Large Language Models (LLMs) Workshop

Introduction and Hands-On Examples

Workshop procedure
Too much too little time > 200 slides

- ~ first 1 – 1,5 hours: theory

- rest "3. Demonstrations"

- Relatively little interaction, more in the form of a monologue

I'm late, I'm
Late, for a very
important
date. No
time to
say hello,
goodbye,

I'm late,
I'm late,
I'm late!

Lewis Carroll
Alice in Wonderland

Workshop procedure
Too much too little time > 200 slides

- the slides can be used for reference and self-study

Two long-term ideas

- Set up working groups regarding X if desired?

> **Project Idea**
>
> Create an speech processing pipeline that mimics Google's Speech-to-Text API by (generating audio from text), transcribing the audio back to text, and refining the transcription using one LLM for enhanced accuracy and one for summarizing.

> **Project Idea**
>
> Develop an integrated system that combines bibliometric analysis with LLM-driven methods for efficiently summarizing academic literature and identifying trends within research.

- Conduct another "hands on" workshop?

Record my screen

Open Broadcaster Software (OBS) Overview:
• OBS Studio is a free, open-source software used for video recording and live streaming.

Screen Recording:
• I will only record the screen from my shared window and record my audio.

Intended Use:
• **For Internal Use Only**: Screen recordings created with OBS are solely for internal documentation.

Goal:
⇒A student will apply LLMs locally to
transcribe the audio file to text
to summarize the text

To write a documentation for students in the future about LLMs, something like https://www.uni-mannheim.de/media/Einrichtungen/Koordinationsstelle_Studieninformationen/Dokumente/Erstsemester/ChatGPT_Handreichung_Studierende_UMA_Stand_Mai_2023.pdf

Does anybody NOT want my split screen to be recorded?

Who am I?

EDUCATIONAL TRAINING

- Sep 2015 – Sep 2018   Applied Statistics (M.Sc.), *University of Göttingen*
- Sep 2012 – July 2015   Psychology (B.Sc.), *Université de Fribourg (CH)*
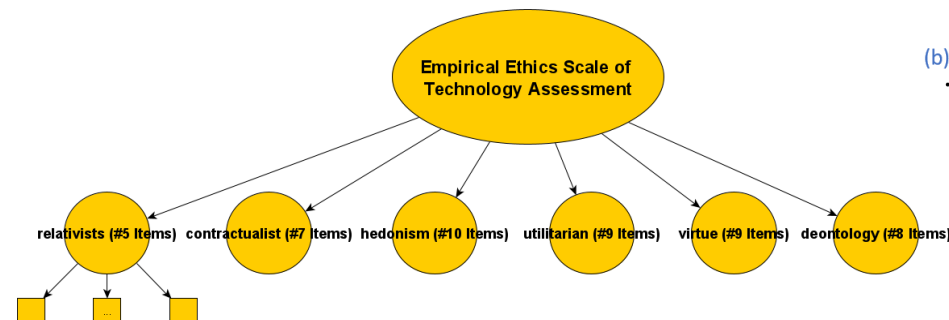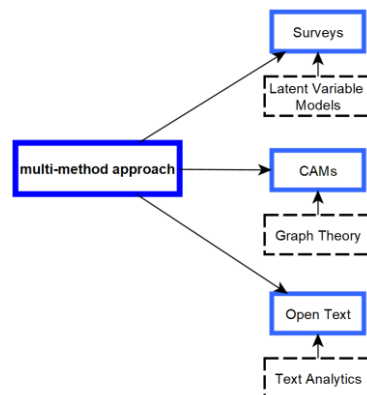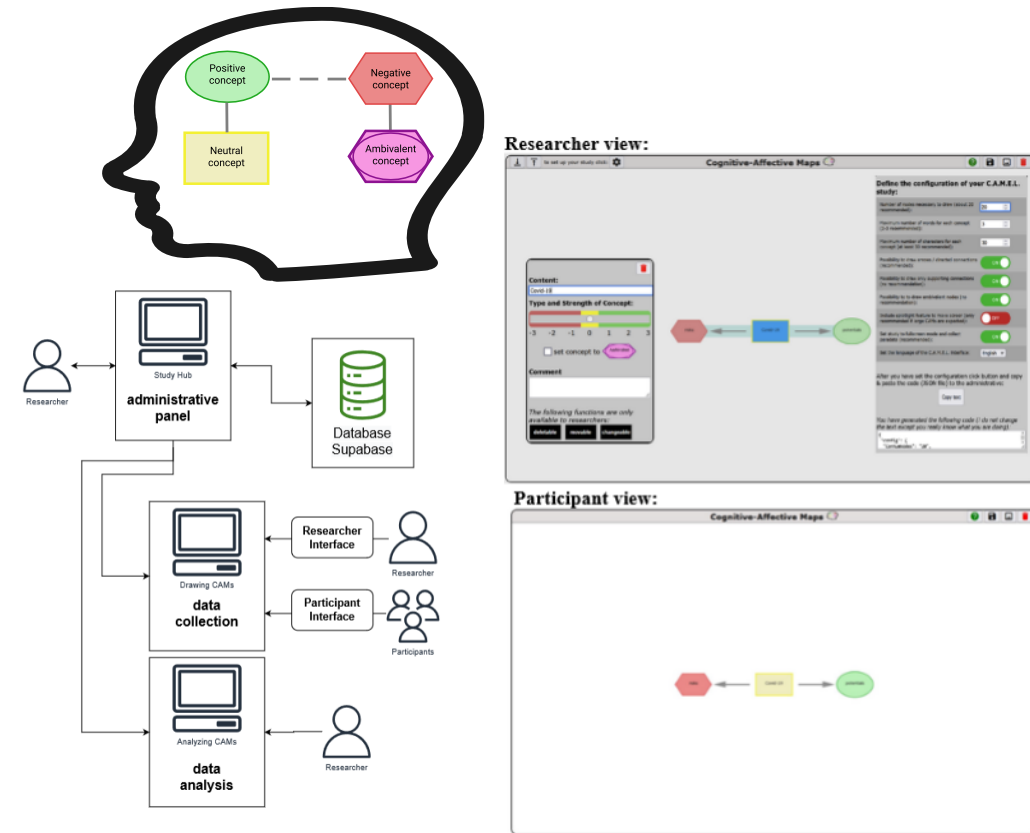
PROFESSIONAL TRAINING

- Since April 2024 PostDoc Project in the DFG Research Training Group "Statistical Modeling in Psychology", primary advisor Prof. Kiesel (University of Freiburg)

- October 2020 – March 2024 PhD Project in the DFG Research Training Group "Statistical Modeling in Psychology", primary advisor Prof. Kiesel (University of Freiburg)

- Since Jan 2020 freelancer: coach for data analysis (supervision of bachelor, master and doctoral theses), private tuition on statistics

My research <-> LLMs

Semantic Networks: Cognitive-Affective Maps
extended logic: Proposing Tools to Collect and
Analyze Attitudes and Belief Systems

Empirical Ethics: Development and Validation of
an Empirical Ethics Scale for Technology
Assessment - Challenges and Perspectives for a
Real Time Ethics for Emerging Technologies

Prospective Technology Assessment: Propose
methodologies to investigate (ethical) concerns
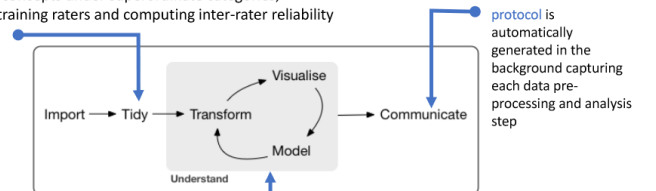of technologies at an early development stage.

Current project <-> LLMs

Evaluate the psychometric properties of the benchmark

MMLU (Massive Multitask Language Understanding

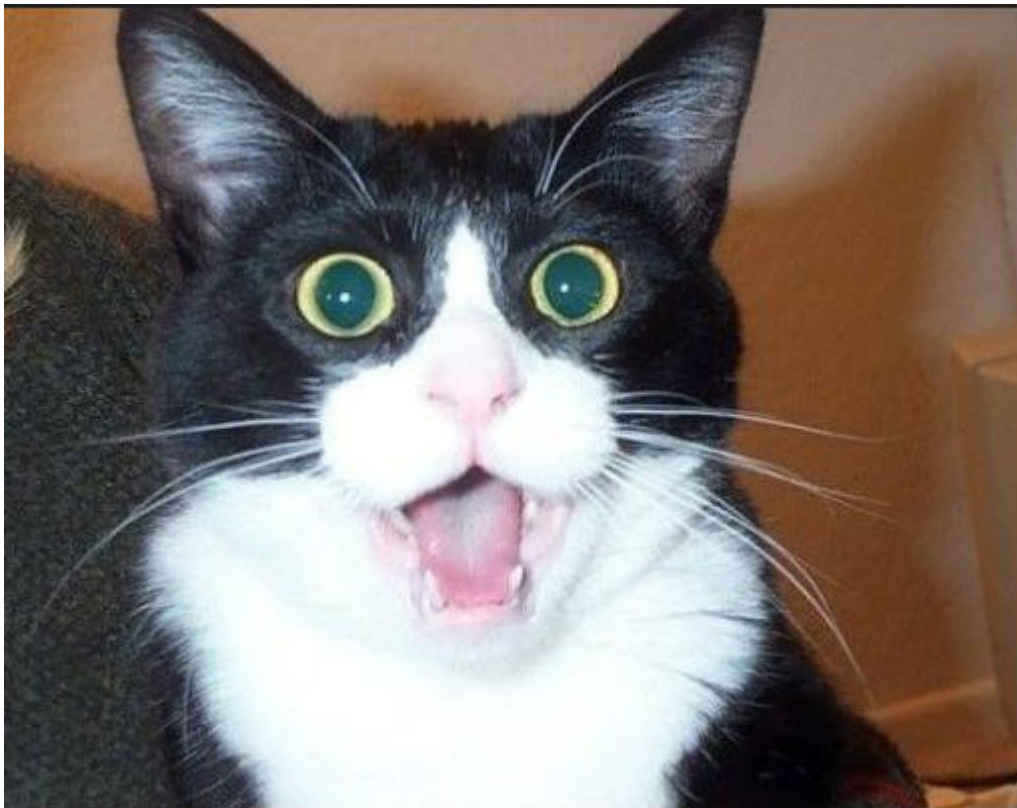➤Analyze over 12.000 items by Latent Variable Models (IRT)

Together with:

- Dr. Aron Fink (https://www.psychologie.uni-frankfurt.de/73548927/Dr__Aron_Fink)

- Tom Sühr (https://tomsuehr.com/)

Let's start

How I felt while preparing the workshop:

- Stunned in disbelief
- How to set up my computer?!?!?!?!?

LEVELS OF SATISFACTION

NEW PROJECT

FIXING A BUG

DELETING CODE

RESOLUTION: NOT A BUG

MONKEYUSER.COM

me trying to find the bug

the actual bug

working code