Information Retrieval

Exercise – Winter term 2025/2026

klara.gutekunst@uni-kassel.de

Agenda

1. Research Questions
2. Hypothesis Testing
3. Assignment
4. Inspiration

Research Questions

What is a good research question?

Research Questions

- A good research question. . . [Bartos 1992]
  - . . . asks about the relationship between two or more variables.
  - . . . is testable (i.e., it is possible to collect data to answer the question).
  - . . . is stated clearly and in the form of a question.
  - . . . does not pose an ethical or moral problem for implementation.
  - . . . is specific and restricted in scope.
  - . . . identifies exactly what is to be solved.

- Examples:
  - *Poor:*
    “What is the effectiveness of parent education when given problem children?”
  - *Good:*
    “What is the effect of the STEP parenting program on the ability of parents to use natural, logical consequences (as opposed to punishment) with their child who has been diagnosed with bipolar disorder?”

Hypothesis Testing

What is a good hypothesis?
How to test a hypothesis?

Hypothesis Testing

- A good hypothesis...
  - ...is founded in a problem statement and supported by research.
  - ...is testable.
  - ...states an expected relationship between variables.
  - ...is stated as simply and concisely as possible.

- Hypothesis testing:
  - Step 1: What are your variables? (nominal, ordinal, scale, ratio)
  - Step 2: Measure the variables (Are aggregated measures enough?)
  - Step 3: Significance test (Null hypothesis? Which α level? Which significance test?) [lecture video 2024]

Assignment

- Come up with a good research question
  - Not too complex
  - Focus on effectiveness, not efficiency
- Formulate a hypothesis for your research question
- Test your hypothesis
  - The previously annotated topics are used for testing (via TIRA)
  - Your experiment can use final effectiveness measures of our 10 baseline systems and the others' approaches (i.e., the full leaderboard)
  - Test measurements are only provided *after* formulating your hypothesis
- Shortly analyze your findings in a written report
- Exercise sheet on temir.org
- Due Date: Monday, 19.01.2025, 23:59
- Deliverable: TIRA submission(s) and short report (1.5–2 pages) written in LaTeX

Inspiration

- System effectiveness from last semesters [summer ’23][winter ’23/’24][summer ’24]
  - Which systems performed well?
  - Which topics were difficult?
  - Where were “good” retrieval systems fooled?

- TIREx components overview [link]

Appendix: Variables

Scale (Operation)	Categories (no order or direction)	Natural Order	Equal Intervals	True Zero	Example
Nominal (=)					Marital status, sex, gender
Ordinal (median)					Student grade
Interval ($a+b$,$a-b$, $\frac{a+b}{2}$)					Temperature, year
Ratio ($a \cdot b$, $\frac{a}{b}$, $\sqrt{a \cdot b}$)					Age, height, weight