University of Stuttgart
Germany

The Independent Research Group in the field of **AI Ethics** headed by **Dr. Thilo Hagendorff** invites applications for a

Ph.D. position (f/m/d) in AI Ethics

(TV-L 13, 100%)

Join the Independent Research Group and work on research related to generative AI, machine behavior, large language models, and their ethical implications. As a Ph.D. candidate, you will have the opportunity to conduct socially relevant research and work on your Ph.D. while contributing to the development of responsible and ethical AI. Apply now and become a part of a dynamic research group! The group is part of the project [Reflecting Intelligent Systems for Diversity, Demography, and Democracy](#) (IRIS3D) and is located at the [University of Stuttgart, Germany.](#)

Details

Within IRIS, Dr. Thilo Hagendorff will build a new research group with an envisaged starting date in early **summer 2023**. We aim at building a dynamic research group dedicated to exploring the ethical implications of AI, especially the ethics of generative AI, such as large language models (LLMs). Our team is passionate about understanding the behavior of these powerful models and ensuring that they align with ethical principles. Potential research questions the candidate would like to tackle are: emerging ethical implications of generative AI, empirically investigating machine behavior in LLMs, applying tools from psychology to investigate reasoning in LLMs, evaluating moral decision-making in LLMs, studying simulated group dynamics of LLMs interacting with each other, etc.

Since AI technologies are constantly evolving, it is also crucial for our team to stay up-to-date with the latest advancements and techniques. Therefore, we are seeking a researcher who is open to new ideas, willing to learn and explore different approaches to tackling ethical issues in the AI field. We expect you to be proactive in identifying new areas of research in response to emerging trends or changes in the field. Ultimately, our goal is to contribute to the **development of responsible and ethical AI**. As a researcher in our group, you will have the opportunity to work on cutting-edge investigations and studies in an emerging and rapidly growing field of AI research. You will publish research findings in high-quality academic journals and present at relevant conferences.

Dr. Thilo Hagendorff invites excellent early career researchers to apply for a **Ph.D. position** to complement the newly built team's research focus. The Ph.D. position will focus on the connection between machine learning, ethics, and psychology related to the aforementioned topics.

Potential areas of research

- AI ethics or ethics of technology
- Large language models and their machine behavior
- The intersection of psychology or cognitive science and machine learning
- Science and technology studies or similar fields

We invite **applications from all fields**. Applicants should have strong competence in bringing societal perspectives into the interdisciplinary/transdisciplinary reflection on intelligent systems and bridging multiple fields of disciplines.

Eligibility criteria

- Ph.D. candidates hold (or expect to complete soon) a Master's or equivalent degree
- Familiarity with large language models, generative AI, or other areas of machine learning
- An interest in the ethical implications of these technologies
- Willingness to learn new skills and knowledge
- Strong analytical and programming skills (being proficient in Python and statistics is a plus)
- Advanced English language skills (knowledge of German is not required)

What we offer

The Ph.D. position offers high-potential researchers an exceptional opportunity to conduct societally relevant research and work on their Ph.D. Additionally, the position benefits from support structures offered by the Graduate Academy of the University of Stuttgart (GRADUS) and the SRF IRIS.

- The position is **TV-L 13** according to the German federal wage agreement. If prerequisites are met, the starting group can be higher depending on the applicant's experience and qualifications.
- The initial contract will be for up to **3 years**.
- The envisaged start date is **01.05.2023** but can be adjusted to suit the availability of the successful candidate.

How to apply

Please submit your application in a **single PDF file** (up to 10 MB) to thilo.hagendorff@iris.uni-stuttgart.de, including the following:

- **Cover letter** describing your background and research interests
- Up to three of your most important **achievements** (this can, but must not only include achievements that are valued by society, but also achievements that you value personally, like caring for a specific cause, learning from a mistake, etc.)
- **Academic CV**, containing a list of publications and lectures (if any)
- **Certificates** (Bachelor, Master). Please include a certified translation of the diploma for languages other than German or English.
- **Reference letters**, if available

Applications submitted before 15.04.2023 will be given priority; however, applications will be accepted **until the position is filled**. Inquiries should be sent to Dr. Thilo Hagendorff at thilo.hagendorff@iris.uni-stuttgart.de.

The University of Stuttgart would like to increase the proportion of women in the scientific field and is therefore particularly interested in **applications from women**. Severely disabled persons are given priority in the case of equal suitability. Information on how the University protects personal data under Article 13 of the GDPR can be found via the following link: https://www.uni-stuttgart.de/en/privacy-notice/jobapplication/

We look forward to your application!