CZECH DIGITAL LIBRARY

Petr Žabička[1], Jan Rychtář[2], Martin Lhoták[3], Filip Jebavý[1], Filip Kersch[3]

Organisations: 1) Moravian Library in Brno, Czech Republic; 2) Trinera s.r.o., Czech Republic; 3) Library of the Academy of Sciences, Czech Republic

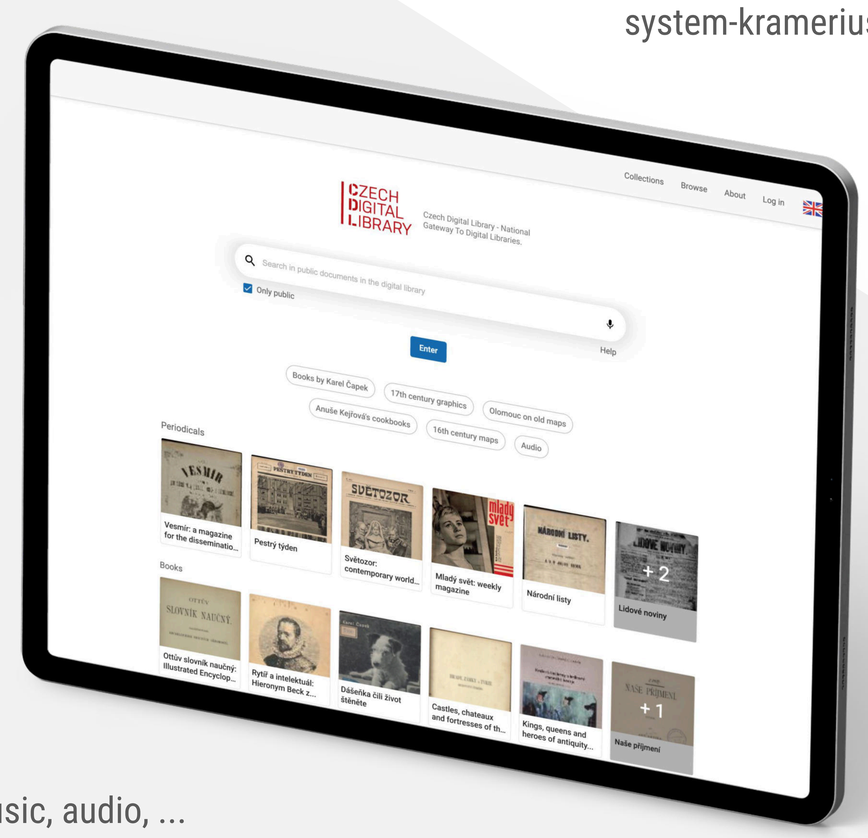
system-kramerius.cz/en

AI Tools for Digital Libraries: Enhancing User Experience and Trust

National Agregator of 15 largest 🇨🇿 digital libraries

based on the open-source Kramerius digital library system.

used by 40+ other libraries



293K	150M	1,5M	50K	20K
BOOKS	PAGES	ISSUES	MAPS	GRAPHICS

+ manuscripts, sheet music, audio, ...

Aggregator for europeana

Access: 14 % Open | 58 % Limited* | 28 % In-library

*Requires login with a valid Czech university or library account

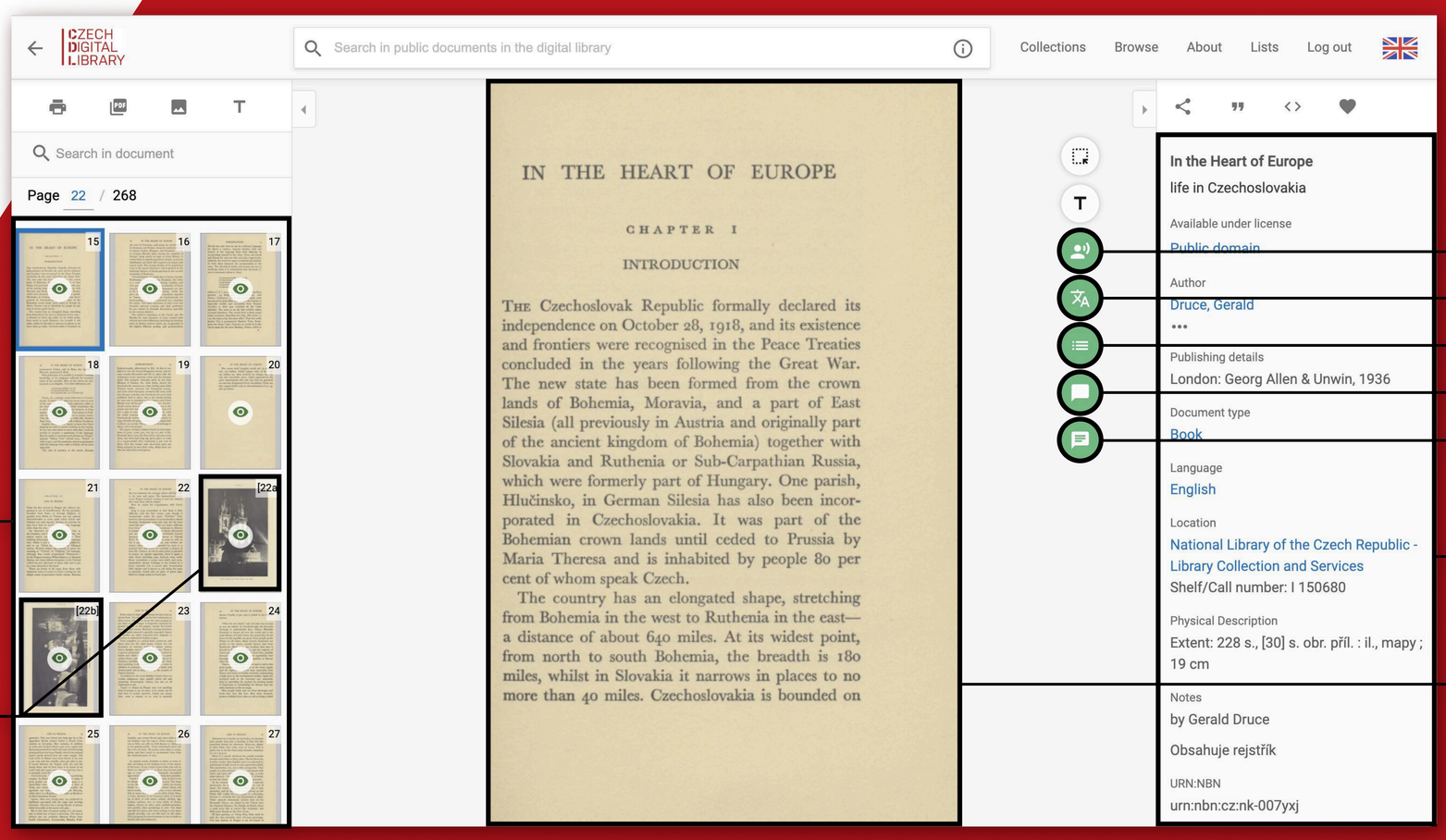
Java Solr

+ METS, MODS, ALTO XML, jpeg2000 images, pdf, epub.

USE OF AI IN CZECH DIGITAL LIBRARY

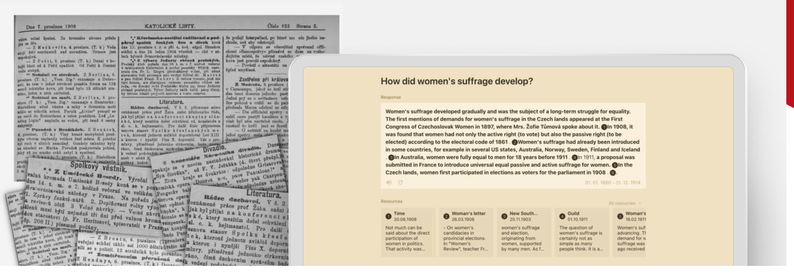
Czech Digital Library has been testing AI-powered features in its interface since 2023.

Users need to log in with their library, university, or Google account to access the AI tools.



Memory of Newspapers

25 historic newspaper titles (1880–1914), over 500,000 pages divided into 10 milion vectorized text chunks. Ask anything in natural language and get accurate LLM answers with clickable references to the original scanned pages.



Discover a Book

Semantic search across millions of bibliographic records from the Czech Union Catalogue (via the Central Portal of Czech Libraries). Books are represented as vectors, enabling natural-language queries to find relevant titles instantly.

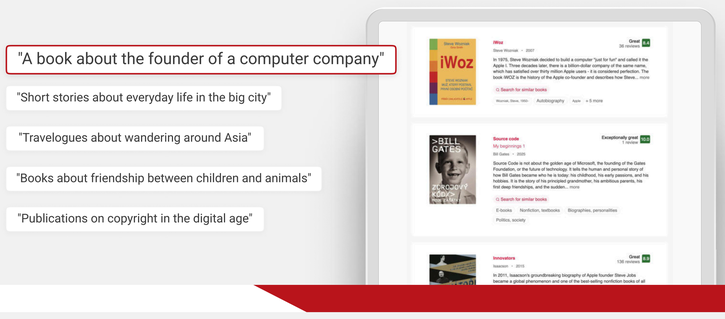
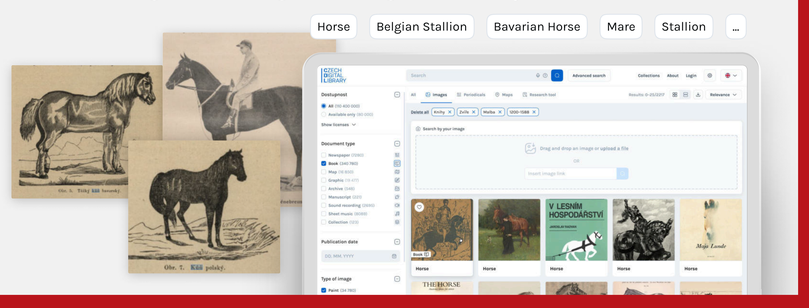


Image Search

Search images by text description or by uploading your own picture. Images are first automatically detected, analyzed, captioned, and converted to vector embeddings for both text-to-image and image-to-image similarity search.



Future Work and Challenges

1. Use of local models on copyright protected documents. Cost of external service vs. local. Legal limitations on copyright protected documents.
2. Pre-generating document summaries.
3. Fusing full text and semantic indexing (hybrid search).
4. Integration of retrieval augmented generation (RAG).
5. Marketing / user education. How to communicate new services to users? How to attract users?

MORAVIAN LIBRARY | LIBRARY CZECH ACADEMY OF SCIENCES | trinera

This research is funded by grants from the Ministry of Culture of the Czech Republic through the NAKI, VISK, and NPO programs and from the The Ministry of Education, Youth and Sports of the Czech Republic through the LINDAT/CLARIAH-CZ research infrastructure.