

Overview of Generative AI

Benjamin Ricaud
Workshop AI for Justice
Oslo, June 2025

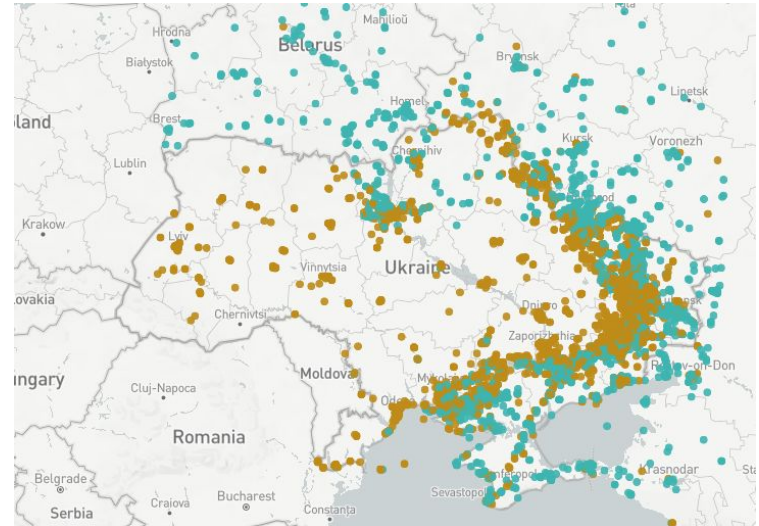
War crimes archiving in the age of social media

A new area with social media

- Enormous amount of documents
- Large variety and diversity of platforms and media (text, audio, video)

New tools are needed

-> AI is made for analysing data



eyesonrussia.org

Generative AI “GenAI”

What is it?

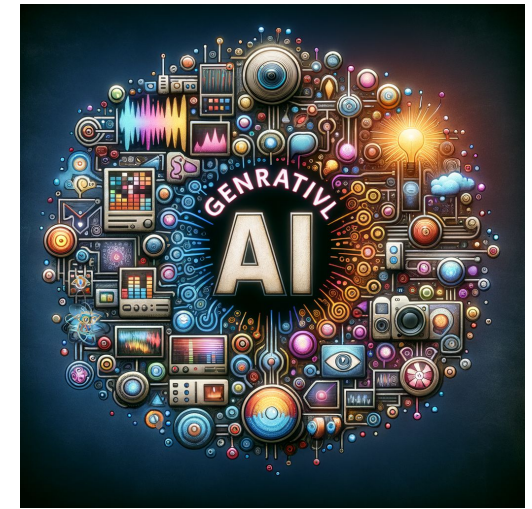
“AI” or “AI models” are computer softwares. It runs on computers, in the cloud or locally.

What can it do?

Generative AI models can output:

- text
- image
- audio
- video

AI can also “understand” text, image, audio, video (input)



First example: image description



Describe this image



02.06.25

14:51

The image shows a person wearing a navy blue sweater with red and white Nordic-style patterns. The background features indoor greenery, including a tree with thin branches and leaves, and a large window that reveals a snowy outdoor scene. The setting appears to be a bright, natural-lit indoor space.

Input: image (+ text)

Output: text



Image description



Describe this image



02.06.25

14:51

The image shows a person wearing a navy blue sweater with red and white Nordic-style patterns. The background features indoor greenery, including a tree with thin branches and leaves, and a large window that reveals a snowy outdoor scene. The setting appears to be a bright, natural-lit indoor space.



- Is it a good description?
- Anything missing?
- How can we obtain a better description?
- Can we use it for war crimes archiving?

War crime archiving use case:

- Generate metadata for each image

Speech to text

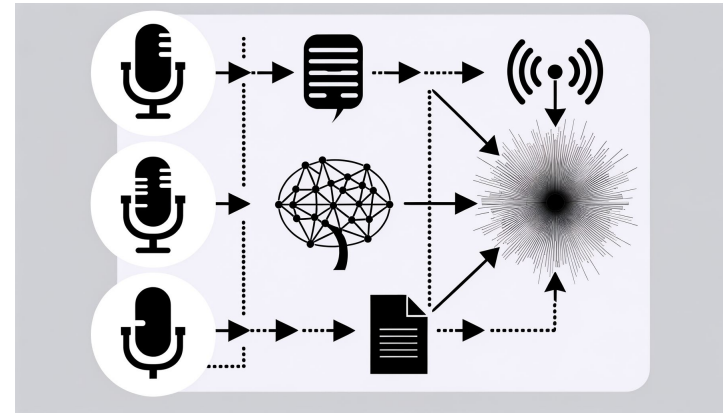
WHISPER model from OpenAI turns speech to text

Applications:

- Transcribing interviews, speeches
- Generating captions for videos (extracting the audio from video)

Use cases for war crime archiving ?:

- create metadata automatically (searchable) for videos
- combine with image description to make even richer video metadata
- Translating spoken content into other languages (using text translation model)

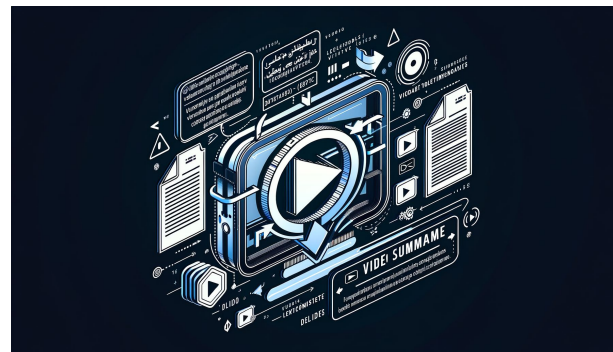



Nonsense AI generated image for
speech to text

Video to text or video-text-to-text

- Ask questions about a video

See [Huggingface.co](https://huggingface.co) :



 **Hugging Face**

[Models](#) [Datasets](#) [Spaces](#) [Community](#) [Docs](#) [Enterprise](#) [Pricing](#) [⌵](#)

Tasks [Libraries](#) [Datasets](#) [Languages](#) [Licenses](#) [Other](#)

[Reset Tasks](#)

Multimodal

[Audio-Text-to-Text](#) [Image-Text-to-Text](#)

[Visual Question Answering](#)


[Document Question Answering](#) [Video-Text-to-Text](#)

[Visual Document Retrieval](#) [Any-to-Any](#)


Computer Vision

[Depth Estimation](#) [Image Classification](#)


Models 139 [Full-text search](#)

 **BAAI/Video-XL-2**


[Video-Text-to-Text](#) • Updated about 8 hours ago • [📄 119](#) • [❤️ 22](#)

 **Haoz0206/Omni-R1**


[Video-Text-to-Text](#) • Updated 7 days ago • [📄 75](#) • [❤️ 10](#)

 **llava-hf/LLaVA-NeXT-Video-7B-hf**


[Video-Text-to-Text](#) • Updated Jan 27 • [📄 158k](#) • [❤️ 101](#)

 **Skywork/SkyCaptioner-V1**


[Video-Text-to-Text](#) • Updated Apr 25 • [📄 731](#) • [❤️ 41](#)


 **lmms-lab/LLaVA-Video-7B-Qwen2**

[Video-Text-to-Text](#) • Updated Oct 25, 2024 • [📄 180k](#) • [❤️ 96](#)

 **OpenGVLab/videochat2**

[Video-Text-to-Text](#) • Updated Aug 14, 2024 • [❤️ 3](#)

 **THUDM/cogvlm2-llama3-caption**

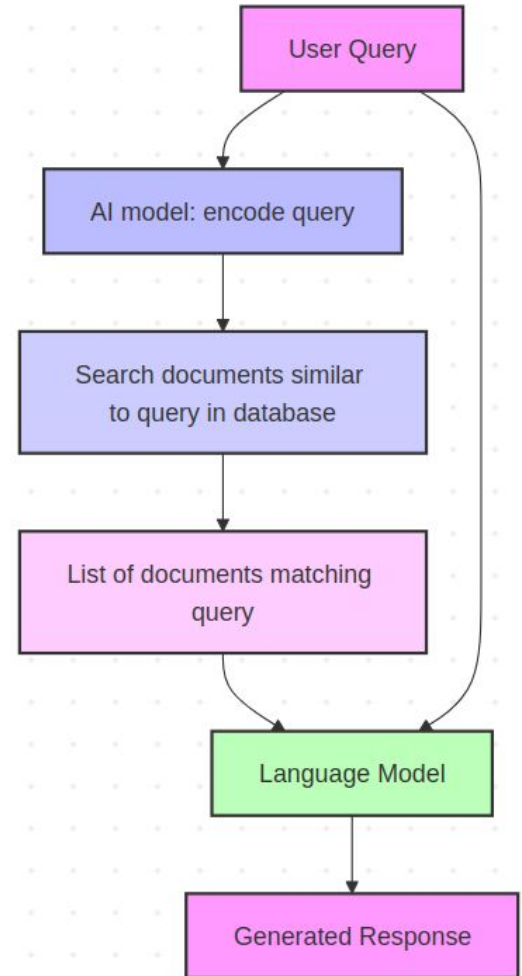
 **OpenGVLab/InternVideo2_5_Chat_8B**

Retrieval-Augmented Generation, RAG

Keyword search becomes chatbot-assisted search

- ask questions to your database
- “AI assistant” helping you search

More details later on!



In the cloud or locally?

Cloud solutions are exploding

- OpenAI
- Google
- ...
- AI assistants for writing
- Chatbots on many websites



Problems:

- which one to choose and trust?
- how to deal with sensitive or confidential data?
- is it safe? does it follow regulations?

Open-source

Most AI models are free and open-source on huggingface.co



**The AI community
building the future.**

These models can run on a local computer:

- no cloud
- data are safe on your computer or server
- needs a few lines of code to work

Future



shutterstock.com · 2326276253

- AI agents working together, with different specialists
- Smarter assistants, reasoning about the query and documents before answering
- New models specialist of war crime archiving (require engineering work)