

**THE (IM)POSSIBILITIES OF AI**  
***for the production of audio description***  
***for film and television content***

Research commissioned by:

Bartiméus Fund, Netherlands Blind Institute/NGO

CFAP, Netherlands Audiovisual film & tv producers

NPO Innovation, Dutch National Public Broadcaster



## Services

- Focus on **online videos** (informative/educational content)
- Subtitles, AD, transcripts and sign language translation
- 17 (EU) languages
- PROs / Freelancers / Partners

## Software

- Online video editor subtitles, AD (TTS), transcript
- Video publication tools for YouTube
- Fully accessible video player
- Integrations with Content Management Systems

**We help organizations meet the - WCAG 2.2 AAA / European Accessibility Act - regulations**

## ***Audio description can be created with AI, right?***

Market in need for a valid answer.

- Researcher: Wim den Herder (professional musician & mathematician)
- Research period: September 2024 to February 2025
- Material in Dutch: Short films & NPO TV content
- LLM AI Models: Gemini, Claude, Malmo, Mistral > Chat GPT 4o (best results in Dutch)
- Comparison of the human written AD with the AI generated AD (scored in %)
- Extra: Feedback round from both professional AD writers and AD-users on AI-AD results

# AI - “An eager, but clueless assistant, that needs micro-management”

Presentation Language at the media 2024

- The more input - the better the output
- ‘Magic’ happens in prompt engineering
- Ask AI: what to do, how to do it (examples/wording/no repetitions), for who to do it etc.
- Ask AI: to improve its own results
- Its output is never the same ! (refresh, refresh, refresh... until satisfied)

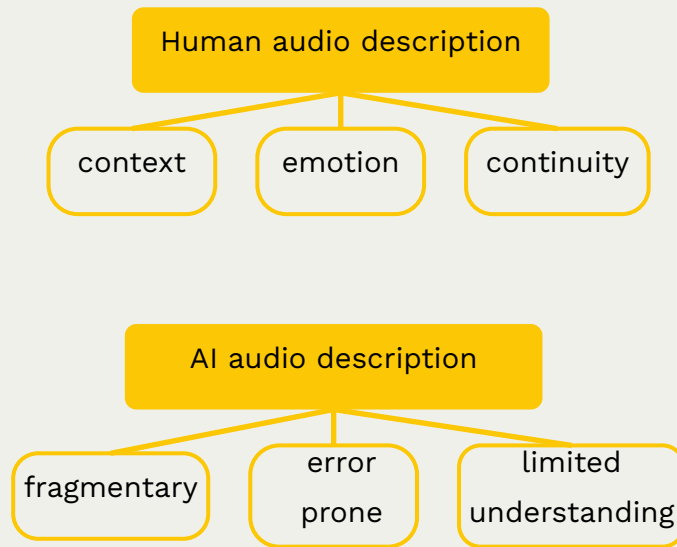
## Methodology

- ‘Phased’ prompt engineering
- Identify AD-areas
  - reversed subtitle track
- Identify scenes within AD-areas
  - based on changed camera viewpoints with a max. 15. sec.
  - send in 40 frames per scene
- Identify (and name) main film characters
  - AI written descriptions of person characteristics
  - screenshots of character with their name
- Send along the film script, the film synopsis and the film subtitle file

## Results - GPT 4o

*Impressive, but...*

- AI has potential, but falls short when it comes to narrative content.
- AI has no overview of the listeners' perspective
  - It doesn't guide the listener A, B, C...
  - Information that's already known
- AI can't pick up on nuanced emotions, small gestures
- AI can't deviate from scene
  - No pre-announcements / ad after a scene
- AI is sensitive to errors
  - AI-AD can add to confusion
  - Trained on most likely scenario - art usually deviates from norm



***Example:***

***A curtain closes in a dramatic scene,  
but AI describes it as “opening”  
—changing the entire narrative!***



*Example AI audio description*



*Example human audio description*

## What professionals think

- Professionals fear for quality and creativity
- To rewrite AI AD will take more time than writing from scratch
- See potential in AI for documentaries where content is less dependent on narrative

## What AD-users think

- Fragmentary - lacks coherence
- Disruptive to the story
- AI-AD ads to the puzzle - it is not a nice experience
- “Better than nothing”



## Conclusion

- AI-AD is an interesting development, but today's model is not usable for (the production of) narrative content
- There is potential for documentaries, provided human editor remains in lead
- AI development is moving in an ultra fast pace, so tomorrow's model can deliver other results

## What's next?

- Extended audio description
  - Add a editor feature to create EX AD
  - Video player that can support EX AD





**Interested in our research?**

**Let's talk!**

[eveline@scribit.pro](mailto:eveline@scribit.pro)

+31 (0)6 14 68 41 08

[www.scribit.pro](http://www.scribit.pro)