

Prof. Dr. Ralph Krüger – Epistemische Verzerrungen und sozialer Bias in sprachbezogener Künstlicher Intelligenz



Prof. Dr. Ralph Krüger

hat die Professur für Sprach- und Übersetzungstechnologie am Institut für Translation und Mehrsprachige Kommunikation an der Fakultät für Informations- und Kommunikationswissenschaften inne.

Dort forscht und lehrt Prof. Dr. Ralph Krüger zu sprachbezogener Künstlicher Intelligenz in Form von Large Language Models und maschineller Übersetzung und deren Einsatzpotenzialen in mehrsprachiger Fachkommunikation und Translation.

Ausgangslage

Die Translationswissenschaft sieht sich traditionell als Hüterin soziokultureller und epistemologischer Vielfalt und befasst sich dementsprechend auch mit Geschlechteraspekten. Gleichzeitig werden in der Forschung zu maschineller Übersetzung und Large Language Models Aspekte von sozialem Bias (u. a. Gender Bias) in dem Output der entsprechenden Systeme untersucht.

Als Reaktion auf neue Entwicklungen im Bereich der sprachbezogenen KI befasst sich Prof. Dr. Ralph Krüger mit einem breiter gefassten Bündel von KI-Kompetenzen (Artificial Intelligence Literacy) und erarbeitet derzeit einen entsprechenden KI-Kompetenzrahmen für Translation und Fachkommunikation, der über diese Berufsbilder hinaus eine breitere gesellschaftliche Gültigkeit beanspruchen könnte.

Erste Ideen zu Forschungsfragen

Der KI-Kompetenzrahmen für Translation und Fachkommunikation ist unterteilt in die fünf Dimensionen 1) Technische Grundlagen, 2) Domänenspezifische Performanz, 3) Interaktion, 4) Implementierung und 5) Ethische/Gesellschaftliche Aspekte.

Im Rahmen des Fellowships soll die Teildimension „Epistemische Gewalt/Verzerrung und Reproduktion von sozialem Bias“ mit Fokus auf Gender Bias in der KI inhaltlich beforscht und didaktisch operationalisiert werden, zunächst durch die Definition entsprechender Kompetenzdeskriptoren, dann durch die Erarbeitung geeigneter digitaler Lehrressourcen und schließlich durch die Erforschung von deren didaktischer Wirksamkeit.

Die in dem Forschungsvorhaben zu entwickelnden Lehrressourcen sollen so breit aufgestellt sein, dass sie über die Berufsfelder Translation und Fachkommunikation hinaus Relevanz für Bürger*innen KI-saturierter Gesellschaften entfalten und einen Beitrag zur gesamtgesellschaftlichen Sensibilisierung für das Thema Gender Bias (und andere soziale Verzerrungen) in aktuellen KI-Technologien leisten können. Die Lehrressourcen werden als Open Educational Resources unter einer Creative-Commons-Lizenz zur Verfügung gestellt.

Literatur

Bianchi, Federico; Fornaciari, Tommaso; Hovy, Dirk; Nozza, Debora (2023). Gender and age bias in commercial machine translation. In: Moniz, Helena; Parra Escartín, Carla (Hg.): *Towards responsible machine translation. Ethical and legal considerations in machine translation*. Cham: Springer, 159–184.

Krüger, Ralph; DataLit^{MT}-Team (2022). Lehrressourcen zum Thema Social/Gender bias in machine translation. DataLit^{MT}-Projektwebseite. <https://itm.k.github.io/The-DataLitMT-Project/resources/#social-bias-in-mt>

Krüger, Ralph (im Erscheinen). Outline of an artificial intelligence literacy framework for translation, interpreting and specialised communication. *Lublin Studies in Modern Languages and Literature*.

Krüger, Ralph (zur Veröffentlichung eingereicht). Künstliche Intelligenz in Translation und Fachkommunikation – Skizze eines Kompetenzrahmens. In: Czulo, Oliver/Hoberg, Felix/Kappus, Martin (Hg.): *Digitale Translatologie*. Berlin: Language Science Press.

Vanmassenhove, Eva (2024). Gender bias in machine translation and the era of large language models. *arXiv*.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Technology
Arts Sciences
TH Köln