Evaluating AI Editing Algorithms for Video News Reporting

Caspian J. Moosburner
RheinMain University of Applied Sciences
Wiesbaden, Germany
email: caspianjade.moosburner@hs-rm.de

Dennis Quandt
RheinMain University of Applied Sciences
Wiesbaden, Germany
email: dennis.quandt@hs-rm.de

Matthias Kowald
RheinMain University of Applied Sciences
Wiesbaden, Germany
email: matthias.kowald@hs-rm.de

Wolfgang Ruppel
RheinMain University of Applied Sciences
Wiesbaden, Germany
email: wolfgang.ruppel@hs-rm.de

Till Dannewald
RheinMain University of Applied Sciences
Wiesbaden, Germany
email: till.dannewald@hs-rm.de

Matthias Narroschke
RheinMain University of Applied Sciences
Wiesbaden, Germany
email: matthias.narroschke@hs-rm.de

Abstract-With video news gaining more and more popularity, Artificial Intelligence (AI) video editing tools could be implemented to accelerate video news production. A challenging issue, however, is the quality of AI-edited video news and the acceptance of such news by media consumers. A survey with 143 participants is conducted in Germany in order to evaluate the quality of video news clips edited by AI models in comparison to news clips produced by professional human editors. All survey participants are recruited by a commercial survey company. The evaluation of the survey reveals that AI editing is widely undetected by the participants. Overall, the evaluation shows that the quality difference of AI edited video new clips and human edited ones is negligible as confidence intervals of measured quality features overlap. Future research can benefit from investigating the influence of clearly labelled AI content in user evaluations.

Keywords-Artificial Intelligence; AI; video editing; video news; algorithms.

I. INTRODUCTION

Video news has been established as a strong contender for attention in the digital landscape, aided by the rise of TikTok and their short form video content that is making a jump to other social media platforms [1]. This creates a need for quick video news reporting, a need that could be aided by implementing Artificial Intelligence (AI) editing tools.

With 40% of young people preferring social media search engines over traditional means [1], the market share of video news reporting on social platforms will continue to grow. This is aided by media publishers who put a bigger focus on digital video production [1] making it clear that the video news reporting demand will increase, and the market is preparing to supply by offering more and more AI tools. One example for this is the launch of the AI writing tool "ChatGPT" in 2022, establishing AI tools as more and more popular and expected to reach a global market share of "more than 2.5 trillion USD by 2032" [2]. Creative industries have implemented generative AI and AI algorithms, for example Adobe Photoshop offering options like generative fill [3], Canva that has its own AI image generation [4] or on countless social media websites via algorithmic AI [5].

The conducted survey is part of a larger survey study. A total of three survey waves are being collected with this paper focussing on the first wave. The first wave aims to gather some general information and results evaluating the algorithm performance.

A second survey wave is planned, aiming to investigate the implications of disclosing AI for quality evaluation in detail by testing audience reception to flagging clips explicitly as AI-edited or human-edited in two different groups. A third wave offers the chance to test further improvements of AI editing algorithms or to re-evaluate the audience reception to AI disclosure.

The here reported first wave questions were designed and conducted with 143 participants to investigate whether the tedious process of editing video news clips could be delegated to AI models and how AI editing technology is perceived by humans. In detail, the study measures subjective receptions of video news and survey participants´ evaluation of the quality of editing algorithms.

In Section 2, the algorithms utilised for video editing will be explained, and a previous experiment will be addressed. Section 3 delves into the survey protocol and video experiment while Section 4 illustrates the results. Finally, Section 5 discusses the survey findings and gives a brief outlook into potential future proceedings.

II. BACKGROUND

Editing video news clips comprises two main steps. In a first step, typically 10 to 15 most suitable scenes for a news story are selected from up to 200 scenes in raw footage. In the second step, the selected scenes are compiled into a video sequence, which is then accompanied by a voice over of a news text to create the final video news clip.

The core of the conducted first survey wave investigates how automated AI-based editing of video news clips scores against human news editing as a high anchor and a random editing as a low anchor. Hereby, two AI models are considered, which both are described in detail in [6]. The first AI model is the CLIP (Contrastive Language–Image Pre-training) model [7], pairing images of video shots with snippets of news text. The second AI model is denoted as KIGVI and has been developed by the RheinMain University of Applied Sciences [6]. The KIGVI model was trained using a dataset of 12354 video clips from news segments that ranged from 30 seconds to 5 minutes covering multiple categories of news from the years 2012 to 2025. The KIGVI model uses shot detection to split footage into scenes and follows learned professional video editing rules. These rules include varying shot sizes, an initial establishing shot illustrating the main topic and the inclusion of a human-like editing rhythm when composing the video news clip. This marks a major improvement over previous algorithms that failed to align information sourced in the visual component as well as the textual one.

A previous experiment with a small sample size (n = 38 participants) scored the KIGVI algorithm against human edits based on professionality of the edit, choice of scenes and how well the video sequences illustrated the voiceover. The findings of this experiment warranted further investigation with a bigger sample size and a larger survey that asks respondents about their opinion on AI, video news consumption habits and general media reception. Thus, a more comprehensive survey was developed.

For the analysis of the first survey wave, CLIP and KIGVI have been combined into the category AI and collectively compared to human video editing and random video editing. Human edited news clips were produced by professional editors. In random editing, scenes are randomly selected and compiled into a video news clip. All edited video news clips are voiced over by a professional German recording studio.

III. SURVEY PROTOCOL

A. Basic principle

The survey is designed as an online survey. The participants of the survey are asked to assess video news clips according to a list of criteria, and to identify the editor of the respective video clip.

B. Samples

The greater RheinMain area is used as survey area. A market research institute was tasked to recruit survey participants living in or around the RheinMain area aged from 16 to 70+ for a first wave of the survey. From this first wave, 150 interviews were conducted between December 2024 and January 2025. The responses of the participants are evaluated in this paper. Surveying the second half of the first wave will continue in the second half of April 2025.

The RheinMain area is a multi-state area in southern Germany including major cities like Frankfurt and Wiesbaden, mid-sized cities such as Russelsheim, as well as a number of smaller municipalities. Since the KIGVI algorithm might improve as it continues being trained on more and more video clips, a future second wave may account for the expected improvements. Completing the online survey took participants 12 minutes on average. The questionnaires were distributed via online links hosted on the platform Qualtrics. The sample size of validly completed questionnaires was 143, with seven interviews that had to be excluded due to incomplete answers and high amounts of item non-response.

C. Survey Structure

In an experiment portion of the questionnaire video news clips are shown, which cover three overall categories of news. These are:

- *Traffic news*
- *Local news*
- *News concerning politics*.

Each of these three categories of news consists of four video news clips, and each video news clip is edited in four different variants, which are:

- *Human editing*
- AI-based editing using the *KIGVI model*
- AI-based editing using the model based on *CLIP*
- *Random editing*.

This results in a pool of 48 video news clips (3 categories × 4 video news clips × 4 variants) as illustrated in Figure 1.

Prior to the experiment portion, the participants received a short introduction before facing a screen-out question that determines whether they agreed to data collection and were inhabitants of the greater RheinMain area. Succeeding led the participants being asked about their news consumption habits and preferences, how familiar they are with AI in general as well as attitudes towards the technology in general, accounting for both algorithmic and generative AI. The experiment portion was introduced to ensure participants understood the task of evaluating the algorithmic AI technology utilised in the editing process.

Matrix questions followed each editing variant in order to assess the quality of a video news clip. A matrix question is the judgement of a specific statement on the Likert-Scale.

Eight statements are used in the survey, which are shown in Table 1. They were inherited from a previous quality survey conducted in [6] with additional questions such as whether participants noticed technical errors like flickering, black screens, timing errors, etc. Furthermore, they were asked whether the video offered additional value to the news report or evoked emotions.

TABLE I. STATEMENTS BEING JUDGED BY A PARTICIPANT ON A LIKERT-SCALE TO ASSESS THE QUALITY OF A VIDEO NEWS CLIP.

Number	Statement
(1)	The voiceover matches the visual material
(2)	The scenes in the video illustrate all the important information in the program
(3)	The video looks professionally edited
(4)	The video is similar to video news I've seen before
(5)	Aspects of the report seem inconsistent / incorrect
(6)	I will probably remember the video report
(7)	The video makes the report more interesting to me
(8)	The report triggers emotion in me

D. Choice Experiment

Responses of the participants to the experiment portion as well as to general attitudes towards news, news reception, artificial intelligence and AI in news are measured on a six-point Likert-scale, which ranges from *Strongly Disagree* (1) to *Strongly Agree* (6). The experiment potion was randomised. A first randomisation assigned a category of news per participant while a second randomisation within each category selected one video news clip per editing variant meaning every participant was shown one random editing variant, one human editing variant, one AI editing variant using the KIGVI model and one AI editing variant using the model based on CLIP.

The video news clips were embedded in the experiment in a way that the participants were able to view them multiple times to secure the responses. Participants were tasked to evaluate the presented video news clip via the matrix questions. Following this evaluation, a choice was presented whether participants believed a human or an AI to be the editor of the before seen video news clip and how confident they were in their choice via Likert scale. Tracking the time, the participants spent on a specific page aimed to make sure that they at least viewed each video news clip once.

IV. RESULTS

In the sample of 143 individuals, 50.4% of survey participants reported identifying as female. A high number of participants, 22.3%, selected an age group of 41-50 years old, 15.8% selected 61-70 years old, 13.7% selected 31-35 years old, and only 10.8% selected ages of 22-25. Ages of 16-22 and 71+ were rare with under 5% of answers. Education levels culminated in 40.6% of participants reporting having passed their A-levels or owning a university degree of some kind while 25.4% completed vocational training.

A key finding of the conducted survey is that human editing and AI editing are rated almost identically on average. Both were ranked much higher than random editing.

When forced to assign a "human-edited" or "AI-edited" label to the different variants, a majority of the participants

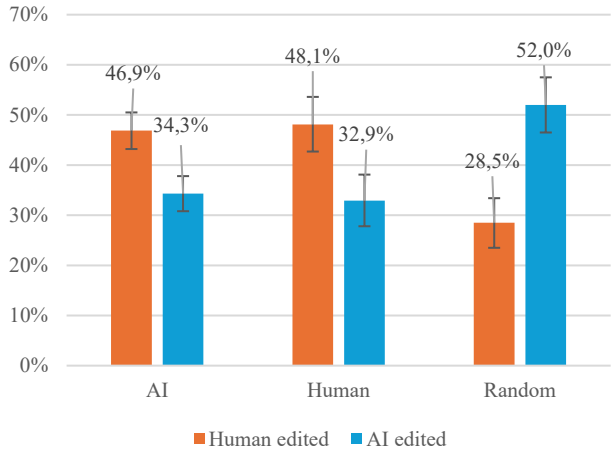

Fig 1. Percentage of survey participants voting either "human edited" (orange) or "AI edited" (blue) for each editing variant.

assigned the label "human-edited" to both, AI edited video news clips and human edited video news clips. Surprisingly, the human edited video news clips were assigned the "AI-edited" label too, as shown in Figure 1. All results shown are

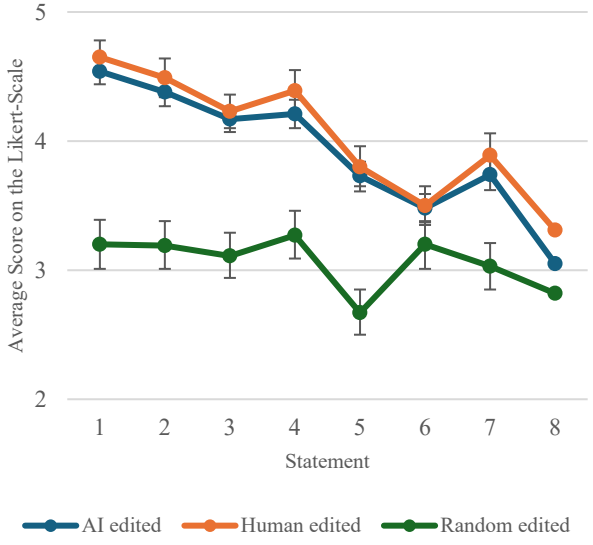

Fig. 2. Average scores on the Likert-Scale versus judged statement of Table 1.

depicted with a corresponding 95% confidence interval.

In Figure 2, the average Likert-Score is illustrated for each of the eight judged statements, (compare Table 1), alongside the corresponding 95% confidence intervals. The Likert-Scores were rounded towards whole values to allow for interpretation.

Overall, AI edited, and human edited video news clips delivered similar average scores on the Likert-scale, recording stronger agreement with the statements except for the inconsistency/incorrectness potential (5) whereas the randomly edited video news clips were scored lower in average for most statements. The more content-related statements are scored highest in average for AI edited and human edited video news clips. Especially the matching of the voiceover to the visual material (1) received highest scores in average, closely followed by the scene selection illustrating the categories of news well (2). Professionality of the edit (3) and resemblance to other news programs (4) also received high scores with AI editing and human editing performing almost identical in average. Inconsistencies / incorrectness (5) were scored much lower than random and met more disagreement. Memorability (6) was scored lower in average than other aspects with the smallest difference between all three editing variants. The human editing scores slightly higher in average when it comes to making the video interesting (7) and emotionality (8) was scored low across all three variants.

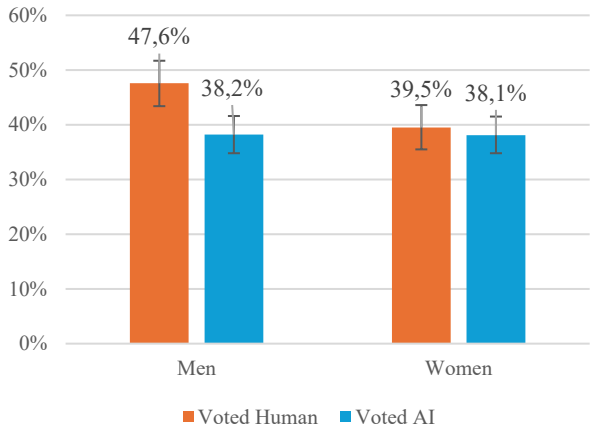

Fig. 3. Percentage of voting for “Human-edited” and “AI-edited” versus the gender of the participants.

Figure 3 illustrates the voting behaviour differentiated into women and men with a 95% confidence interval, showing a higher overall likeliness of men to vote for human-editing as the editor of the video news clip. For women, the choice between AI editing and human editing is much more evenly split. However, a high number of recipients who reported the usage of AI in media as positive assumed the creator of the videos as human rather than AI whereas all other stances had a more even distribution of creator assumptions. Since participants were given the option to cast a neutral vote in addition to the options “Human” or “AI”, Figures 1, 3 and 4 do not sum up to 100%.

Figure 4 details the attitude of the participants towards AI in media landscapes ranging from *positive* to *negative* on a four-point scale combined with the likeliness to identify the video editor as human or AI.

Generally, the attitude towards AI in media industries was evaluated as rather positive and rather negative much more

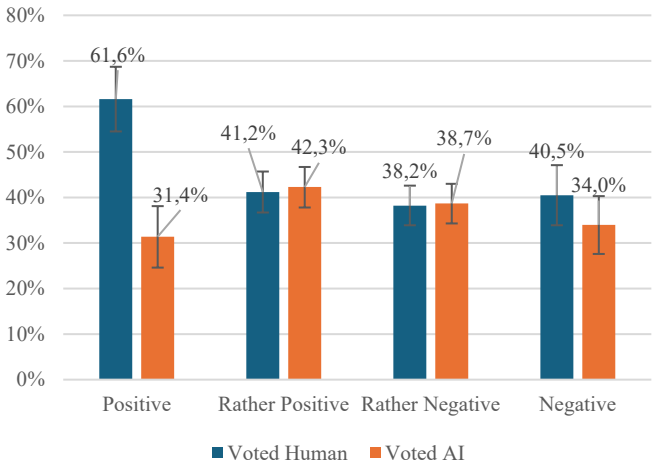

Fig. 4. Attitude of the survey participants towards AI usage in media spaces and voting behaviour.

than the definitive stances of positive or negative as shown in Table 2.

TABLE II. OPINIONS TOWARDS AI IN MEDIA IN PERCENTAGE.

Opinion towards AI	Number of Answers (%)
Positive	13,6
Rather Positive	34,6
Rather Negative	36,0
Negative	15,8

It is noted that Figures 4 and 5 display voting behaviour without accounting for whether the selected choice (human-edited or AI-edited) corresponds to the actual editor of the video news clips. Furthermore, Figure 5 shows that a negative attitude towards AI does not influence the rating of the participants supporting the evaluation results. However, a high number of recipients that reported the usage of AI in media as positive assumed the creator of the videos as human rather than AI whereas all other stances had a more even distribution of creator assumptions.

V. CONCLUSION

In this paper, a survey is designed and conducted with 143 recruited participants in order to evaluate the quality of video news clips edited by AI models in comparison to the one of news clips produced by professional human editors. A total of 48 video news clips is used in the study, containing three categories of news, four video news clips per category, and four editing variants of each video news clip. Two AI editing variants are applied. In addition, two anchor variants are used. The professionally human edited video news clips serve as a high anchor. Randomly edited video news clips serve as a low anchor. The participants were asked to judge eights statements on a Likert-Scale to assess the quality of a video news clip.

Overall, the evaluation shows that the quality difference of AI edited video news clips and human edited ones is statistically insignificant as confidence intervals of measured quality features overlap. AI editing is widely undetected by the participants.

The evaluation results point towards the conclusion that an AI can already perform the task of video editing as well as, or at least similarly to a professional editor in the industry. The lower evaluation of results for statements (6) to (8) of the question matrix question the relevance of these statements when it comes to the performance of the algorithms since the strictly technical evaluation generally scored higher provided a high separation of AI editing and Human editing versus random editing.

Additionally, the low scoring random edit could indicate faithful responses, as the random variant videos were lower in quality and included clips that did not fit the categories of news of the overall videos. Seeing this represented in the matrix questions allows to interpret, that users filled out the video experiment portion genuinely paying attention to the videos.

Reference [8] already found social media users to be incapable of distinguishing AI and human made content, in their case generative AI created images and Instagram captions. These findings align with our findings of recipients not being able to distinguish the AI from the human edit.

As of the present, results are based on a first wave of data of $n = 147$ participants. A second survey wave is currently in the field. For this second wave, the video experiment portion was pulled up in the survey flow to precede the questions about news habit and co. in an effort to potentially rule out effects of fatigue on the data. Furthermore, the two AI editing tools, the CLIP and the KIGVI algorithm, were combined in this paper to focus only on the performance of AI versus human. However, the evaluation in the surveys were separated into each algorithm to ensure the possibility of a separate assessment as well.

Additionally, the survey asks for user's socio-demographic characteristics, individual habits of news and video consumption, as well as views on AI. The according information allows deeper and group specific analysis in the future. For example, it will be possible to analyse whether bias towards AI exists, how it shows in the evaluation of the videos, and whether disclosing the use of AI influences audiences' perception of it.

VI. OUTLOOK

The present research gives a brief outline of key findings from a first survey wave. Data collection is ongoing and will allow for more in-depth results concerning the multitude of topics (news habits, attitudes towards AI) recorded in the survey later on. The present findings already highlight the improvements of AI editing in news broadcasting and pave the way towards bigger strides of the technology. As of now, statistical choice models determining the author of the clips are being estimated, aided by the aforementioned matrix with the addition of sociodemographic factors and attitudes towards AI. Employing a logistic regression model will inform about multiple effects and contributing factors affecting the recipient's assumed identification of the creator of the videos whether it be human or AI.

ACKNOWLEDGMENT

The authors thank the German Federal Ministry of Research, Technology and Space, and Qvest GmbH for funding this work via the program FH-Kooperativ 2-2019, contract number 13FH544KX9.

REFERENCES

[1] N. Newman, "Journalism, media and technology trends and predictions 2023," Reuters Institute for the Study of Journalism. https://doi.org/10.5287/bodleian:NokooZeEP [retrieved: 2025-06-05]

[2] Y. Hayashi, "Prospects for Revolutionary and Popular AI Technology following the Launch of ChatGPT in 2023," Electronics, 13(2), 290, pp. 1–2. https://doi.org/10.3390/electronics13020290 [retrieved: 05, 2025]

[3] J. O. Adigun et al., "Enhancing Academic Performance with AI-Powered Tools: A Comparative Study of Adobe Photoshop and Lightroom in Educational Technology Photography," Journal of Science Research and Reviews, 1(1), pp. 43–48. http://dx.doi.org/10.70882/josrar.2024.v1i1.9 [retrieved: 2025-06-05]

[4] Free online AI image Generator. (n.d.). Canva. https://www.canva.com/ai-image-generator/ [retrieved: 05, 2025]

[5] E. Theophilou et al., "AI and narrative scripts to educate adolescents about social media algorithms: insights about AI overdependence, trust and awareness," In European Conference on Technology Enhanced Learning, pp. 415–429. Cham: Springer Nature Switzerland. http://dx.doi.org/10.1007/978-3-031-42682-7_28 [retrieved: 2025-06-05]

[6] D. Quandt, P. Altmeyer, W. Ruppel, and M. Narroschke, "Automatic Text-based Clip Composition for Video News," In Proceedings of the 2024 9th International Conference on Multimedia and Image Processing (ICMIP '24). Association for Computing Machinery, New York, NY, USA, pp. 106–112. https://doi.org/10.1145/3665026.3665042 [retrieved: 2025-06-05]

[7] A. Radford et al., " Learning Transferable Visual Models from Natural Language Supervision," In Proceedings of the 38th International Conference on Machine Learning, in Proceedings of Machine Learning Research 139, pp. 8748–8763, Available from https://proceedings.mlr.press/v139/radford21a.html [retrieved: 05, 2025]

[8] J. Park, C. Oh, and H. Y. Kim, "AI vs. human-generated content and accounts on Instagram: User preferences, evaluations, and ethical considerations," Technology in Society, 79, 102705, pp. 10–11. http://dx.doi.org/10.1016/j.techsoc.2024.102705 [retrieved: 2025-06-05]