

Foundation Models - Anwenderkreis

Foundation Models und Grundlagen für einen vertrauenswürdigen Einsatz von KI

Spätestens der mediale Hype um ChatGPT, Dall-E und Co. haben sog. *Foundation Models* (auch *Large Language Models* bzw. *Große oder Generative Sprachmodelle*) in das Licht der Öffentlichkeit gerückt. Die Gründe sind naheliegend: Generative Modelle, die nicht konkret für einen spezifischen Zweck trainiert werden, sondern aus großen Mengen nicht-annotierter Daten eigenständig Strukturen erlernen und variabel auf eine Vielzahl von Fragestellungen angewendet werden können, kommen menschlicher Intelligenz so nahe wie kein anderes Paradigma künstlicher Intelligenz. *Foundation Models* versprechen neue Erfolge und Chancen in vielen Anwendungsfeldern, angefangen bei der Sprachverarbeitung über Computer Vision bis hin zu biomedizinischen Anwendungen. Gleichzeitig ergeben sich bereits neue Herausforderungen hinsichtlich der Vertrauenswürdigkeit dieser Modelle und ihrer Auswirkungen auf unsere Gesellschaft und unseren Arbeitsalltag, während auch die Diskussion um die Prüfung der Vertrauenswürdigkeit von neuronalen Netzen allgemein noch nicht abgeschlossen ist.

Ziel dieser Workshopreihe ist es einerseits, an die Technologie der *Foundation Models* heranzuführen und gemeinsam mit Wissenschaft, Industrie und Anwender*innen zu erarbeiten, welche Chancen, Herausforderungen und Lösungsmöglichkeiten sich durch den Einsatz dieser ergeben. Andererseits geht die Workshopreihe aber auch auf allgemeine Grundlagen ein, die sowohl für den vertrauenswürdigen Einsatz von einzelnen anwendungsspezifischen „klassischen“ KI-Modellen als auch für große generative Modelle notwendig sind. Die Workshopreihe ist in vier voneinander unabhängige Workshops aufgeteilt und richtet sich neben den Partnern und assoziierten Partnern des ZERTIFIZIERTE KI-Projektes an alle interessierten Expert*innen.

1. Workshop: Foundation Models – Chancen und Herausforderungen

Im ersten Workshop wird ein technischer Überblick über die Großen Sprachmodelle bezüglich der Sprachverarbeitung gegeben. Des weiteren werden ethische Anforderungen hinsichtlich der Vertrauenswürdigkeit, Absicherung und Zertifizierung diskutiert, sowohl aus wissenschaftlicher als auch aus Unternehmens- und Anwender*innenperspektive.

Termin: 20. Juni 2023

2. Workshop: Multimodale Foundational Models – Funktionsweise und Einsatz als Werkzeug zum semantischen Testen

Multimodale Foundation Models sind große generative KI-Modelle, die nicht nur Sprache verarbeiten, sondern auch andere Datentypen wie beispielsweise Bilder. Durch einfache textuelle Anfragen lassen sich innerhalb von Sekunden überraschend spezifische und realistische Bilder generieren. Die dahintersteckende Technologie verspricht jedoch auch weitere Einsatzmöglichkeiten, bspw. zum semantischen Beschreiben von Metadaten, wiederum eine Grundlage zum effektiven Testen von neuronalen Netzen. In diesem Workshop wird aus wissenschaftlicher Perspektive in Methodik und Verfahren von multimodalen Foundation Models eingeführt und der Einsatz als Hilfestellung zum semantischen Testen neuronaler Netzwerke diskutiert.

Termin: 27. September 2023

3. Workshop: Datenqualität – Grundlage für Vertrauenswürdigkeit verschiedenster KI-Systeme

Daten sind der essenzielle Rohstoff sowohl für KI-Systeme, die allein für eine spezifische Aufgabe trainiert werden, als auch für große Foundation Models, die riesige Mengen an Daten verarbeiten. Im Entwurf für den EU AI Act wird gefordert, dass die Trainings-, Eingabe- und Ausgabedaten von KI-Systemen relevant, repräsentativ, fehlerfrei und vollständig sein müssen. Ziel dieses Workshops ist es zu verstehen, welche die Merkmale von qualitativen Daten sind, wie Datenqualität verbessert werden kann und wie sie die Verlässlichkeit und Fairness von KI-Systemen beeinflusst.

Termin: 30. Oktober 2023

4. Prüfframeworks und Infrastruktur als allgemeine Basis zur Gestaltung von vertrauenswürdiger KI

Um KI-Anwendungen zu prüfen, werden neuartige und hochwertige technische Software-Werkzeuge und Testverfahren von verschiedenen Herstellern entwickelt und angeboten. Diese müssen anwendbar sein für eine Varianz von einfachen linearen Entscheidungsbäumen bis hin zu komplexen generativen DL-Modellen sowie für eine Vielzahl von verschiedenen Anwendungsbereichen mit spezifischen Anforderungen. Die Herausforderung besteht darin, die verschiedenen Prüftools zu kombinieren und einfache, reproduzierbare und nachvollziehbare Tests zu ermöglichen. Dazu werden in diesem Workshop Werkzeuge, Schnittstellen sowie das Konzept für ein Prüfframework vorgestellt.

Termin: 29. November 2023

Gefördert durch:

Ministerium für Wirtschaft,
Industrie, Klimaschutz und Energie
des Landes Nordrhein-Westfalen

