



FH Münster | Hüfferstraße 27 | 48149 Münster

Landtag Nordrhein-Westfalen
Wissenschaftsausschuss

- per E-Mail an:
anhoerung@landtag.de -

LANDTAG
NORDRHEIN-WESTFALEN
18. WAHLPERIODE

STELLUNGNAHME
18/589

Alle Abgeordneten

Der Vorsitzende

Prof. Dr. Bernd Kriegesmann
Telefon: 0209 - 9596 461
Telefax: 0209 - 9596 562
E-Mail: praesident@w-hs.de

Geschäftsführer

Robert von Olberg
Telefon: 0251 - 83 64019
E-Mail: robert.von-olberg@fh-
muenster.de

Gelsenkirchen/Münster, 05.06.2023



Anhörung am 19.06.2023 zum Antrag „Wo bleibt ein deutsches ChatGPT? – Nordrhein-Westfalen zur Deep-Tech-Fabrik machen!“, Drucksache 18/3285

Sehr geehrte Damen und Herren Abgeordnete,

für die Einladung zur o.g. Anhörung und die Gelegenheit zur Stellungnahme bedanken wir uns als Hochschulen für Angewandte Wissenschaften (HAW) sehr herzlich. Wir teilen die Auffassung, dass KI im Bereich Forschung und Transfer jetzt und auch in den nächsten Jahren eine wesentliche Rolle in NRW spielen wird. Daher begrüßen wir die Befassung des Ausschusses mit der Thematik im Rahmen einer Anhörung.

Standortbestimmung

Zu der im Prinzip positiven Standortbestimmung gehört, dass wir kein „deutsches“ ChatGPT im engen Sinne brauchen, sondern gute rechtliche Rahmenbedingungen und einen breiteren Blick auf die komplexe KI-Landschaft. „Berge“ wie ChatGPT, Bard und Midjourney versperren den Blick auf die kleingliedrige, aber trotzdem auch effiziente Landschaft in Europa, Deutschland und NRW. Es ist dabei tatsächlich essentiell, sich nicht auf amerikanische Lösungen zu verlassen, sondern die europäische Dimension in NRW zu fördern.

Ein Beispiel, um Probleme mit Lösungen wie GPT-3 zu illustrieren: Europa zeichnet sich durch seine Vielfalt an Sprachen aus. Die Sprachmodelle, auch wenn vielfach keine Angaben vorliegen, holen sich ihre Daten aus dem Netz und hier ist die Verteilung sehr ungleichmäßig. W3Techs schätzt den englischen Anteil der Webinhalte auf über 60 %. Sprachen wie Deutsch oder Französisch liegen bei ca. 3 %, kleinere europäische Sprachgemeinschaften darunter. Daher sind die Modelle, wenn man nicht aktiv auf die Förderung solcher Sprachen hinwirkt, immer schlechter in jeder anderen Sprache als Englisch.

Bis zum Start des Hypes um ChatGPT 4 sind die meisten der sogenannten Large Language Models (LLM) auch als Open Source Modelle veröffentlicht worden und gut dokumentiert. GPT-2 hatte 1,5 Milliarden trainierbare Parameter, LLaMA 65 Milliarden, GPT-3 175 Milliarden und über GPT 4 schweigt sich OpenAI aus. Aktuell schaffen es jedoch Forschungseinrichtungen und Open Source Communities, durch weiterentwickelte Modelle mit nur 17 Milliarden Parametern sehr nah an die Leistung von GPT-3 heranzukommen. Ein Faktor 10 also, der hier faktisch eingespart wird

und durch den das Large Language Model nicht mehr ganz so „Large“ ist. Es ist folglich nicht zwangsläufig so, dass Größe der primäre Maßstab ist. Es gibt Untergrenzen, aber Wachstum entsteht durch Innovation bzgl. der besten Ausnutzung vorhandener Kapazitäten. Kleinere Modelle, die auf gewisse Anwendungen spezialisiert sind, haben weitere Vorteile bezüglich ihres CO₂-Fußabdrucks. Ein Modell mit sehr vielen Parametern benötigt sowohl im Einsatz als auch beim Training weit mehr Ressourcen. Dies führt aktuell dazu, dass mit ChatGPT noch kein Geld gemacht werden kann, sondern es momentan noch ein Verlustgeschäft ist.

Darüber hinaus sind auch Technologien wie LoRA, also Low-Rank Adaptation of Large Language Models, auf dem Vormarsch und erlauben nachträgliche Anpassungen existierender großer Foundation Modelle mit übersichtlichem Rechenaufwand. Dass die Innovation bei unveränderten Rahmenbedingungen eher von Open Source und KMU ausgeht, sehen auch Anbieter wie Google so. Neben Techniken wie LoRA gibt es bekannte Ansätze wie Continual Learning, die auch zu aktiven Themen im Kontext der NRW-Forschungslandschaft gehören.

Innovation in NRW kann also auch entstehen, indem offene Modelle verwendet, durch Unternehmen oder Forschungseinrichtungen in NRW angepasst und dann ggf. vermarktet werden. Schaut man auf die Open Source Modelle, muss man jedoch feststellen, dass diese leider auch nicht primär in Deutschland oder NRW erstellt wurden. Hier eine Auswahl von Modellen, die im Jahr 2023 released wurden:

- Alpaca (Stanford University): veröffentlicht im März 2023; Model size: 7B to 65B models parameters; non-commercial license
- Vicuna (UC Berkeley, CMU, Stanford, and UC San Diego): veröffentlicht im April 2023; Model size: 13B; Apache 2.0
- RedPajama (Together, Ontocord.ai, ETH DS3Lab, Stanford CRFM, Hazy Research, and MILA Québec AI Institute): veröffentlicht im April 2023; Model 200B; LICENSE: Apache-2.0
- StableLM (Stability-AI, UK): veröffentlicht im April 2023; Model sizes: 1.37B parameters, 7B parameters, 137B parameters; Non-commercial CC BY-NC-SA 4.0 license
- OpenAssistant (LAION-AI, Deutschland): veröffentlicht im März 2023; Model size 1.37B parameters; Apache-2.0

Die Liste ist bei Weitem nicht umfassend, vermittelt aber einen Eindruck. Die oben vermerkte Apache-2.0 Lizenz ist sehr gut sowohl bei der Entwicklung von Close-Software als auch von Open Source-Produkten einsetzbar. Wenn man aber europäische Innovation sucht, findet man diese zum Beispiel bei OpenAssistant (<https://open-assistant.io/team>). Yannic Kilcher, ein Absolvent der ETH Zürich, Christoph Schuhmann, ein Österreicher, und Andreas Köpf aus Münster (NRW) sind hier z.B. von Bedeutung. Auch bei Together (<https://www.together.xyz/>) spielen besonders schweizerische Forscher eine wichtige Rolle. Auf jedes dieser Open Source Modelle mit einer Apache-2.0 können Firmen aus NRW sehr unterschiedliche Geschäftsmodelle aufbauen, auch da es sich hier nicht um Copyleft-Lizenzen handelt und somit Patentfragen mitgedacht sind.

Zu den europäischen Gelingensbedingungen, auf die NRW nur einwirken, aber bei denen NRW nicht direkt mitbestimmen kann, gehört die Umsetzung des EU AI Act, so dass er die Innovation, die durch Open Source Modelle ausgeht, nicht versehentlich hemmt. Wir werden vermutlich in NRW keine großen „From-Scratch-Modelle“ sehen, sondern gemeinsame (Weiter-)Entwicklungen mit europäischen und außereuropäischen Partnern. Das bedeutet, dass die aktuell geplanten Regelungen zu den Foundation Modellen im EU AI Act großen Einfluss auf die Umsetzung in NRW

haben werden. Wenn das Land NRW Einfluss nehmen kann, sollte es das tun, um Open Source Foundation Modelle von der Regulierung auszunehmen bzw. diese in innovationsfreundliche Bahnen zu lenken.

Zu den Aspekten des Antrages

Die Frage ist also, wie Forschung, Innovation und Transfer in NRW bzw. für NRW gefördert werden können. Unten sind einige Gedanken aus dem Antrag aufgegriffen und aus Sicht der HAW kommentiert, wobei viele Sichtweisen auch auf kleine und mittlere Universitäten zutreffen dürften.

Vernetzung bezüglich KI in NRW fördern

Der Antrag nimmt insbesondere Bezug auf die Struktur des KI.NRW. Wir möchten diesen gerne auf die sehr erfolgreichen dezentralen Strukturen erweitern und insbesondere die Perspektive der HAW einbringen. Ebenso wie Deutschland durch seine kleinen und mittelständischen Unternehmen und deren dezentrale Innovationskraft stark ist, so ist die Hochschullandschaft in NRW vielschichtig, in ihrer Dichte einzigartig und als solche für die Partner vor Ort besonders leistungsfähig. Hierzu tragen viele Professuren an HAW und klassischen Universitäten in der Informatik und außerhalb bei, die nicht den Begriff "KI" in ihrer Denomination führen.

Innerhalb des Promotionskollegs NRW (PK NRW) ist die Abteilung „Informatik und Data Science“ sehr aktiv im Bereich Maschinelles Lernen und künstliche Intelligenz. Es wäre wünschenswert, das PK NRW hier für Vernetzungsgedanken konsequent mitzudenken. Dazu muss es weiter gestärkt werden. Aktuell fehlt das HAW-Äquivalent zu Landesstellen an Universitäten für die Durchführung von Promotionen. Mit diesem notwendigen Aufbau von Mittelbaustellen für die Forschung könnte z.B. im Bereich KI begonnen werden.

Ebenso wichtig ist es, die Vernetzung von Universitäten und HAW weiter zu fördern, um die unterschiedlichen Profile für mehr Innovation zusammenzubringen. Förderlinien, die konsequent eine Vernetzung einfordern – am besten jeweils drei Partner Universität, HAW bzw. PK NRW und regionale Wirtschaft – wären ein wichtiges Instrument. Sie können mit Blick auf die kommenden Programme im Rahmen der Deutschen Agentur für Transfer und Innovation (DATI) ein wichtiger Nukleus für regionale Cluster mit starker HAW-Beteiligung sein und so Geld für den Transfer der Forschung zu Start-Ups und KMU nach NRW holen. Zwar ist das DATI-Konzept noch nicht abgeschlossen, doch sind die großen Linien mit einem Schwerpunkt auf HAW, Regionalität und den Transfer angewandter Forschung klar erkennbar. Es ist davon auszugehen, dass KI und deren Anwendung ein wichtiges Feld für die DATI sein werden. Stehen gute dezentrale Kollaborationsmöglichkeiten zur Verfügung, hat NRW großes Potenzial. Einige sehr innovative Firmen unterschiedlicher Größenordnungen aus NRW seien beispielhaft genannt: Bei den eher großen fällt sicherlich die Firma Aptiv als Automobil-Zulieferer mit hoher KI-Affinität auf. Dann gibt es mittlere Firmen wie z.B. ControlExpert oder auch kleine innovative Firmen wie Lorent IT-Lösungen, die sich schon lange mit Chat-Bots beschäftigen; alle haben ihren Sitz in NRW. Bei all diesen Firmen gilt, dass diese auch in Innovationsnetzwerken aus Hochschulen – Universitäten und HAW, oft gemeinsam – eingebunden sind. Diese regionalen, innovativen Netzwerke gilt es zu stärken.

Rechenkapazitäten für KI-Anwendungen erhöhen

Um im Rahmen solcher möglicher Landes-, aber auch existierender Förderungen der Bundesministerien und der EU leistungsfähig zu sein, sind in der Tat mehr Rechenleistungen ein wichtiger Aspekt. Hierzu sollte man sich klarmachen, dass der Marktwert des Trainings des Generativen Bilderzeugungsmodells Stable Diffusion mit 600000 \$ angegeben wurde. Auch hier gilt: Deutsch-

land ist sehr innovativ, denn die Grundlagen für dieses Modell wurden an der LMU München gelegt. Aber um die Anwendungen zu skalieren, fehlen im Land oft die Möglichkeiten. Für HAW, aber auch für kleinere Universitäten ist die eigene Bereitstellung von Rechenleistung, wie sie heute nötig ist, wirtschaftlich schwer umsetzbar. Jede einzelne Bildungseinrichtung müsste Personal für Wartung und Weiterentwicklung bereitstellen. Darüber hinaus sind die nötigen Ressourcen in der Breite besser auszuschöpfen. Es gibt entsprechende Überlegungen im Rahmen von HPC.NRW, dies zu fördern. Hier wäre es wichtig, zu beschleunigen und sehr schnell die nötigen Strukturen zu schaffen. Wir brauchen hier Umsetzungszeiten von maximal einem Jahr. Das gilt auch für die anderen diskutierten Initiativen zu GitLab etc. für NRW. Die Infrastrukturprojekte müssen im IT Bereich mit den dortigen Innovationszyklen mithalten und schneller kommen.

Teilweise um Diskussionen über Datenschutz mitzudenken, sollte man überlegen, ob man nicht klarer die Hochschulen mandatieren kann und/oder bei größeren Projekten Datenschutzfragen zentraler regeln kann, um Verzögerungen wie aktuell beim Forschungskernsatz zu verringern. Um auch Start-Ups und KMU mitzudenken, wäre es sinnvoll, diesen im Rahmen von Projekten mit Hochschulen z.B. in öffentlich geförderten Projekten Rechenzeit auf diesen Clustern für Prototypen und Vorentwicklung zu ermöglichen. Ein garantierter Zugriff auf skalierende Rechenleistung würde auch die Chancen von Hochschulen aus NRW – unabhängig von ihrem Typ – in Ausschreibungen erhöhen, da diese ggf. sogar als Eigenleistung eingebracht werden kann.

KI-Professuren in Nordrhein-Westfalen erhöhen

Die Studienanfänger_innenzahlen – gemittelt über alle Fachbereiche – werden in NRW in den nächsten Jahren vermutlich sinken. Dazu kommt das Umstellungsjahr von G8 auf G9, in dem fast keine Absolvent_innen von Gymnasien erwartet werden. Diese demografische Dividende sollte man nutzen und Forschungsprofessuren an HAW kapazitiv wirksam einrichten und hier Schwerpunkte setzen, wie z.B. KI & Maschinelles Lernen. Ein Ansatz könnte sein, z.B. 10% - 15% der Professuren einer HAW als Forschungsprofessuren mit 9 - 12 SWS Lehrverpflichtung zu ermöglichen. Hierdurch könnten existierende forschende Kolleg_innen mehr Aktivitäten freisetzen sowie bei Neuberufungen könnten Anreize für NRW als Arbeitsort gegeben werden. Würde man darüber hinaus den HAW zusätzliche solcher Forschungsprofessuren dauerhaft finanziell ermöglichen, so wäre der Effekt entsprechend größer und die zuvor und im Anschluss geschilderten Ansätze hätten die Chance, angewandte Forschung in NRW zu stärken.

Gewinnung (ausländischer) Wissenschaftler_innen verbessern

Die Gewinnung von Wissenschaftler_innen für den Standort NRW ist ein wichtiger Aspekt. Dabei geht es zum einen um den internationalen Wettbewerb, und zum anderen um den nationalen. Durch die 3,5 Milliarden Euro starke Hightech Agenda Bayern hat NRW einige Forscher_innen nach Süddeutschland verloren. Das gilt für Universitäten und HAW. Im internationalen Wettbewerb stellen die Bürokratie und die Kapazitäten der Ausländerämter oft einen Standortnachteil dar. Jede Veränderung, die dazu führt, dass sich internationale Wissenschaftler_innen hier als Gäste fühlen und leichter aufgenommen werden, ist hilfreich. Im Rahmen der immer umfangreicher werdenden Verfahren bei der Aufnahme von Wissenschaftler_innen wäre es wünschenswert, das Serviceangebot des ZAW.NRW auszubauen und hier vollumfängliche Prüfungen (Exportkontrolle) durchzuführen. Dies würde insbesondere kleineren Hochschulen helfen, da AI in diesem Umfeld ebenfalls ein Thema ist und die Gesetzeslage an Komplexität zunimmt.

Open Data fördern

Open Data ist, wie im Antrag bemerkt, ein wichtiger Baustein für Innovationen im Bereich Maschinelles Lernen und Künstliche Intelligenz. Voraussetzung für diese gute Ablage von Daten nach

den FAIR-Prinzipien ist ein erfolgreich umgesetztes Forschungsdatenmanagement. Es gibt bereits mit FDM-Scouts.nrw eine erfolgreiche Initiative für das Forschungsdatenmanagement an HAW des Landes. Um die Hochschulen bei dieser neuen und immer umfangreicher werdenden Daueraufgabe zu entlasten, sollte diese Initiative verstetigt und darüber hinaus um den Aspekt Open Data ergänzt werden. Dieser ist leider in den FAIR-Prinzipien nicht automatisch inkludiert. Setzt man dies konsequent um und versucht noch stärker, auch öffentliche Einrichtungen wie Kommunen und das Land selbst mitzudenken, können hier Datenschatze für zukünftige Innovationen entstehen.

Wir hoffen, mit diesen Hinweisen Anregungen für Ihre Diskussion zu dem vorliegenden Antrag gegeben zu haben, und freuen uns, die Diskussion in der Anhörung im Ausschuss vertiefen zu können.

Mit freundlichen Grüßen



Prof. Dr. Bernd Kriegesmann
Vorsitzender



Prof. Dr. Jörg Frochte
Vizepräsident für Forschung, Digitalisierung
und Internationalisierung der
Hochschule Bochum