



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



RESEARCH ASSOCIATE FOR THE PROJECT “EVALUATING, EXPLAINING, AND ENABLING ETHICAL MULTI-AGENT SYSTEMS OF LARGE LANGUAGE MODELS (E4-MALM)” § 28 SUBSECTION 3 HMBHG

Institution: Faculty of Mathematics, Informatics and Natural Sciences, Department of Informatics, Trustworthy Artificial Intelligence

Salary level: EGR. 13 TV-L

Start date: 01.06.2026 pending approval of external funding, fixed until 31.05.2029 (This is a fixed-term contract in accordance with Section 2 of the academic fixed-term labor contract act [Wissenschaftszeitvertragsgesetz, WissZeitVG]).

Application deadline: 2026-04-08

Scope of work: full-time position suitable for part-time

Your responsibilities

Duties include academic services in the project named above. Research associates may also pursue further academic qualifications outside of their work responsibilities. They may also pursue doctoral studies outside of working duties.

The position is located at the Professorship of Trustworthy AI that focuses on the computational processing of unstructured data, mainly text. Our goal is to conduct holistic research in artificial intelligence: while most research in natural language processing and artificial intelligence focuses on obtaining state-of-the-art results on established benchmark data sets only, we seek to couple research on improving the effectiveness of our systems with the consideration of ethical aspects. That is, we aim to develop effective but also environmentally sustainable, fair, and inclusive solutions.

The employee will be hired within the project “Evaluating, Explaining, and Enabling Ethical Multi-Agent Systems of Large Language Models (E4-MALM)”, which is funded by the German Research Foundation (DFG) as part of the Priority Program LaSTing (SPP 2556) on Language Science and Technology. The aim of this project is to advance the safety and ethical reliability of multi-agent systems of large language models (MALMs). Concretely, E4-MALM pursues three core objectives: (1) developing robust evaluation frameworks and stress-testing methods to assess ethical and emergent behaviors in MALMs, (2) explaining collective failure modes (e.g., toxic agreement, groupthink) at a mechanistic level using interpretability methods, and (3) enabling ethical behavior through parameter-efficient fine-tuning and activation steering techniques guided by these mechanistic insights.

Within this project, the employee is highly encouraged to conduct a Ph.D. covering fundamental research in natural language processing and artificial intelligence with a focus on one or multiple of the following topics:

- Large Language Models

- Multi-Agent Systems of LLMs
- Mechanistic Interpretability
- Ethical Artificial Intelligence (Fairness, Sustainability, Inclusion, etc.)

The research should result in international top-tier publications (conference or journal articles) in natural language processing (e.g., ACL), artificial intelligence (e.g., AAAI), and/or interdisciplinary venues. The supervisors of the Ph.D. candidate will be Prof. Dr. Anne Lauscher and Dr. Jae Hee Lee. Both will be available for questions related to the position via email.

Our research group is dynamic, ambitious, and nationally as well as internationally very well connected. Diversity and the support of individuals identifying as members of marginalized groups are important to us.

Your profile

A university degree in a relevant field.

- master's degree (or similar) in computer science, computational linguistics, artificial intelligence or another relevant field (e.g., information systems, data science, etc.)
- experience with Python programming
- very good proficiency of the English language
- interest and joy in seeking creative solutions
- experience in machine learning and/or natural language processing is a plus
- publications or theses relevant to the topic of the project (e.g., large language models, multi-agent systems, ethical AI, etc.) are a plus

We offer



Reliable remuneration based on wage agreements



Continuing education opportunities



University pensions



Attractive location



Flexible working hours



Work-life balance opportunities



Health management, EGYM Wellpass



Educational leave



30 days of vacation per annum

Universität Hamburg—University of Excellence is one of the strongest research educational institutions in Germany. Our work in research, teaching, educational and knowledge exchange activities is fostering the next generation of responsible global citizens ready to tackle the global challenges facing us. Our guiding principle “Innovating and Cooperating for a Sustainable Future in a digital age” drives collaboration with academic and nonacademic partner institutions in the Hamburg Metropolitan Region and around the world. We would like to invite you to be part of our community to work with us in creating sustainable and digital change for a dynamic and pluralist society.

The University of Hamburg is committed to equity. Diversity enriches our university life, whether in our studies, research, teaching, education, or workplace. We therefore welcome all applications, regardless of gender, gender identity, sexual orientation, ethnic or social background, age, religion or belief, disability, or chronic illness.

The University of Hamburg strives to increase the number of women in academia, and encourages qualified female academics to apply.

Severely disabled and disabled applicants with the same status will receive preference over equally qualified non-disabled applicants.

Instructions for applying

Contact

Prof. Dr. Anne Lauscher

Miriam Döring

anne.lauscher@uni-hamburg.de
[+49 40 2395-27967](tel:+4940239527967)

miriam.doering@uni-hamburg.de
[+49 40 2395-26106](tel:+4940239526106)

Location

Bundesstraße 56b
20146 Hamburg
[Zu Google Maps](#)

Reference number

80

Application deadline

2026-04-08

Use only the online application form to submit your application with the following documents:

- cover letter
- CV
- copies of degree certificate(s)

If you experience technical problems, send an email to bewerbungen@uni-hamburg.de.
More information on [data protection](#) in selection procedures.

VIelfALT 
GESTALTEN
RE-AUDIT
DES STIFTERVERBANDES

ZERTIFIKAT 2024

Die Universität Hamburg ist zertifiziert. audit
familiengerechte hochschule

