LaSTing: **La**nguage **S**cience & **T**echnology

Robust Assessment & Safe Applicability of Language Modeling: Foundations for a New Field of Language Science & Technology

DFG priority area 2026–2029

Michael Franke
Eberhard Karls
Universität Tübingen

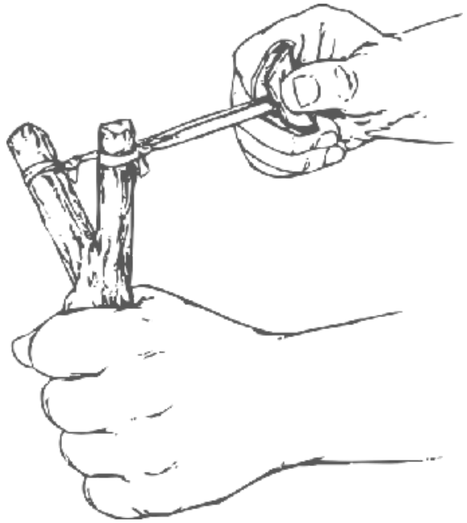
Vera Demberg
Saarland Universität
MPI Informatik

Gerhard Jäger
Eberhard Karls
Universität Tübingen

Barbara Plank
Ludwig Maximilian
Universität München

David Schlangen
Universität Potsdam

Agenda



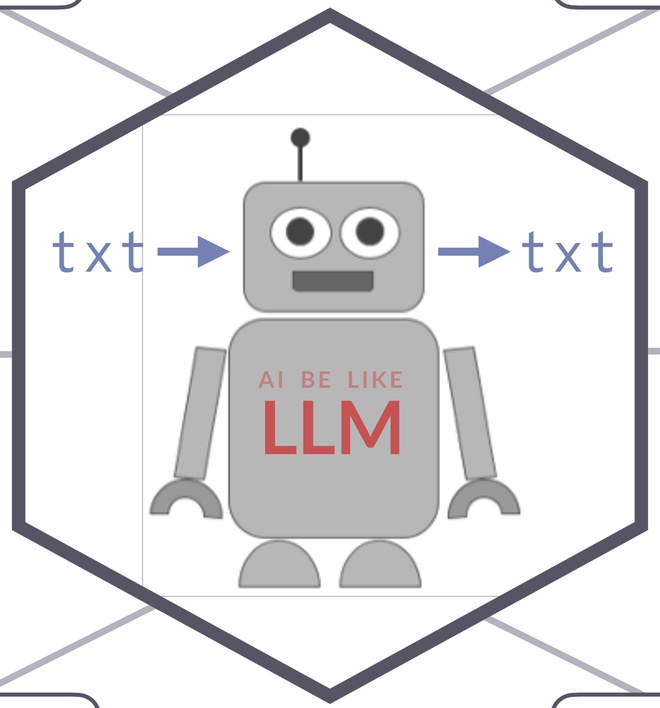
- motivation
  - why this SPP initiative?
- content
  - which projects to include?
- practicalities
  - how to apply for what?
- Q&A

vision

SOCIETY

INDUSTRY

EDUCATION



ENGINEERING

HUMANITIES

SCIENCE

Challenges & chances for the cognitive language sciences

language processing / human cognition

[T]o learn to predict text, is to learn to predict the causal processes of which the text is a shadow.

Eliezer Yudkowsky "GPTs are Predictors, not Imitators", April 8th 2023 on lesswrong.com, attributed to **Ilya Sutskever**

nature of language

[L]anguage models should be treated as bona fide linguistic *theories*.

Modern language models refute Chomsky's approach to language

Steven T. Piantadosi[a,b]
[a]UC Berkeley, Psychology [b]Helen Wills Neuroscience Institute

Motivation

Problems

- fast-paced field
- Wirtschaftlichkeit ≠ Wissenschaftlichkeit
- limits of current understanding
  - representations & mechanisms
  - what *does* an LM model actually?
- lack of methodological standards
  - e.g., how do we know what an LM knows or can do?

Motivation

Problems

- fast-paced field
- Wirtschaftlichkeit ≠ Wissenschaftlichkeit
- limits of current understanding
  - representations & mechanisms
  - what *does* an LM model actually?
- lack of methodological standards
  - e.g., how do we know what an LM knows or can do?

Goals

- LMs as tool for the cognitive language sciences
- insights into
  - human linguistic processing / general cognition
  - nature of language (in the abstract)

Motivation

Problems

- fast-paced field
- Wirtschaftlichkeit ≠ Wissenschaftlichkeit
- limits of current understanding
  - representations & mechanisms
  - what *does* an LM model actually?
- lack of methodological standards
  - e.g., how do we know what an LM knows or can do?

Goals

- LMs as tool for the cognitive language sciences
- insights into
  - human linguistic processing / general cognition
  - nature of language (in the abstract)

Path forward

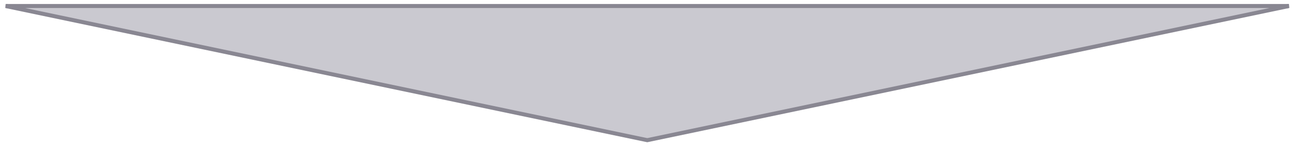
- interdisciplinary reflection on methods and foundational issues
- theoretically and empirically anchored "Philosophy of LMs"

LaSTing

Relevance

same questions arising in multiple research contexts

- theoretically-informed ‘mechanistic interpretability’
- linguistically-informed benchmarks
- evidence from synthetic data
- architectural ‘inductive biases’
- cognitive modeling with LMs
- ….



robust assessment

behavior, representations, mechanisms

safe applicability

in the cogn. language sciences | for explanation

foundational understanding

nature of models and their predictions

How to tell that your project fits this SPP?

non-exhaustive examples

✅

project contributes to a (methodological / foundational) reflection on the role of language technology* in the cognitive language sciences**

* main but non-exclusive focus on language modeling

** linguistics, CL, NLP, CogSci, psychology, neuroscience, philosophy, …

✅

project addresses a concrete research question from a specific domain, but the problem statement / results / methods (…) are relevant in and transferable to other domains

✅

we learn something about langTech using methods / insights from the cognitive language sciences OR

we learn something about human cognition / language from the models

How to tell that your project does not fit this SPP?

non-exhaustive examples

new method to improve performance for some task X (w/o deeper conceptual justification except that it works)

engineering solution to make training / inference more resource efficient (unrelated to the cognitive language sciences)

new benchmark data set with material vaguely reminiscent of the human ability usually referred to as X

abstract discussion of aspect X of LMs (abilities, societal impact, ethical issues, …) without engaging concretely with the technology

Deliverables

non-exhaustive examples

- deeper technical understanding (of langTech)
  - theoretically informed benchmarks or training sets
  - formal (limit) results
  - mechanistic interpretability integrating insights from psycho- and neurolinguistics
- novel & safe applications (of langTech)
  - for downstream practical tasks
  - as tools assisting scientific inquiry
- foundational questions
  - language models as theories of language
  - trustworthy evidence from LMs in scientific debate
- robust methods
  - LM-ology 101
  - experimental methods informed by standard best-practices from the behavioral sciences

Network, community building, structural measures

20 projects to be assigned *ex post* to 4-6 thematic areas

- annual meetings
- workshops
- short-term collaboration
- outreach program
- autumn schools
- coaching / mentoring
- equal opportunity measures
- PhD progress trajectories
- PostDoc start-up grants
- Mercator Fellows

practicalities

Who can apply?

2.1 Eligibility

Researchers in Germany, or those working at a German research institution abroad, who have completed their academic training (a doctorate as a rule) are eligible to apply.

Proposals may also be submitted by researchers working at a non-university research institution without being subject to a cooperation requirement regarding the individual proposal.

Furthermore, project proposals may be submitted by researchers based at foreign research institutions if their project offers added value to the Priority Programme as a whole. This must be explained in the proposal. In addition, participation of researchers at research institutions abroad is possible subject to the conditions described under Special Provisions (B III 1).

https://www.dfg.de/resource/blob/168092/a766b77fd205967d25e1ea711e963aae/50-05-en-data.pdf

How to apply?

- essentially: same as for any individual DFG project
  - read the **Proposal Preparation Instructions** [link]

What makes an application successful?

- thematic fit
  - interdisciplinarity
  - contribution to foundational / methodological question
  - network-ability
- modest budgeting
- **caveat:** selection is made by DFG (senate) based on external reviews
  - neither the coordinator (Franke), nor the board (Demberg, Jäger, Plank, Schlangen) have any influence on this

What / How much to apply for?

- €7m in total
  - for initial funding phase of 3 years
- intended for ~20 projects
  - at most; possibly fewer
- calculated for a total of
  - 10 PostDocs (E13 100%)
  - 7 PhDs (E13 100% | e.g., computer science)
  - 8 PhDs (E13 65% | e.g., humanities)
- ⇒most projects have one PhD or one PostDoc
- ⇒ average per-project budget: ~ € 323k
  - average per-project running costs: ~ € 40K

Additional benefits from SPP

short-term collaboration	30k
workshops	48k
equality measures	45k
day care	30k
autumn schools	54k
annual meeting	9k
start-up program	90k
Mercator fellows	60k
public relations	10k

Timeline

- deadline for project proposals: **Sep 30 2025**
- unofficial decision on accepted projects: **~ January 2026**
- official decisions on accepted projects **~ March 2026**
- planned project start: **May 1 2026**
- duration of first funding phase: 3 years

LaSTing

DEEP
CALM
OPEN

LANGUAGE
SCIENCE
on & with
Language Modeling