idw - Informationsdienst Wissenschaft
Nachrichten, Termine, Experten

Pressemitteilung

Agentur für Innovation in der Cybersicherheit GmbH
Michael Lindner

06.06.2025
http://idw-online.de/de/news853491

Forschungsprojekte, Wettbewerbe / Auszeichnungen
Geowissenschaften, Informationstechnik, Mathematik, Physik / Astronomie, Wirtschaft
überregional

Making ChatGPT & Co usable for security domain applications

Making ChatGPT & Co usable for security domain applications HEGEMON tackles the evaluation and adaptation of generative foundation models for security-critical applications On 4 June 2025, the Agentur für Innovation in der Cybersicherheit GmbH (Cyberagentur) launched the call for HEGEMON - a new research programme aimed at the co-development of holistic benchmarks and AI model adaptations for safety-critical applications. In a competitive "everyone vs. everyone" setting, for the first time generative foundation models are to be adapted and implemented as prototypes for complex geoinformation tasks while being repeatedly evaluated with newly developed holistic measurements.

A common dilemma: You want to generate an image or summarise a long text. Taking the quick route, you task a generative AI service based on a foundation model (often a large multimodal language model such as GPT-4o), which is supposed to generate the desired image or summary by means of text prompting. Unfortunately, the first results can be frequently unsatisfactory or even incorrect.

The broad application possibilities of foundation models and associated process acceleration potentials are also of interest to German security authorities. However, if the example of text summarisation is transferred to the security context - e.g. a soldier using such a tool to summarise a long command - the potential source of error takes on a much broader meaning. To date, no comprehensive tests (so-called benchmarks) exist specifically for the German security context in order to evaluate the universal and application-specific impacts of pre-trained foundation models.

Against this background, the Cyberagentur's HEGEMON research programme ("Holistic Evaluation of Generative Foundation Models in the Security Context") was launched on 4 June 2025. Universities, colleges, research institutions, companies and start-ups are invited to make an offer with their innovative ideas. The programme is designed as a multi-year research competition with several phases and high disruptive potential.

"With HEGEMON, we are creating a unique experimentation environment to systematically enable the comprehensive evaluation of generative AI models in the security sector - beyond previous benchmark routines," says Dr Daniel Gille, project manager of the programme and Head of Artificial Intelligence at the Cyberagentur. "We invite all research-intensive players to participate in this challenge."

The research programme addresses a key gap in current AI development: Foundation models such as GPT-4, Midjourney or Claude that are developed internationally - mainly by private companies, mostly in the USA and China - are often opaque in their training data and model architectures and are increasingly dominating security-critical areas. In the German and European context in particular, there is a lack of opportunities to evaluate these models comprehensively and comparatively - especially with regard to complex, multimodal tasks and applications with high demands on safety and facticity.

Aim of HEGEMON is the development of domain-specific, holistic benchmark sets (consisting of tasks, metrics and test data sets) as well as customised AI models for defined use cases. This combination of benchmark and model

development will be tested in a competitive scenario in which all participants evaluate each other's solutions - an "everyone versus everyone" structure with integrated red/blue teaming for robustness testing.

"We are rethinking the evaluation of large models," continues project manager Dr Daniel Gille. "The desired benchmarks should not only measure what is technically possible, but also what is safe, explainable and relevant to the application."

The focus is on three use cases from the geoinformation sector:
- The generation of comprehensible text summaries on country-specific topics,
- the conversion of remote sensing data into vector data,
- and a map chatbot with intelligent text output (e.g. "Are there medical facilities on this map? Please share the coordinates if they exist.").

Cyberagentur uses the tried-and-tested pre-commercial procurement procedure (PCP) for implementation. It allows for risky but fully financed research outside of regular procurement law. PCP is ideal for disruptive projects where no mature market products yet exist - i.e. exactly for the research subject of HEGEMON. Participants retain their usage and exploitation rights, which further strengthens their innovative capacity.

The call for proposals is the first step in a multi-phase research process that extends over a total of three years. Following an evaluation phase, three interaction points will be defined at which benchmarks and models will be systematically compared, improved and re-evaluated.

The tender was published in the Supplement to the Official Journal of the European Union with the contract notice number TED 358520-2025 (https://ted.europa.eu/de/notice/-/detail/358520-2025). The deadline for submitting the short concept is 31 July 2025, 10:00 am. Participation is possible both alone and in a consortium.

Further information can be found at
https://www.cyberagentur.de/en/programs/hegemon/

Contact:

Agentur für Innovation in der Cybersicherheit GmbH
Große Steinstraße 19
06108 Halle (Saale)

Michael Lindner
Press Officer

Phone: +49 151 44150 645
E-Mail: presse@cyberagentur.de

Background: Cyberagentur

The Agentur für Innovation in der Cybersicherheit GmbH (Cyberagentur) was founded by the German Federal Government in 2020 as a fully in-house company of the Federal Government under the joint leadership of the German Federal Ministry of Defence and the German Federal Ministry of the Interior and Community with the aim of adopting an application-strategy-related and cross-departmental view of internal and external security in the field of cybersecurity. Against this background, the work of the Cyberagentur is primarily aimed at the institutionalised implementation of highly innovative projects that are associated with a high risk with regard to the achievement of objectives, but at the same time can have a very high potential for disruption if successful.

The Cyberagentur is part of the National Security Strategy of the Federal Republic of Germany.

The Cyberagentur is headed by Prof Dr Christian Hummert as Scientific Director and Managing Director and Daniel Mayer as Commercial Director.

wissenschaftliche Ansprechpartner:
Dr Daniel Gille, project manager of the programme and Head of Artificial Intelligence at the Cyberagentur

Originalpublikation:
https://www.cyberagentur.de/en/press/hegemon-erforscht-bewertung-und-anpassung-generativer-foundation-models-fuer-sicherheitskritische-anwendungen/

URL zur Pressemitteilung: https://www.cyberagentur.de/en/programs/hegemon/

Launch of the call for proposals for the HEGEMON research programme to evaluate generative AI models in a security context.
Nancy Glor
Cyberagentur