

Pressemitteilung

Alexander von Humboldt Institut für Internet und Gesellschaft

Frederik Efferenn

09.09.2024

<http://idw-online.de/de/news839368>

Forschungsergebnisse, Forschungsprojekte
Gesellschaft
überregional



Simba: Berlin Research Institute Presents New Free AI Application to Reduce Language Barriers in Online Texts

Simba has a clear mission: to make the internet more understandable for everyone. This web app, designed to simplify German-language texts, was developed by Freya Hewett, a researcher at the Alexander von Humboldt Institute for Internet and Society (HIIG). Simba offers two AI-powered solutions: a web app for simplifying personal texts and a browser extension that automatically summarises texts on websites. This reduces digital language barriers, enhances the reading experience and facilitates access to the German language.

To the web app: <https://publicinterest.ai/tool/simba/simplifier?lang=en>

To the browser extension: <https://publicinterest.ai/tool/simba/extension?lang=en>

People use the internet every day to keep up to date with current affairs. However, many articles on websites and online publications are difficult to understand because they have complex sentence structures and use technical jargon. According to Freya Hewett, a computational linguist at HIIG and the initiator of Simba, this particularly disadvantages individuals with learning difficulties or those learning German as a foreign language. "Our research shows that the websites of government bodies and institutions in the education and science sectors often exclude a significant proportion of the population from important information due to their complicated language," Hewett explains. "Simplified language can help bridge these gaps. Participation in society depends on being able to access information and services online."

Simba is the first free solution specifically designed to target end users, enabling them to simplify texts in their everyday life. Its AI-powered apps replace long words with shorter ones that have similar or identical meanings. They reduce sentence length and add additional information to make connections clearer. Until now, similar automated language-simplification solutions in Germany were fee-based and primarily used by institutions and businesses. For many people who need to have easily understandable texts, these services were inaccessible due to paywalls. Simba fills this gap by allowing everyone to simplify texts themselves. "Our goal is for Simba to become an everyday tool that helps all people who want to use text simplification in their daily lives", emphasises Dr. Theresa Züger, head of the research group where the AI application was developed.

Dr. Theresa Züger's team explores what qualities artificial intelligence should have for the common good of society. These insights are implemented in practical prototypes, with Simba being the first to be released. You can find more information on this on the Public Interest Ai website: <https://publicinterest.ai/>

One distinctive feature of Simba in this regard is its focus on common good. The AI app is operated on a non-commercial, not-for-profit basis, and the developers are available for inquiries. The source code and underlying models are openly accessible, allowing for transparent collaboration. Freya Hewett emphasises: "We want to

collaborate with a community of researchers, inclusion experts and users to continuously develop and improve Simba. The target groups we address – such as people with learning difficulties or those who do not speak German as their first language – are very diverse. Our goal is to improve the language model through ongoing feedback and create simplifications that genuinely benefit many people.”

Simba’s two applications are based on a text generation model. These models, also known as large language models or foundation models, are trained with large amounts of text data. They predict which word is most likely to come next in a sequence. Examples of such models include GPT-4, Mistral 7B and Llama. For Simba, the team used the Llama-3-8B-Instruct foundation model, which was fine-tuned with German-language newspaper articles.

“As with all text generation models, there is a possibility that automatically generated summaries may contain incorrect information,” explains Freya Hewett. “Nevertheless, we are confident that Simba provides valuable support.” Hewett advises carefully comparing the input and output text of the AI application to ensure the accuracy of the information.

Simba is available as a prototype free of charge for the time being, but the ongoing costs of operating such an AI application are significant. To ensure the continuous availability and development of Simba, HIIG is looking for additional partners. Companies and organisations are invited to get involved and support either the hosting or further development of the AI application. These partnerships aim to advance Simba’s mission and promote digital inclusion.

The beta version of Simba was developed as part of the “Public Interest AI” research group funded by the Federal Ministry of Education and Research (BMBF) at HIIG, under the leadership of Dr. Theresa Züger. Freya Hewett trained the language model and received technical implementation support from Hadi Asghari and Christopher Richter. Larissa Wunderlich designed the user interface and created the website.

wissenschaftliche Ansprechpartner:

Freya Hewett freya.hewett@hiig.de, Dr. Theresa Züger theresa.zueger@hiig.de

URL zur Pressemitteilung: <https://publicinterest.ai/tool/simba?lang=en> Overview of Simba

URL zur Pressemitteilung: <https://publicinterest.ai/tool/simba/simplifier?lang=en> To the web app

URL zur Pressemitteilung: <https://publicinterest.ai/tool/simba/extension?lang=en> To the browser extension

URL zur Pressemitteilung: <https://www.hiig.de/en/simba-text-simplification/> Instructions for Simba in simplified language



Simba: AI-supported text simplification

Prince David

<https://unsplash.com/photos/brown-lion-looking-up-in-macro-lens-photography-MMKAbQPIXg8>