Lernende Systeme

GERMANY'S PLATFORM FOR ARTIFICIAL INTELLIGENCE

Large Language Models

White Paper

Löser, A., Tresp, V. et al.
Working group Technological Enablers and Data Science

Executive Summary

Whether writing essays, poems, or programming code, large language models such as ChatGPT, BARD, and BLOOM are rapidly transforming how we work, communicate, and interact with information and knowledge. These AI models are more flexible and powerful than their predecessors. As a key technology, they serve as the core of many important applications: recognizing, generating, translating, and processing language. However, the underlying AI technology is not limited to language processing. Large language models therefore offer tremendous potential across all areas of life and drive fundamental and lasting changes not only in many industries but also in research and development.

Experts from the working group Technological Enablers and Data Science of the Plattform Lernende Systeme provide an overview in this white paper of the foundations on which large language models are based, their unique characteristics, and the research and development opportunities they present. The focus is primarily on examining the opportunities, challenges, and perspectives of research in Germany.

Explanation box

Overview of language models using ChatGPT as an example:

ChatGPT can …

- *Engage in natural-sounding conversations*
- *Generate texts of various kinds (essays, poems, summaries, recipes, etc.)*
- *Imitate desired styles (factual, poetic, etc.)*
- *Generate program code*
- *Translate texts, and much more*

It is based on …

- *General text data from the internet (Wikipedia, etc.)*
- *Books*

It operates …

- *Based on probabilities*
- *Using AI algorithms and extensive training, leveraging human evaluations of generated content*

Limitations …

- *Always provides a response (even if the data may not have sufficient information for the query)*
- *Occasionally fabricates content or sources („hallucinates")*
- *And others*

Source: Own compilation.

Perspectives from Research and Development

The rapid and disruptive development and dissemination of large language models present a dynamic field of action with numerous opportunities, research areas, and needs, as well as challenges for research and development. Research from Germany is actively involved in the development of large AI models and makes a significant contribution to topics such as language application services following European standards, design and augmentation for medical language models, explainability, benchmarking for evaluation, and much more. Since there are only a few large language models available for the German language, research and the (further) development of modern, open multilingual models as well as models for the German language should be promoted and supported in Germany and Europe. This gives rise to key research fields:

- **Pretrained large language models for the German language:** It is reasonable to develop large language models specifically for smaller language communities, such as the German language. Particularly in critical domains like medicine, justice, or security, further development allows for implementation while considering German and European laws as well as local values.

- **Multimodality:** Large language models benefit from additional contextual information that can be integrated through multimodality. This means going beyond textual data and incorporating images, ontologies, sets, tables, time-variant data, motion trajectories, and so on.

- **Multilinguality:** Multilingual models (trained on data from multiple language families) can potentially compensate for missing contextual knowledge in one language by leveraging existing contextual knowledge in another language. This is particularly helpful in domains with a global nomenclature or standard, such as medicine.

- **Combined AI or hybrid AI:** Significant progress has been made in language models by combining multiple methods that provide additional contexts. There is potential in combining knowledge-based AI with language models, for example, by incorporating knowledge graphs for faster adaptation to facts and circumstances.

- **Understanding context in long texts:** Language models typically condense information across texts and recognize connections. Therefore, language models that can process particularly long texts and contexts, such as those from prompts, are interesting for specific domains in the German language.

- **Explainability:** Research on explainability of AI aims to understand how a model arrives at its results, what happens during data processing, and whether results are based on unnoticed biases or correlations. This method allows for model improvement and further development, as well as strengthening users' trust in AI.

- **Differentiable Tokenizers:** Tokenizers are a method for breaking down character strings and multi-words into logical and cohesive units, making them essential as the foundation for large language models. They heavily depend on the domain and are currently often rule-based, which makes it challenging to adapt them to specific domains.

- **Benchmarks:** In research, benchmarks are used for evaluating large language models to investigate their limits, the emergence of capabilities, or to test their performance on specific tasks such as logical reasoning, solving mathematical problems, or answering questions. However, there is a lack of such benchmarks for numerous important domains in the German language, including the healthcare, industry, service, education sectors, mechanical and automotive engineering, as well as the legal field.

From the limitations of such models, additional research fields emerge that need to be addressed to enhance the language understanding of such models. For instance, accurately **relating multiple texts to each other**, **understanding temporal aspects** or **knowledge about knowledge**, as well as **modularity**, which involves separating knowledge about language and reasoning on the one hand and factual knowledge on the other hand (see Goldberg, Y. 2023)[1].

Moreover, it is crucial to place stronger emphasis on **context and grounding** and **detecting and addressing bias**. This is due to the prevailing fact that current language models often lack grounding and rely heavily on self-referentiality. Consequently, the description of real-world situations is solely based on language (as a single modality) and word references. In order to better understand and generate human language, additional modalities beyond texts could be integrated. However, actual grounding, in a narrower sense, can only be achieved by connecting models with reality through perception and action.

Various methods are available to researchers to detect bias in text collections used for training large language models, as well as within the models themselves, to address or minimize prejudices. For example, bias can be identified by adapting pretrained models to text collections specific to online communities. By adopting such an approach, potential biases in generative language models for dialogues can be addressed even prior to their deployment. Furthermore, differences between models can provide insights into the texts they were trained on, enabling the detection of hidden biases within a text collection.

1 Goldberg, Y. (2023): Some remarks on Large Language Models.
Online: https://gist.github.com/yoavg/59d174608e92e845c8994ac2e234c8a9 (retrieved 01/21/2023)

Fields of action

Research in Germany is in a good position to address the challenges in the (further) development of large language models, to develop them in line with European values, and to fully exploit their potential for Germany and Europe. However, in this dynamic research field, which is strongly influenced by non-European actors and large technology companies, it must not lose ground. Although German researchers are involved in the development of large language models in many areas, it is important to address the most pressing research needs and expand the foundations for harnessing the economic and societal potential of this AI technology for Germany and Europe, in line with European values. This requires a European ecosystem for the development of large language models and research that is closely aligned with practical applications, facilitating the transfer to industrial practices.

- **Options for promoting application-oriented research:** In order to efficiently and rapidly facilitate the industrial application of modern language models through targeted transfer research and funding, research efforts should focus on the automation of processes for model creation, adaptation, and maintenance. There is a specific need for models and methods for cost-effective adaptation of language models. Accordingly, research should be aligned with both a model-centric and a data-centric focus.

- **Access to publicly available, German-language training data:** Training data should follow broad distribution and cover a wide range of application scenarios, from domain-specific tasks such as legal documents and requirement analyses to colloquial free-form documents.

- **Evaluation and testing of large language models:** Language models should be tested for performance, robustness, bias, resource consumption, and other factors. The systematic collection of such tests for language models can serve as a basis for certifying large language models.

- **Strengthening the community for large AI models:** In order to advance the development of large language models across modalities in line with European values and norms, as well as application-oriented research that facilitates transfer to industrial practice, a European ecosystem is necessary, with interdisciplinary collaboration at its core. This ecosystem aims to address the technical challenges of AI technology, fully leverage its economic potential, and catch up with the major players from the USA and China.

Summary – Seizing opportunities and meeting challenges

Seizing Opportunities

- Utilize the adaptability of these models by developing efficient methods that facilitate domain adaptation, thus enabling transfer to applications: applying AI models to new domains or tasks.
- Better represent the German language community by contributing to multi-lingual models.
- Address a faster adaptation to new factual situations and circumstances through hybrid approaches, such as combining language models with knowledge graphs.
- Implement and further develop technical solutions for bias detection.
- Enhance the language understanding of models through multimodal approaches.
- Improve the grounding of language models by connecting them with perception and action.

Meeting Challenges:

- Improve methods for creating effective training datasets more efficiently, minimizing the need for human annotation.
- Address "hallucination" in generated content.
- Enhance logical abilities of models and ensure robust reasoning.
- Further develop approaches for explainability, consistency, and coherence.
- Counter recency bias in models.
- Identify and resolve problem areas that arise in multilingual models due to specific characteristics of the German language.
- Improve energy efficiency in model training.

Source: Own compilation.

Imprint
Editor: Lernende Systeme – Germany's Platform for Artificial Intelligence | Managing Office | c/o acatech | Karolinenplatz 4 | D-80333 Munich | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Follow us on Twitter: @LernendeSysteme | Status: May 2023 | Photo credit: AREE/Adobe Stock/Title

This executive summary is based on the whitepaper: *Große Sprachmodelle – Grundlagen, Potenziale und Herausforderungen für die Forschung. Munich*, 2023. The authors are members of the working group Technological Enablers and Data Science of Plattform Lernende Systeme. https://doi.org/10.48669/pls_2023-3

SPONSORED BY THE

Federal Ministry of Education and Research

acatech
NATIONAL ACADEMY OF SCIENCE AND ENGINEERING