Evaluating Stability of Information Needs

Christin Katharina Kreutz[1], Philipp Schaer[1], Ralf Schenkel[2]

[1] TH Köln - University of Applied Sciences, Cologne, Germany
[2] Trier University, Trier, Germany

24^{th} March 24

Technology
Arts Sciences
TH Köln

CIR UNIVERSITÄT TRIER

- In general, what makes a paper relevant for you?
- In an ideal world, what would you want to base your relevancy decision regarding a paper on?
- When using dblp/Semantic Scholar/... what do you actually look at?

- In general, what makes a paper relevant for you?
- In an ideal world, what would you want to base your relevancy decision regarding a paper on?
- When using dblp/Semantic Scholar/... what do you actually look at?

... Was your answer always the same?

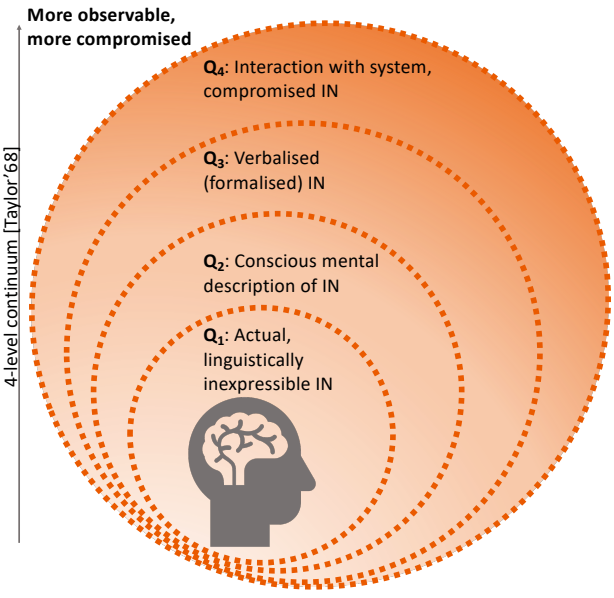
- In general, what makes a paper relevant for you?
- In an ideal world, what would you want to base your relevancy decision regarding a paper on?
- When using dblp/Semantic Scholar/... what do you actually look at?

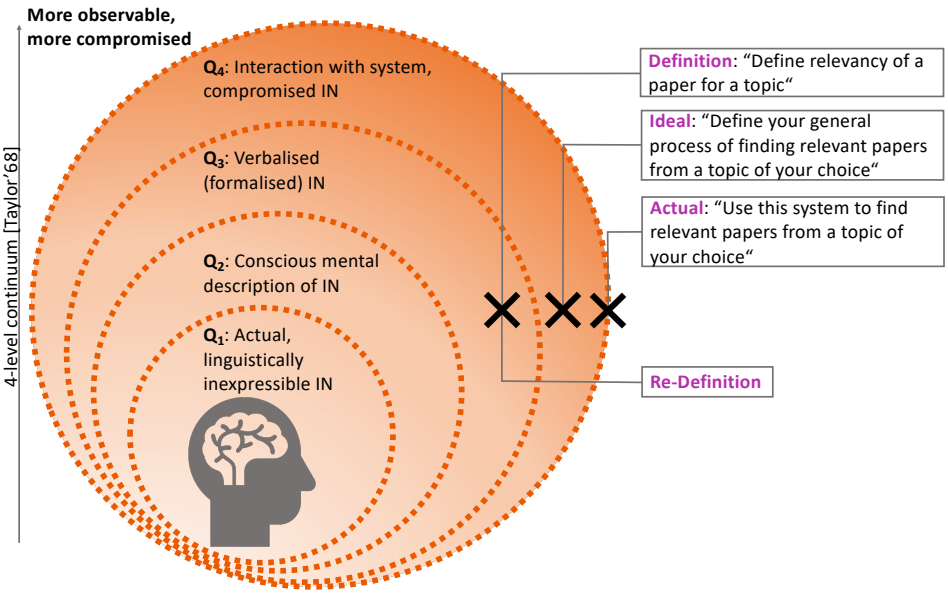
... Was your answer always the same?

If you were asked some months later, would your answer change?

Objective

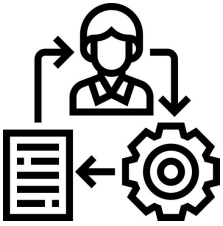
How can we observe stability of motives of users' information needs in different expressions?


More observable,
more compromised
4-level continuum [Taylor'68]
Q_4: Interaction with system, compromised IN
Q_3: Verbalised (formalised) IN
Q_2: Conscious mental description of IN
Q_1: Actual, linguistically inexpressible IN


More observable,
more compromised
4-level continuum [Taylor'68]
Q_4: Interaction with system, compromised IN
Q_3: Verbalised (formalised) IN
Q_2: Conscious mental description of IN
Q_1: Actual, linguistically inexpressible IN
Definition: "Define relevancy of a paper for a topic"
Ideal: "Define your general process of finding relevant papers from a topic of your choice"
Actual: "Use this system to find relevant papers from a topic of your choice"
Re-Definition

- Motivation
- Concept
- Datasets: FIND and Re-FIND
  - Study
  - Manifestations
  - Categorisation
- Analysis
- Conclusion

- Reuse of an existing dataset (FIND, on Zenodo) we composed for an earlier paper
- Extension of dataset with additional data (Re-FIND, on Zenodo) with the same participants some months later
- Analysis of data under stability viewpoint



Tasks

- **Expert search**: Find two experts on a topic of your liking.
- **Paper search**: Find relevant papers from a topic of your liking which appeared after 2017.

Participants

- 13 computer/information scientists, differing expertise in using DLs for research tasks:
  - 2 Master's students
  - 6 PhD students (first to last year)
  - 1 industry researcher
  - 1 dblp staff member
  - 1 postdoc
  - 2 professors
- Code names for anonymity, e.g. *green_deer*

Definition (from FIND)

- Person's definition of what the desired result would look like
- Independent of finding result
- **Used:** Transcription of definitions

Definition (from FIND)

- Person's definition of what the desired result would look like
- Independent of finding result
- **Used:** Transcription of definitions
- *How would a person define the result satisfying an information need?*

Ideal Strategy (from FIND)

- Person's description of them ideally solving a task
- Independent of DL, fuzzy
- **Used:** BPMNs of ideal task solutions

Ideal Strategy (from FIND)

- Person's description of them ideally solving a task
- Independent of DL, fuzzy
- **Used:** BPMNs of ideal task solutions
- *How would a person ideally solve a task if they were free to do it any way they wanted?*

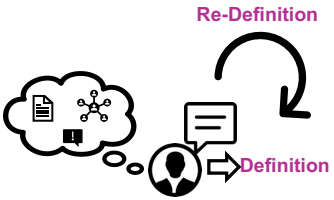
Actual Strategy (from FIND)

- Strategy shown by person actually using one DL (SchenQL)
- Mostly limited by options provided by DL, descriptions what would be searched for
- **Used:** BPMNs of actual task solutions

Actual Strategy (from FIND)

- Strategy shown by person actually using one DL (SchenQL)
- Mostly limited by options provided by DL, descriptions what would be searched for
- **Used:** BPMNs of actual task solutions
- *How does a person's strategy actually look like using one specific system?*

Re-Definition (from Re-FIND)



- Person's definition of what the desired result would look like some months later
- Independent of finding result
- **Method:** Audio-recording, transcription → Transcripts of re-definitions
- *How would a person define the result satisfying an information need some months later?*

Factor Extraction

- Consider 4 manifestations: 3 existing, 1 new
- Expert in DLs extracted factors
- Semi-normalised factors
  E.g., *'resulting paper newer than put in paper/keywords'* and *'check recency of paper'* (of papers fitting keywords on topic)
  $\rightarrow$ *'published recently on topic'*

Categorisation

- 4 meta-categories (SOUP)
- Self-determined (S) - 8 sub-categories for expertise, 3 for relevancy of papers
- Other-directed (O) - 2 sub-categories for expertise, 3 for relevancy of papers
- Under-specified (U)
- Personal (P)

Categorisation

- 4 meta-categories (`SOUP`)
- `Self-determined (S)` - 8 sub-categories for **expertise**, 3 for relevancy of papers
  → academic, author-dependent, collaboration, knowledge, paper-dependent, productivity, quality, venue
- `Other-directed (O)` - 2 sub-categories for **expertise**, 3 for relevancy of papers
  → citation, external
- `Under-specified (U)`
  → under-specified
- `Personal (P)`
  → searcher's context

Categorisation

- 4 meta-categories (`SOUP`)
- `Self-determined (S)` - 8 sub-categories for expertise, 3 for **relevancy of papers**
  → author-dependent, paper-dependent, venue
- `Other-directed (O)` - 2 sub-categories for expertise, 3 for **relevancy of papers**
  → citation, expert-dependent, external
- `Under-specified (U)`
  → under-specified
- `Personal (P)`
  → searcher's context

Categorisation

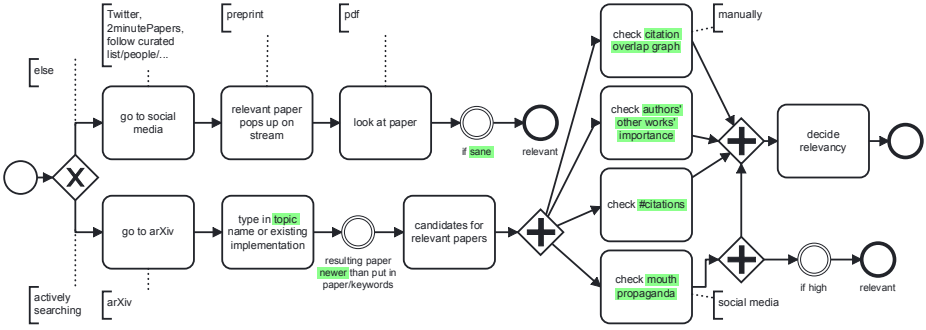
- 4 meta-categories (`SOUP`)
- `Self-determined (S)` - 8 sub-categories for expertise, 3 for relevancy of papers
- `Other-directed (O)` - 2 sub-categories for expertise, 3 for relevancy of papers
- `Under-specified (U)`
- `Personal (P)`

After this step: For all participants indication which factors from which categories are relevant in which manifestations

investigator: Okay. And how would you define relevancy?
green_deer: It's actually pretty hard to, so it's a visual topic again. So, relevancy here is, does it improve the outcome? And as a human, you can basically just look at the outcome of the algorithm and decide if this is relevant or relevant improvement or not. So, that's. Yeah. I think that's how this whole visual community is driven.
So, yeah, I'd say it's basically looking at it.

Factor:

- Does paper improve the outcome (S - paper-dependent)



Factors:

- Sanely written (S - paper-dependent)
- Fitting topic (S - paper-dependent)
- Recency (S - paper-dependent)
- Overlapping references with other relevant papers (O - citation)
- Importance of other works of authors (S - author-dependent)
- Number of citations (O - citation)
- Mouth propaganda (U - under-specified)

- Motivation
- Concept
- Datasets: FIND and Re-FIND
  - Study
  - Manifestations
  - Categorisation
- Analysis
- Conclusion

Research Questions

How can we observe stability of motives of users' information needs in different expressions?

RQ_1 What factors do users of DLs define expertise and relevancy of papers with?

RQ_2 How do individual users (intend to) apply their general definitions?

RQ_3 How stable are individual users' general definitions over time?

RQ_4 How stable are individual users' manifestations of information needs?

RQ_1: What factors do users of DLs define expertise and relevancy of papers with?

- Observe definitions
- 3.08 (expert) and 2.69 (relevancy of papers) factors on average
- Factors are very diverse and highly individual
- Mostly factors from `self-determined` meta-category

RQ_2: How do individual users (intend to) apply their general definitions?

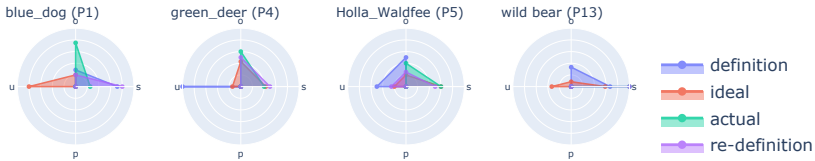
- Observe definitions – ideal, definitions – actual
- Many overlaps between definitions + ideal, but ideal processes more detailed
- *„I strongly idealized my search behaviour. (...) I had the impression that my real search behaviour is much simpler.“*
- Many overlaps in definition + actual, especially for relevancy of papers
- Under-specified factors disregarded in actual

RQ_3: How stable are individual users' general definitions over time?

- Observe definitions – re-definitions
- 4.3 (expert) and 5 (relevancy of papers) factors on average
- Considerable similarities in categories (7, 11) and individual factors (5, 4)

RQ_4: How stable are individual users' manifestations of information needs?

- Observe definitions – ideal – actual – re-definitions
- Different usage patterns of participants, e.g., ignoring searchers' context (P, see below) or clearly describing relevant factors



- Use or disregard complete categories
- Considerable stability of (meta-) categories

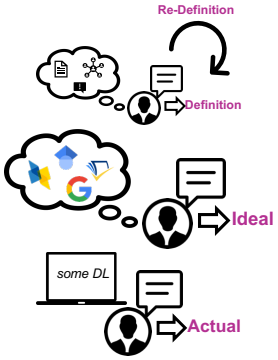
- Motivation
- Concept
- Datasets: FIND and Re-FIND
  - Study
  - Manifestations
  - Categorisation
- Analysis
- Conclusion

Recap:

- Capturing information need in different manifestations
- General stability of meta-categories
- Satisfying information needs required multiple factors

Future Work

- Observe importance of factors
- User model → user simulation



Thank you for your kind attention!

@kreutzch

✉ christin.kreutz@th-koeln.de