

Evaluating the Usefulness and Effectiveness of a Visualization for a Field of Study Hierarchy Graph to Explore NLP Research

Patrick Kufner

05.01.2024, Master's Thesis Kick-Off

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. Target User – Persona
2. Motivation
3. Approach
4. Literature Findings
5. User Study
6. Evaluation
7. Research Questions
8. Timeline

1. Target User – Persona

Name: Anna Mayer
Age: 23
Work: Informatics Student
University: TUM



Setting

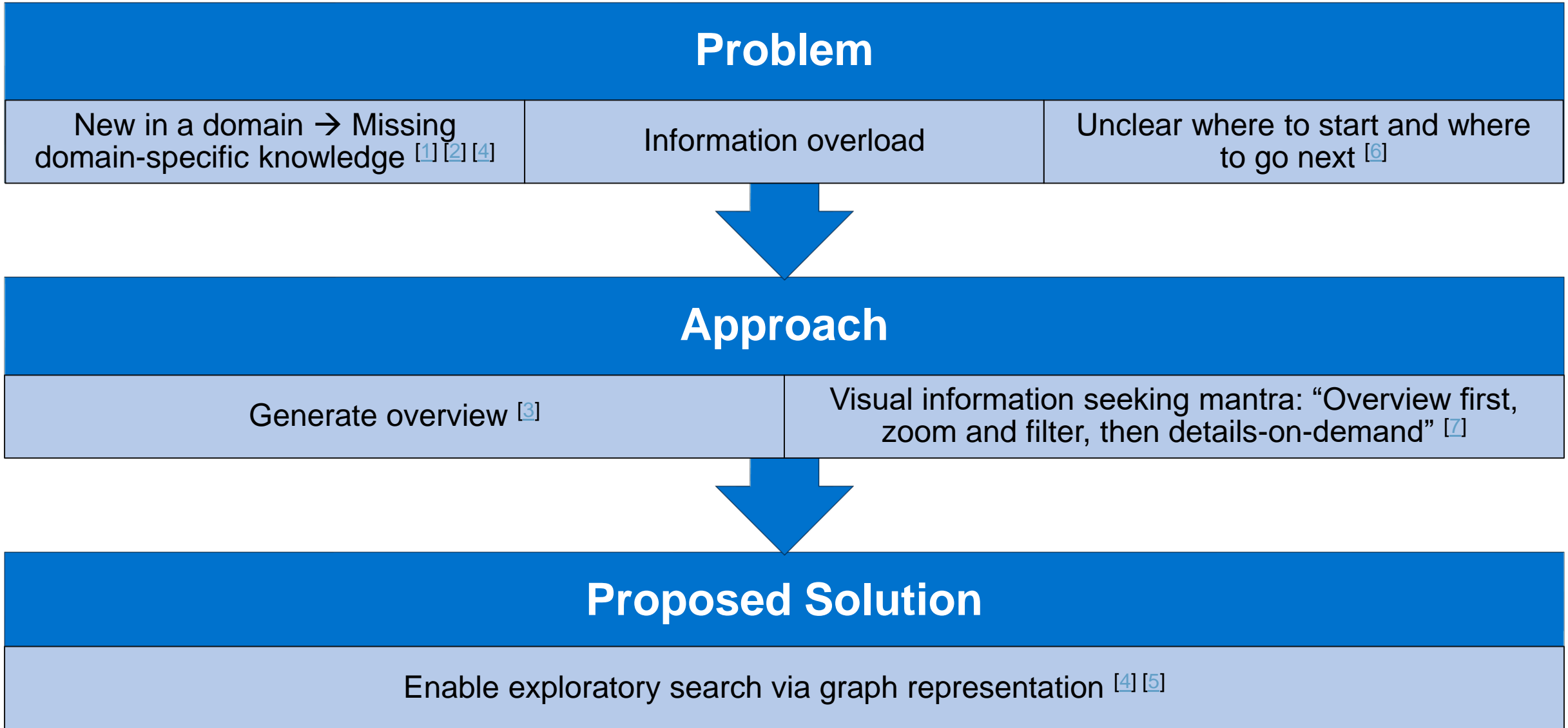
Anna has followed the curriculum of TUMs Bachelor of Science Informatics degree. She has finished all the subjects and only needs to write her Bachelor's Thesis. She found a topic within the NLP domain and must start her thesis with a literature review.

Goals

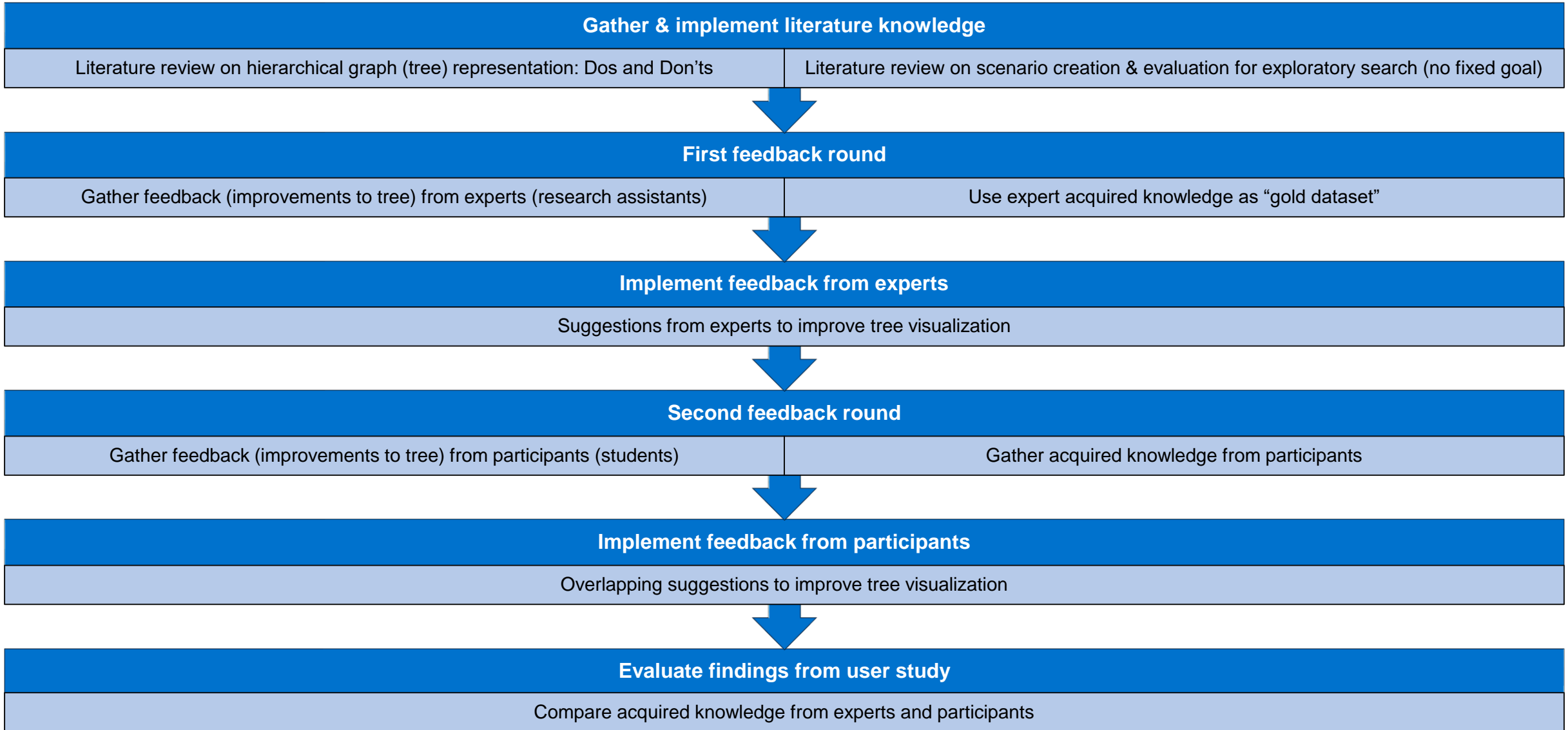
- Write Bachelor's Thesis
- Achieve Bachelor's Degree
- Gain Knowledge in NLP Domain

Frustrations

- Limited NLP specific knowledge – doesn't know keywords yet
- Lacking research skills – haven't had to scientifically research yet

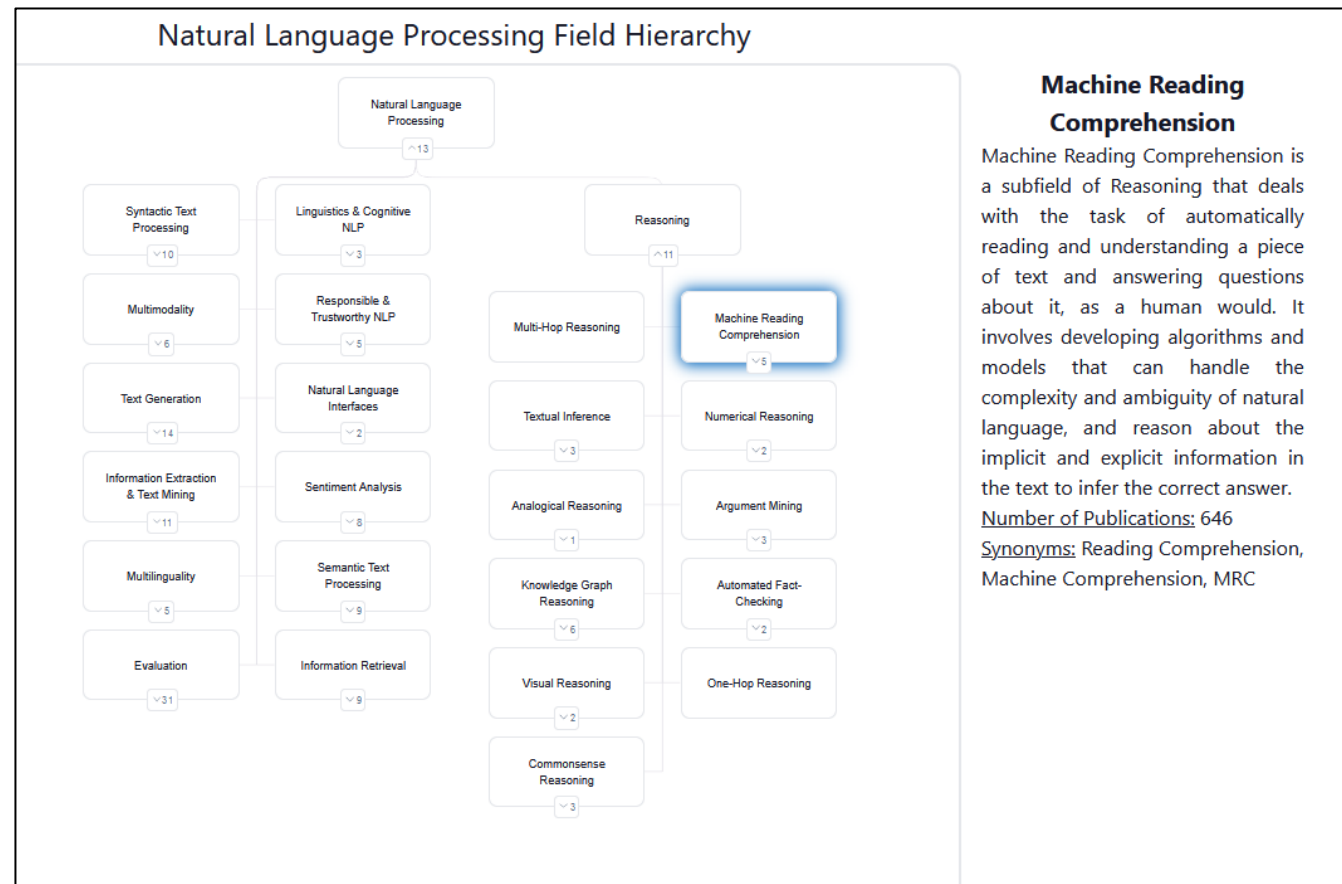


3. Approach



4. Literature Findings – Graph visualization

- Have basic information on hover [10]
- Create structure → Hierarchical format [3]
- Limited size of graph → Papers not included in the graph [1] [13]



4. Literature Findings – User Study

- Distinguish between “borderline” and “core” exploration [9]
- Core exploratory search lack defined goals
 - Questionnaire evaluation is not possible
 - Information retrieval metrics (precision, recall, F1) not applicable

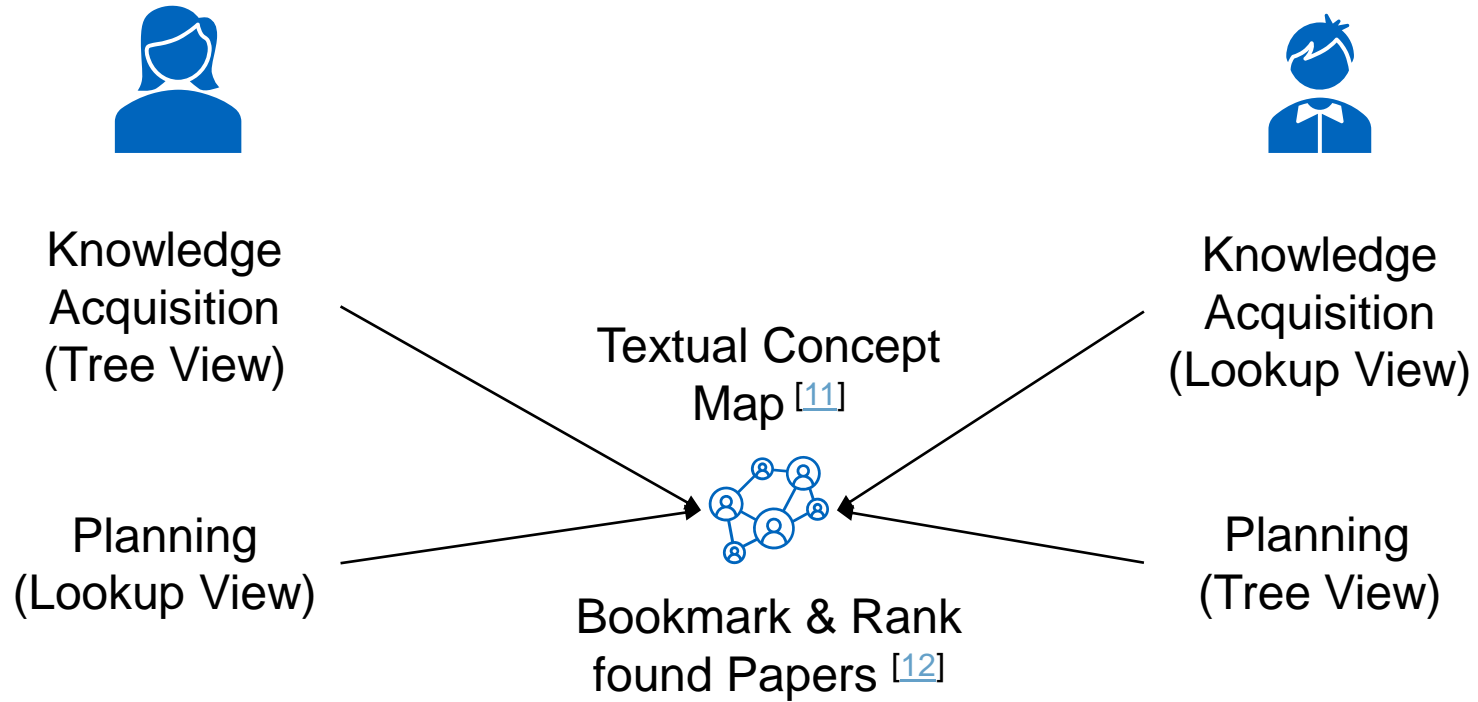
TABLE 3. Our categorization of exploratory and lookup tasks according to their primary facets. [9]

	Low complexity	High complexity
Precise goals	Core lookup	Borderline lookup
Open-ended goals	Borderline exploratory	Core exploratory

TABLE 4. Tasks used in the study. We created two tasks per type, that is 12 in total, and randomly assigned one task per type to every participant; every participant attempted six tasks in total (within subject design). The topic of the second task in the same type is given in brackets. [9]

Task (Abbrev.)	Tasks
Knowledge acquisition (Know)	You are going to start a new research project on the topic <i>Reinforcement learning</i> (or <i>Active learning</i>). You would like to learn as much information as possible about this topic, e.g. applications, problems, specific algorithms
Planning (Plan)	You are planning to give a talk on the topic <i>Deep neural networks</i> (or <i>Clustering techniques</i>). Plan the structure of your presentation, including short titles of the headings of your slides and using bullet points describe the content
Comparison (Comp)	Collect literature to write a short essay describing similarities and differences between <i>Supervised learning and Unsupervised learning</i> (or <i>Transfer learning and Multitask learning</i>).
Fact-finding (Fact)	Define the term SVM (or UCB) as in the first article that proposed it.
Navigation (Navi)	Navigate to the article that presents the most commonly used topic model—latent Dirichlet allocation—for the first time (or Navigate to the article that solves the—multi-armed bandit—problem for the first time.)
Question answering (Question)	What are the most common sampling methods used in machine learning? (List three) (or What are the kernels used in machine learning? (List three))

5. User Study – Overview



5. User Study – Scenarios

1. Knowledge acquisition

- “You are going to start a new research project on the topic [*Righteous (correct, proper, decent, morally right) usage of NLP*]. You would like to learn as much information as possible about this topic, e.g. applications, problems, specific algorithms.” [\[9\]](#)

2. Planning

- “You are going to give a talk on the topic [*Voice & Sound in NLP*]. Plan the structure of your presentation, including short titles of the headings of your slides and using bullet points to describe the content.” [\[9\]](#)

5. User Study – Concept Map

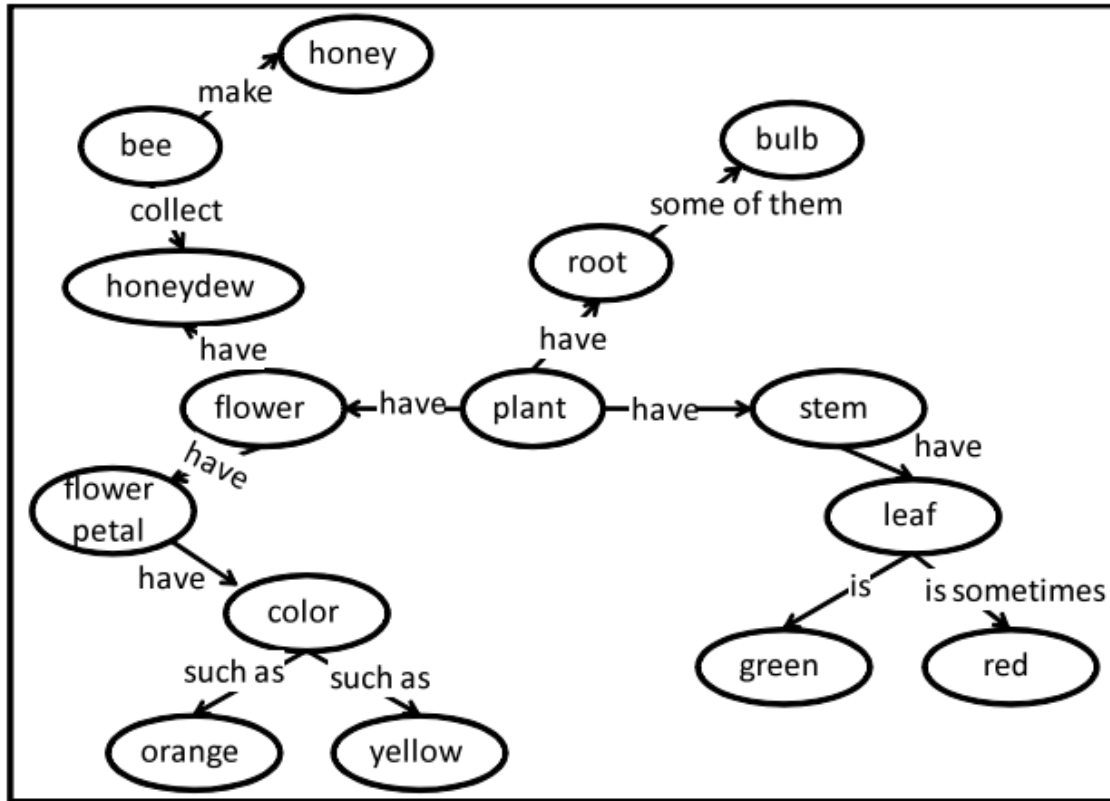


Figure 1: Example concept map about plants [\[11\]](#)

Style changes

- Text field instead of canvas for answers
- Headings and bullet points instead of nodes and edges

Reasons

- Whole evaluation done via Google Forms
 - No graph solution for Google Forms
- Graph solutions are more time consuming
- No training required for new interface

6. Evaluation – Usability & Likeability

- SUS questionnaire for tree view [8]
- Personal Evaluation for lookup & tree view
- Open Feedback for tree view

Abschnitt 4 von 7

System Usability

I think that I would like to use this system frequently *

I found the system unnecessarily complex *

I thought the system was easy to use *

I think that I would need the support of a technical person to be able to use this system *

I found the various functions in this system were well integrated *

I thought there was too much inconsistency in this system *

I would imagine that most people would learn to use this system very quickly *

I found the system very cumbersome to use *

I needed to learn a lot of things before I could get going with this system *

Abschnitt 5 von 7

Personal Evaluation Tree View

Please evaluate the tree view answering the following questions

I had a great experience with the search tool *

I was able to make new discoveries *

The tool helped me to make unexpected discoveries *

I found interesting information *

The user interface was easy to navigate *

The user interface was visually appealing *

I am satisfied with the information I found *

Abschnitt 7 von 7

Open Feedback (Tree view)

List the most **negative** aspect(s) *

List the most **positive** aspect(s) *

Are there features you missed on your search?

What could be improved?

Is there anything else you would like to share?

Content map

- For each Scenario and View...
 - Summarize experts' gathered knowledge to "gold dataset"
 - Compare participants' gathered knowledge to "gold dataset"
 - Report on similar findings
 - Report on extra findings
- Which view enables participants to gather more knowledge?
- Which view enables participants to gather qualitative better (according to "gold dataset") knowledge?

Bookmarked papers

- For each Scenario and View...
 - Compare # of experts' bookmarks to participants bookmarks
 - Compare # of bookmarks from lookup to tree view
 - Compare rating of experts' findings to participants findings
 - Compare Rating between lookup and tree view
- Bookmark more papers via lookup or tree view?
- Relevance (rating) of bookmarked papers higher in lookup or tree view?

7. Research Questions



What are the contemporary trends and advancements in graph-based exploratory search approaches?



How can the graph view be optimized to suit the distinctive characteristics of NLP research papers?

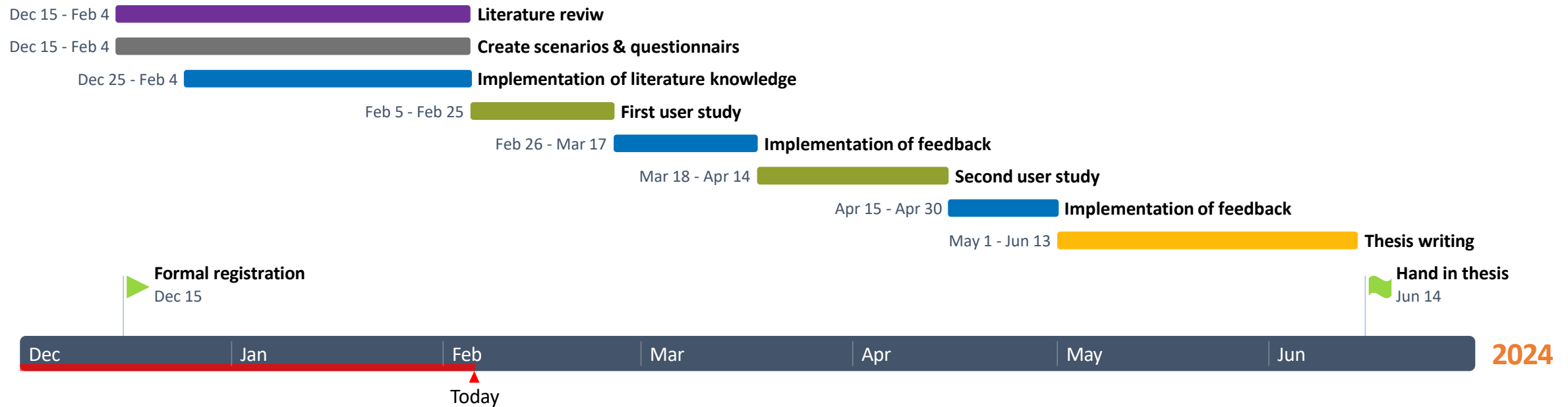


What methods are effective for evaluating exploratory search through a hierarchical graph visualization approach?



What insights can be gleaned from a user study comparing the effectiveness and usefulness of graph view and lookup search in NLP paper discovery?

8. Timeline



References

- [1] Lehmann, Simon; Schwanecke, Ulrich; Dörner, Ralf (2010)
- [2] Medlar, Alan; Ilves, Kalle; Wang, Ping; Buntine, Wray; Glowacka, Dorota (2016)
- [3] Nazemi, Kawa; Breyer, Matthias; Forster, Jeanette; Burkhardt, Dirk; Kuijper, Arjan (2011)
- [4] Nedumov, Y. R.; Kuznetsov, S. D. (2019)
- [5] Palagi, Emilie; Gandon, Fabien; Giboin, Alain; Troncy, Raphaël (2017)
- [6] Sarrafzadeh, Bahareh; Vechtomova, Olga; Jokic, Vlado (2014)
- [7] Ben Shneiderman (1996)
- [8] Laugwitz, Bettina; Held, Theo; Schrepp, Martin (2008)
- [9] Athukorala, Kumaripaba; Głowacka, Dorota; Jacucci, Giulio; Oulasvirta, Antti; Vreeken, Jilles (2016)
- [10] Christian Hirsch; John Hosking; John Grundy (2010)
- [11] Egusa, Yuka; Saito, Hitomi; Takaku, Masao; Terai, Hitoshi; Miwa, Makiko; Kando, Noriko (2010)
- [12] Athukorala, Kumaripaba; Medlar, Alan; Oulasvirta, Antti; Jacucci, Giulio; Glowacka, Dorota (2016)
- [13] Schraefel, M.; Karger, David R. (2006)



B.Sc.

Patrick Kufner

Master's Student Information Systems

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

+49.89.289.17132
matthes@in.tum.de
www.matthes.in.tum.de



Backup: System Usability Scale

Abschnitt 4 von 7

System Usability ⌵ ⋮

I think that I would like to use this system frequently *

I found the system unnecessarily complex *

I thought the system was easy to use *

I think that I would need the support of a technical person to be able to use this system *

I found the various functions in this system were well integrated *

I thought there was too much inconsistency in this system *

I would imagine that most people would learn to use this system very quickly *

I found the system very cumbersome to use *

I needed to learn a lot of things before I could get going with this system *

Backup: Personal Evaluation Tree & Lookup View

Abschnitt 5 von 7

Personal Evaluation Tree View ⌵ ⋮

Please evaluate the tree view answering the following questions

I had a great experience with the search tool *

I was able to make new discoveries *

The tool helped me to make unexpected discoveries *

I found interesting information *

The user interface was easy to navigate *

The user interface was visually appealing *

I am satisfied with the information I found *

Abschnitt 7 von 7

Open Feedback (Tree view) ⌵ ⋮

List the most **negative** aspect(s) *

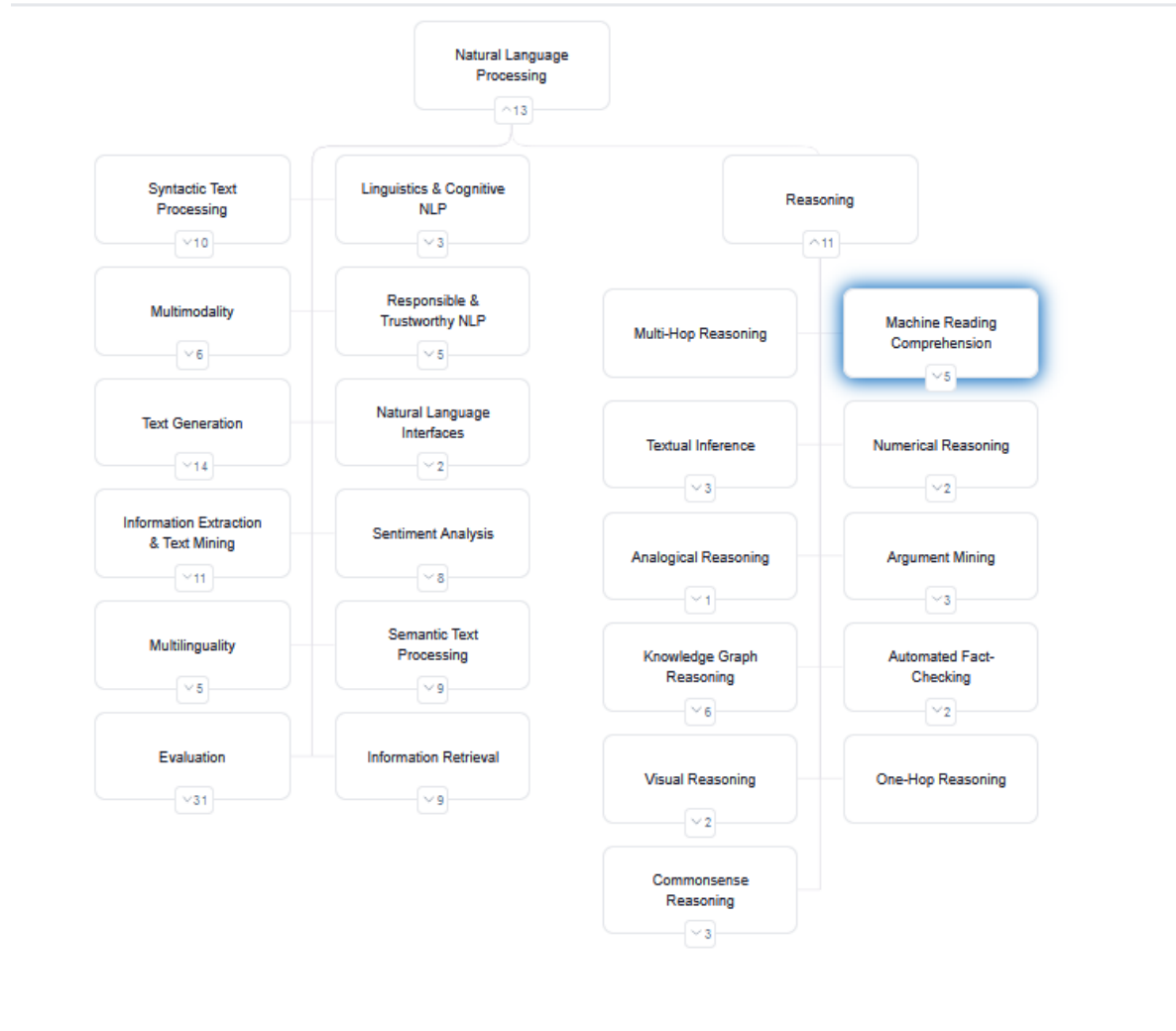
List the most **positive** aspect(s) *

Are there features you missed on your search?

What could be improved?

Is there anything else you would like to share?

Natural Language Processing Field Hierarchy



Machine Reading Comprehension

Machine Reading Comprehension is a subfield of Reasoning that deals with the task of automatically reading and understanding a piece of text and answering questions about it, as a human would. It involves developing algorithms and models that can handle the complexity and ambiguity of natural language, and reason about the implicit and explicit information in the text to infer the correct answer.

Number of Publications: 646

Synonyms: Reading Comprehension, Machine Comprehension, MRC