steps

Timeline





Ferdy Dermawan Hadiwijaya

ferdydh.hadiwijaya@tum.de

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business
Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München



Appendix



About the User

Wesley is a munich-based PhD candidate specializing in NLP. He focuses on Large Language Model and their applications. Wesley is passionate about technology and likes to spend his free time playing with gadgets.

Problems

- Finding relevant and significant papers.
- Getting a comprehensive overview of new fields of study.
- Looking for a more modern search engine for publications.

Name	Wesley
Age	28 years old
Occupation	PhD Candidate
Years of Experience	3
Personality	Tech Savvy, Enthusiastic

Challenges

- Not enough time to sort and read all the papers.
- Extracting the knowledge and managing his findings take up a lot of his time.

Goals and Needs

- To contribute to top NLP journals to help develop innovative approaches.
- To filter publications more effectively and recognize which papers are most relevant to be read



About the User

Emily is a professor specializing in NLP. She leads a research group that studies Sentiment Analysis. Emily likes to work despite her tight schedule.

Problems

- Not enough time to sort and select papers to read
- Determining the significance and relevance of papers before investing substantial time and effort.

Name	Emily
Age	45 years old
Occupation	Professor
Years of Experience	20
Personality	Detail-Oriented

Challenges

- It is hard for her to prioritize the significant publications.
- She often finds it hard to have an overview of a paper and understand what research subfield it focuses on before reading it thoroughly.

Goals and Needs

- Exploring recent publications more efficiently.
- Seeing how a new publication is related, relevant, and significant for her research interests and expertise
- Having a holistic understanding of a paper's content without having to invest time in reading.