
Special issue

Jorge Caballero

<https://orcid.org/0009-0009-3535-8866>

jorge.caballero@upf.edu

Universitat Pompeu Fabra

Carles Sora-Domenjó

<https://orcid.org/0000-0003-2761-2384>

carles.sora@citm.upc.edu

Universitat Politècnica de Catalunya

Submitted

February 5th, 2024

Approved

March 20th, 2024

© 2024

Communication & Society

ISSN 0214-0039

E ISSN 2386-7876

www.communication-society.com

2024 – Vol. 37(3)

pp. 201-218

How to cite this article:

Caballero, J. & Sora-Domenjó, C. (2024).

Automation and creativity in AI-driven film

editing: the view from the professional

documentary sector. *Communication &*

Society, 37(3), 201-218.

<https://doi.org/10.15581/003.37.3.201-218>

Automation and creativity in AI-driven film editing: the view from the professional documentary sector

Abstract

This article explores the impact of integrating artificial intelligence (AI) into film editing, a field that has gradually adopted semi- and fully automated techniques, marking a turning point in film narrative and montage techniques. The study undertakes a systematic literature review and conducts fifteen semi-structured interviews with professional film editors in Spain to investigate how AI might transform workflows in film editing, in general, and documentary filmmaking, in particular. AI's potential as a complementary tool in the management and organization of large volumes of material is discussed, while growing interest is highlighted in the development of more interactive, more dynamic interfaces. Finally, the study stresses the need for future research to broaden its current scope, further exploring the practical applications of AI-driven film editing and addressing both the ethical and creative concerns associated with its use.

Keywords

Artificial intelligence, film editing, qualitative analysis, automation in editing, editing tools, audiovisual narrative.

1. Introduction

Film editing or montage is a key tool for articulating audiovisual narratives, on the one hand, and triggering emotional responses in viewers, on the other. Indeed, montage techniques, which embody influences from a broad spectrum of the arts –ranging from literature and theatre to painting and even music– have been widely studied in fields that include cognitive psychology and perception (Morante, 2017; Cutting, 2016; Smith, 2012).

The above definition of film editing adheres closely to Morante's (2017) dual approach: the first, operative in nature, involving the selection, cutting and montage of fragments of image and sound to create the final version of an audiovisual product and, the second, conceptual, involving the fixing of spatio-temporal relationships through the combination and duration

of scenes, to facilitate a fluid, coherent transmission of the audiovisual message to the viewer. In this respect, film montage plays a critical role in shaping the narrative and pacing of a film (Reisz, 1957) and, over the years, the focus has been fixed firmly on the achievement of such fundamental aspects as narrative continuity and connecting with the viewer, which shapes the emotional and psychological experience of the public (Murch, 2001).

At the intersection of cinema and the digital humanities (Heftberger, 2019; Chávez Heras, 2023), new perspectives for film montage have been opened up, above all with the systematization of extensive, complex databases (Rohrbach *et al.*, 2015; Winer *et al.*, 2021), and their analysis with different computational models. This new approach is reflected in such practices as “distant viewing” (Arnold & Tilton, 2019), which highlights visual interpretation in the analysis of large visual corpora, and in the framework of “cultural analytics” (Manovich, 2020), which highlights the need to compile and analyse large sets of cultural data for their systematic interpretation using quantitative and computational methods, and which is closely aligned with previous works of film analysis that have exploited methods of statistical analysis (Salt, 1974; Tsivian, 2009; Butler, 2014).

The confluence of these areas has given rise to what Ronfard (2017) calls “intelligent cinematography and editing,” combining as it does computational techniques and data analysis for audiovisual editing. These methods are enriching and transforming visual storytelling practices in the digital age; yet, despite being a recurrent concern in the field of academic research, their adoption among film editors and other professionals is less common, due to their limited practical applicability (Soe, 2021). An important distinction lies in the fact that montage professionals often lack programming skills, which limits their ability to incorporate computational methods into their work. Hence, there is a clear need to explore and apply computational techniques that might complement traditional editing practices. Indeed, adopting a combined approach could significantly improve film narratives and offer new creative possibilities (Zhang *et al.*, 2022).

The use of artificial intelligence (AI) models has the capacity to revolutionize film editing, especially with regards the efficient management of large amounts of visual material. Advanced computational tools allow editors to rapidly classify, sort and analyse images and videos, accurately identifying such elements as people and objects, and optimizing the extraction of metadata and the adaptation of video formats (Ferguson, 2017; Hielscher, 2020; Cooper *et al.*, 2021; Liu *et al.*, 2020; Ding *et al.*, 2021; Ding *et al.*, 2023). However, in addition to improving technical efficiency, AI has the potential to make film montage a more creative process. By facilitating the analysis and manipulation of audiovisual materials, AI can become an indispensable collaborator of film editors, opening up new creative avenues and allowing the exploration of innovative narratives. This technology not only simplifies tasks that might be deemed monotonous (Bieda & Panchenko, 2022), it also enriches the intuition and creativity of professionals by fusing algorithmic precision with human sensitivity. This collaborative approach between machine and creators can push the art of film editing towards new creative horizons.

Over the last thirty years, video editing has evolved towards the use of semi-automated or fully automated techniques. These techniques, which include the automatic detection of cuts and camera motion, as well as the identification of objects and people, facilitate editing by making audiovisual material more accessible and intuitive to handle (Ueda *et al.*, 1993). These innovations seek to optimize the editing process (Zhang *et al.*, 2022) and are especially useful for making documentaries. In this genre, editing is considered a complex creative and cognitive act that requires close collaboration between the editor, the director and the uncut filmed material. Pearlman (2018) describes it as the work of an “extended mind,” where artificial intelligence can be particularly beneficial, enhancing these cognitive abilities. The reason why these innovations may be more relevant to documentaries than fiction lies in the nature of the material with which the former work. Documentaries often involve handling

large volumes of raw, unstructured material, where finding the desired narrative can be more complex.

Despite significant advances in automation and the potential that AI offers for video editing, their impact on the creative industries, particularly cinema, has yet to achieve the levels reached in other areas (Chow, 2020). While in the business and commercial sphere AI has shown notable progress, in cinema, especially in the production of documentaries where its application could be revolutionary, its adoption has been somewhat more cautious. However, the growth recorded by on-demand video streaming platforms and the use of artificially intelligent personalization algorithms are pushing film studios to review and possibly adopt new methods in their processes of production and distribution (Chow, 2020).

The incorporation of AI in film production promises not only the possibility of innovating creatively and technically, but also of significantly increasing box office revenues (Singh *et al.*, 2023). Such progress is a sign of the industry's interest in exploring how AI can contribute, for example, to the improvement of visual effects, in its efforts to increase the quality and commercial impact of productions. However, as we move towards this integration, we should not lose sight of the limitations and ethical dilemmas involved in using AI in filmmaking, especially if we bear in mind its far-reaching potential impact on genres such as documentary, where it could enrich the narrative and editing in hitherto unimaginable ways.

Films are not simply a source of entertainment but can serve as important educational and communication tools that significantly impact our visual-spatial perception and cognition, surpassing, in this regard, the influence of texts (Bordwell, 2015). This impact highlights the relevance of understanding just how we interpret visual experiences and how AI might enrich the creation of complex, engaging narratives (Suchan & Bhatt, 2016). Indeed, AI can emerge as a potential collaborator in film production, capable of capturing an audience's attention and of interacting with it at deeper, more meaningful levels.

This article investigates the potential of AI to transform the workflows of film editing. To do so, it explores the work of Spanish film editors by conducting and reporting 15 semi-structured interviews. The study focuses on the methods they employ and on their needs, and with this information explores how AI could influence their creativity and autonomy, with a particular concern for its impact on documentary film. The rest of the article is structured as follows. Following on from this introduction, a review is undertaken of the literature dedicated to the use of AI in video editing and a description is provided of the methodology used herein to analyse the interview data collected. The conclusions seek to highlight the most significant findings and propose paths for future research, while reflecting on how AI has the potential to significantly transform film editing.

2. A literature review: the background

AI applications in video editing have increased significantly, driven by ease of access to massive volumes of visual content (Dancyger, 2014). Initially used to generate content for social platforms, AI is today being employed for more advanced tasks, including film editing. In this latter field, it allows editors to both analyse and enrich narrative elements with a much greater degree of sophistication and precision. Advances in areas such as computer vision, machine learning and deep learning have spurred research into, and the development of, specific editing tools, including those of video segmentation, intelligent clip manipulation, and the creation of advanced transitions (Bieda & Panchenko, 2022).

Several innovative solutions, focused specifically on narrative and text, have been developed to improve video editing. These include technologies for analysing scripts and dialogues, constructing narrative event chains through the application of video semantics and theme representation (Nack & Parkes, 1997), ordering narrative event chains (Chambers & Jurafsky, 2008), applying the rules of video grammar to extracted metadata to connect shots (Kumano *et al.*, 2002), and facilitating automatic, collaborative editing using audio transcripts

or subtitles (Outtagarts & Mbodj, 2012). Other solutions of note include those techniques to segment long videos into a chapter structure using automated transcripts (Pavel *et al.*, 2014), analyse relationships between narrative events (Hu *et al.*, 2017), generate video summaries from voice transcripts (Taskir *et al.*, 2006), automate editing through dialogues and scripts (Leake *et al.*, 2017), and use deep learning tools to edit videos as one might a text (Barakat *et al.*, 2021).

Research in this field has evolved over decades, expanding its initial focus from textual analysis to include object and person recognition, camera organization and shot and scene identification. Innovations such the automatic structure visualization for video editing (Ueda *et al.*, 1993), and improvements in intelligent interaction with semantic zoom and multiple lens techniques (Casares *et al.*, 2002; Long *et al.*, 2002), along with automated clip organization (Girgensohn *et al.*, 2001), have contributed significantly to the development of the field.

With the arrival of AI, advanced video-editing solutions have been developed, including segment annotation and editing based on narrative events, actions and the characters' emotions (Lombardo & Damiano, 2010), and an intelligent film assistant for mobile devices that offers expert knowledge on scene composition and camera motion using tailored story boards (Schörkhuber *et al.*, 2017). Recent innovations include automatic video editing methods that use neural networks to extract visual semantic and aesthetic features to create brief, coherent visual narratives (Podlesnyy, 2019), attention models to automatically edit recordings of meetings in corporate environments (Wu *et al.*, 2020), technologies to generate talking faces in videos (Gupta *et al.*, 2021), and the use of AI to automatically focus on and edit specific faces in videos (Xie & Sun, 2022).

Following the AI-powered transformation of this field, the concept of database cinema, as introduced by Manovich (2000), represents a natural evolution and a step forward. This concept highlights how these technologies allow multimedia data to be manipulated and organized in an innovative way, offering new aesthetic and narrative possibilities by using algorithmic editing. This technique distances itself from traditional editing approaches and encourages the exploration of visual narratives that rely on computational capacity to re-define story creation.

The importance of databases has become crucial in the era of neural networks, an aspect that has been explored in the research, for example, of Matsuo *et al.* (2002) and Soe (2021). Their work demonstrates the significant impact of data mining on video editing, identifying patterns and rules that facilitate more efficient automated editing systems. Such an approach not only optimizes processes of film montage, but also paves the way for the customization of AI-driven editing tools, thus responding to the specific needs of the sector's professionals.

These advances herald a future in which video editing moves towards the integration of traditional methods and the latest AI-powered technologies (Zhang *et al.*, 2022). The fusion between cinematic tradition and technological innovation reflects the field's unceasing adaptation to the new possibilities that AI and data analysis present.

Within video editing, AI-driven innovations can redefine not only film editing techniques, but also the understanding and generation of content. This paradigm shift is evident in apps such as Project Blink (Adobe Labs, 2023), which illustrates how the application of AI models on online platforms can radically transform the editing process, by allowing video manipulations via the textual editing of its transcripts (Adobe Research, 2023). This technology facilitates the automatic detection of objects and sounds, as well as the classification of narrators, providing editors with powerful tools to organize and edit content more intuitively and fluidly.

In this context, the ability to search through and classify footage efficiently becomes a critical component in the montage process. Traditionally, this work required the labour-intensive manual viewing and selection by the editing team. However, with the advent of AI, innovative methods are emerging that allow the automatic identification of highlights, based on pre-trained models. An example of this is the video highlights via photogenic moments

generation project (Lin *et al.*, 2022a), which uses a convolutional neural network (CNN) and CLIP network semantic coding (Radford *et al.*, 2021) to discern and highlight significant visual elements based on comparisons with professional image databases, such as Adobe Stock.

Similarly, Netflix has implemented machine learning models to automate the search for compatible sequences, facilitating the montage of trailers and promotional videos by identifying shots that complement each other in terms of composition and framing (Chen *et al.*, 2022). Such advances do not seek to replace human creativity, but rather to provide tools that can improve film editing efficiency, allowing editors to focus on the more creative aspects of montage.

Deep learning pushes back the frontiers of audiovisual editing, especially as regards content exploration and organization. Traditionally restricted to linear and chronological structures, the introduction of “VideoMap: Video Editing in Latent Space” (Lin *et al.*, 2022b) represents a paradigm shift towards a more flexible, multidimensional approach. This method allows editors to manipulate and organize content in a latent space, where frames are encoded using deep neural networks, facilitating innovative visual exploration via the city metaphors of nodes, paths, districts and landmarks. This approach not only breaks with traditional temporal linearity, but also promotes discoveries of new narrative associations for video montage.

AI-powered film editing promises to revolutionize the way we interact with visual material via more flexible, intuitive interfaces. By simplifying the organization, editing, and categorization of content, these technologies herald the expansion of creative possibilities, furnishing editors with new tools to overcome semantic limitations and discover new narrative horizons. This progress cannot only improve workflows, but it can also trigger significant innovation, opening up paths towards more intuitive editing tailored to the specific needs of each project.

In documentary making, above all, these innovations constitute exceptional tools for managing the complexity of nonfiction narratives. The organization of massive volumes of footage and the ability of neural networks to discern and organize narrative sequences significantly enhance this genre. This allows documentary filmmakers to address more effectively the unique challenges of narrating authentic stories, enriching them with new layers of narrative depth and creative exploration.

3. Methodology

This study focuses its attention on determining the possibilities AI has to offer film editing. In so doing, it first identifies the methods and needs of editors as revealed through semi-structured interviews. This flexible technique allows us to capture the experiences, visions and expectations of professionals, thereby laying the foundations to explore how AI might satisfy their specific needs and enrich their practice.

3.1. Sample and profile of interviewees

For the purposes of this study, film editors were chosen based on their experience and geographical location. A gender balance was also ensured. Participants must have worked on at least two feature films, including fiction, non-fiction and experimental, and received recognition in festivals or had their work shown in commercial cinemas. Experience with AI or automation was not considered essential, the focus being on their editing techniques rather than on their use of advanced technology.

Table 1. Profile of the 15 film editors interviewed in this study.

Gender	Professional experience (in years)	Genres of specialization	Awards	Other roles
Male	15+	Documentary, Fiction	Emmy and Gaudí Prizes	Direction, Screenplay
Female	15+	Documentary, TV, Fiction	International Film Festival Prizes	Direction, Teaching
Female	20+	Documentary, TV, Fiction	International Film Festival Prizes	Teaching, Direction
Male	10+	Documentary, Fiction	International Film Festival Prizes	Direction, Screenplay
Male	20+	Documentary, Fiction, Advertising	International Film Festival Prizes	Direction Creative
Male	15+	Documentary, Fiction	International Film Festival Prizes	Direction, Photography
Female	25+	Documentary, Fiction	International Film Festival Prizes	Teaching
Male	15+	Documentary	International Film Festival Prizes	Curation, Writing
Female	20+	Documentary, Experimental	International Film Festival Prizes	Screenplay
Male	25+	Documentary, Fiction	International Film Festival Prizes, Goya Prize	Direction, Teaching
Female	10+	Documentary, Fiction	International Film Festival Prizes, Goya and Gaudí Prizes	Teaching
Female	15+	Documentary, TV, Fiction	International Film Festival Prizes	Management
Female	10+	Documentary, Fiction	International Film Festival Prizes	Direction, Writing
Female	10+	Documentary, Fiction	International Film Festival Prizes	Teaching
Male	15+	Documentary, TV, Fiction	International Film Festival Prizes	Teaching

Source: Own elaboration.

Fifteen film editors were selected for interview. These were conducted in a controlled environment, prioritizing the participants' comfort and privacy. Each session was recorded and transcribed for subsequent analysis. To examine the data collected, various content analysis techniques were applied –including textual and thematic analyses– to identify structures and interpret themes in the interviews, in an effort at determining the editors' common concerns and experiences. Throughout the whole process, interviewee confidentiality was maintained and informed consent was obtained from the participants, the data being used exclusively for academic and research purposes while respecting the privacy and rights of all the editors involved.

Interview design comprised three phases: 1) introductory questions to collect demographic and professional data; 2) semi-structured questions focused on the detailed breakdown of montage practices, with a particular emphasis on the organization of materials, the stages and substages of the editing work, and the structure of the teams in each of these phases; and 3) final questions aimed at recording any additional opinions or comments the participants might want to make. This design ensured a comprehensive coverage of the editors' practices and preferences.

Having completed the interviews, and as an additional element of this study's exploratory approach, it was decided to automate the content analysis using AI algorithms, a technology that has established itself as a valuable resource in qualitative research, supporting literature reviews, and thematic and content analyses (Christou, 2023) and helping qualitative researchers obtain complex insights from large data volumes. Advanced AI tools, including such models as GPT, have been successfully incorporated into these methodologies, improving the efficiency of the research process (Şen *et al.*, 2023), and have opened up avenues for fruitful debates on future research strategies, particularly as regards collaboration between humans and machines. Other large language models (LLMs), such as BARD, have been used to explore the feasibility of AI in qualitative analyses and essential guidelines have been drawn up for their effective use (Dengel *et al.*, 2023).

The findings indicate that LLMs can be useful exploratory tools for qualitative research, offering diverse perspectives and facilitating the identification of patterns and themes in large volumes of data. However, their unpredictability may limit their applications in formal research. Despite their advantages in terms of efficiency and analytical capacity, their use also raises concerns about bias and ethical issues, areas in which the researcher's skill in the interpretation and ethical use of results takes precedence (Christou, 2023).

3.2. Methodological experimental design

Here, computational tools were used in the qualitative analysis of the interviews, helping in evaluating the automatic transcription and in the semantic analysis. Adopting an experimental approach, AI and natural language processing techniques were adapted, adjusting the methodology to the needs and findings. Constant iteration enabled a synergy to be developed between precision, automation and human oversight, providing a better qualitative data analysis than that achieved using conventional methods.

The 15 interviews were transcribed using the automated app, Trint. A manual coding process was then initiated, but a tendency to simplification and loss of nuances was quickly observed, leading to generalized and ambiguous concepts. To address these limitations, the OpenAI GPT-4 model was implemented in an effort to improve the identification and analysis of concepts by means of knowledge graphs (Fensel *et al.*, 2020) that connect entities with nodes and edges. However, this approach proved excessively detailed, impeding the synthesis of general themes. An alternative strategy was therefore chosen applying the same model, using a specific prompt to extract key concepts from the segments, presented in JSON format (see attachment in Annex 1). This technique captured data complexity and richness more effectively, while maintaining levels of efficiency and overcoming the restrictions evident in the initial manual coding.

Following the automated transcription of the interviews and a preliminary analysis with GPT-4, the concepts identified were refined and examined. They were then vectorized using SpaCy, a natural language processing library, transforming them into numerical vectors for deep semantic analysis. These vector representations, or embeddings, facilitated the clustering of concepts according to their density in the vector space using the HDBSCAN algorithm, a noise-robust method that proved to be effective in organizing concepts into relevant categories (McInnes *et al.*, 2017; Tran *et al.*, 2013).

However, the accuracy of the initial results did not meet expectations, and so higher-dimensional embeddings were incorporated using OpenAI's text-embedding-ada-002 model. This integration provided considerable progress to be made in the detection of semantic nuances. The use of the K-Means algorithm and the cosine similarity measure improved the classification of concepts, further enriching the identification of semantic connections.

To further refine precision, a manual list of key concepts was selected after extensive examination, iteratively adjusting it to highlight relevant themes and contribute to the development of a semantic search engine. A similarity threshold was defined to include only those closely linked concepts in the visualization. This strategy provided a clear and precise visual illustration, revealing patterns and links in the data for a more complete analysis. In this way, an ideal balance was struck between the preservation of the original concepts and the automatic detection with AI of significant links, highlighting its value for exploring and visualizing semantic relationships in texts.

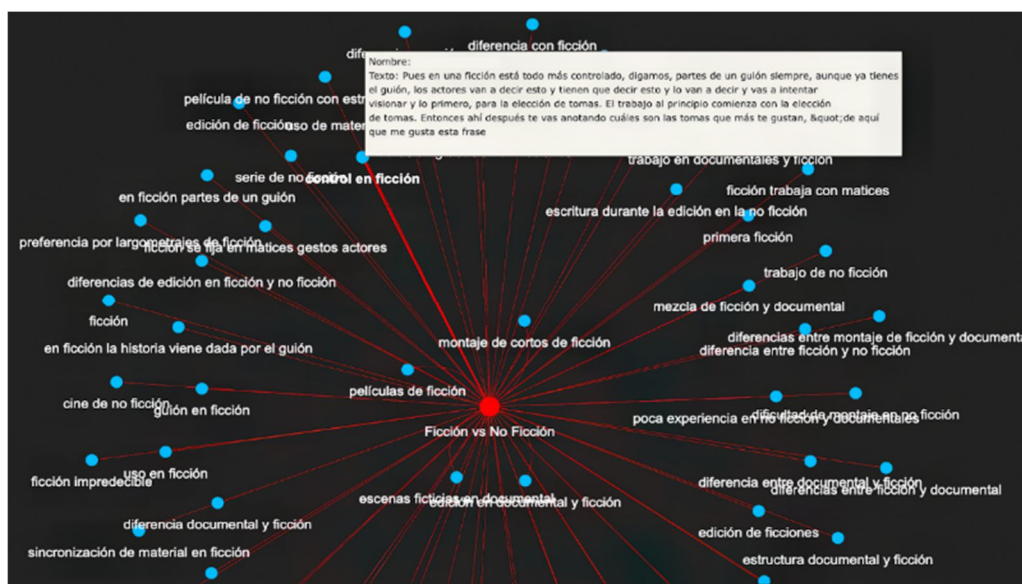
4. Results

4.1. Differences between editing fiction and non-fiction films

The first result to emerge from using automated AI-powered techniques to analyse the interview data was the significant differences found between editing fiction and non-fiction, particularly as far as using a script is concerned. This outcome is visualized in Figure 1 and is critical for the development of subsequent editing approaches using AI.

Figure 1. Visualization of the concept: Fiction vs. Non-fiction. Extract from one of the interviews.

Name: Text: Well, with fiction everything is more controlled, I mean, the starting point is always the script. And although you already have the script, the actors are going to say whatever and they say it, and you're going to try to visualize this first, so as to choose the takes. The job at first begins with deciding on the takes. Then you note down which takes you like best, " and from here I like this phrase.



Source: Own elaboration.

In the montage of fiction films, the script is fundamental, serving to guide the narrative structure and the filming logic. In this case, the editing process is a daily, sequential montage, responding to what was filmed each day, enabling an editor to concentrate on such details as

the actors' gestures and other subtle features in the footage. Although this methodology is less flexible structurally, it ensures complete immersion in the narrative and visual details of fiction cinema.

In contrast, the editing of non-fiction or documentaries usually commences once filming is complete or nearly complete. Such a practice results in the loss of certain opportunities that fiction editing offers, a point highlighted in various comments –see Figure 2 – made by the editors. This differentiation in editing practices highlights unique challenges for the application of AI and automation in the editing of non-fiction films.

Figure 2. Visualization of the concept: On finishing shooting. Extract from one of the interviews.

Name: Text: For two reasons. One, because it's extremely frustrating to become aware of things after the event, things you can no longer intervene in. That is, as I see it, editing while you're shooting is somehow like having a safety net, but with reason, it's like saying, I've edited yesterday's footage. And I think it'll be good; and it's not just for things that went wrong, as we say, no, it's for correcting bad things and implementing good things. "listen this character is working brilliantly, perhaps we should give him a bit more time"



Source: Own elaboration.

4.2. Editing and narrative in non-fiction

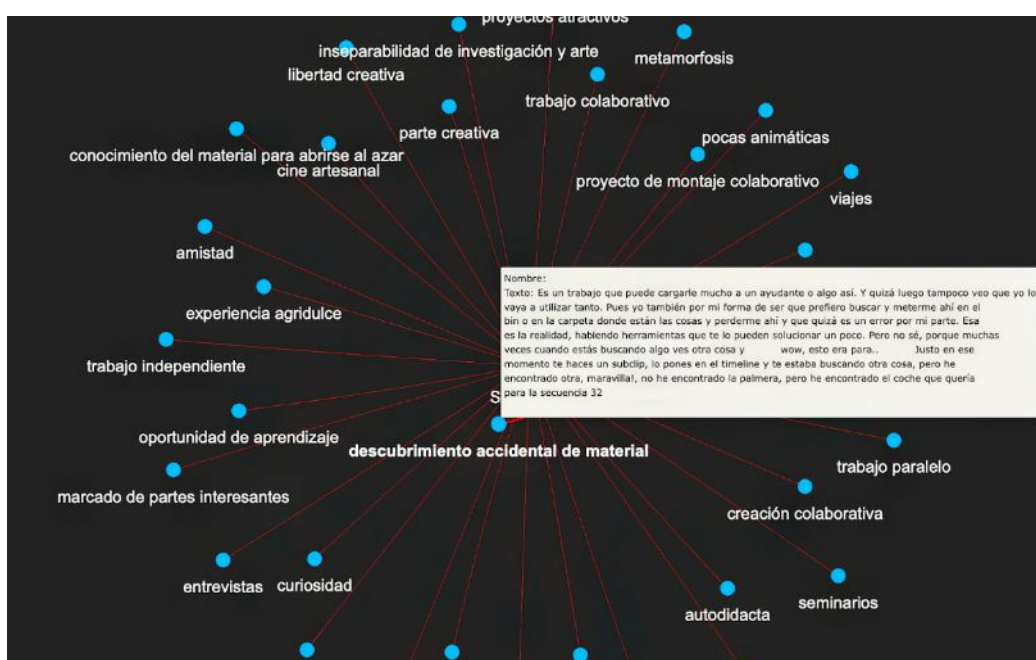
For this reason, in non-fiction film editing, the construction of the narrative acquires particular importance during the editing stage precisely because of the absence of a predefined script. In this way, editors enjoy greater narrative and structural flexibility. Indeed, editors often describe this phase as a deep immersion in their footage and as a time when they are open to emerging themes and hidden narratives in their material, as highlighted in Figure 3. This flexibility contrasts with the more predetermined structure of fiction cinema and suggests that certain AI-automated processes could significantly assist in this work.

Yet, some editors do not see such a marked distinction between editing fiction and non-fiction, considering film a creative continuum. This view is more common among editors of more experimental (or *d'auteur*) fiction projects. In contrast, those who work on fiction series

or productions for a broader audience tend to draw a more marked distinction in their work methods. Additionally, in non-fiction projects based on archival material or interviews, editors emphasize the importance of becoming very familiar with the content that has to be woven to form the narrative. Repeated reviewing of the footage is deemed essential to internalize the content and to make unexpected findings; in short, it is considered a key process in editing non-fiction in order that salient narrative elements might be discovered.

Figure 3. Visualization of the concept: Deep immersion in the footage. Extract from one of the interviews.

Name: Text: It's a job that can be too much for an assistant or the like. And perhaps later, I also see that it's not something I'm going to have much use for. And, because of the way I am, I like to rummage around in the bin or folder where the things are and lose myself in it, and perhaps I'm making a mistake. This is how it is, having tools that can help you out a bit. But I don't know, because often when you're looking for something you come across something else and... wow, that was for.... So there and then you make a subclip, you put it on the timeline but you'd been looking for something else, but I've found another thing, great! I've not found the palm tree, but I've found the car I wanted for sequence 32, bloody great.



Source: Own elaboration.

4.3. Editing tools and techniques

The professional editors interviewed stated a preference for using analogue tools such as post-its, frame printouts, whiteboards and manual notetaking while viewing footage and in prior phases dedicated to organizing that footage. These techniques permit the physical handling of materials and a more focused visualization, reflecting their preferences for a more direct, sensory connection with their materials. Additionally, the use of printed cards to physically visualize the structure of the film is common, especially in advanced stages of the montage. Taking notes in notebooks is another widespread practice, while very few of the interviewees mentioned using digital notes or digital data tables.

For the classification and organization of material, each editor appears to have developed their own system, which varies according to personal preferences and the needs of the project. Some opt for a chronological structure, others for daily sequences, themes or characters. Within these categories, they often create subfolders with more detailed classifications, which are essential for managing large volumes of material and for quickly accessing specific moments. In non-fiction projects with large numbers of folders, editors use a system of labels and keywords to facilitate organization. However, none of the professionals had used automated classification or labelling systems in their projects.

In the editing of non-fiction films, the technique of placing all the raw material on a timeline and then cutting it down predominates, in a process that is compared by some to sculpting in stone or modeling in clay. This technique reflects the artisanal, creative nature of film montage, where editors eliminate or incorporate material to forge the final narrative. Additionally, it is common for them to create subclips from longer master clips for more efficient management.

The editors interviewed primarily rely on four computer applications: namely, Avid, Premiere, Final Cut and DaVinci Resolve, each valued for its own specific characteristics or features. While largely satisfied with these tools, many expressed the need to evolve towards more intuitive, more dynamic interfaces, and suggested the use of touch screens and voice controls that could incorporate AI systems for a more immersive approach and a faster, more efficient working method.

4.4. *Automatization and views on AI*

Regarding the adoption of automated technologies for organizing cinematographic material, the professional editors expressed a remarkable degree of openness to such innovations, albeit that many admitted to not having actively sought out tools of this nature. And, despite the potential of these technologies to meticulously capture the elements of a scene, the editors were on the whole somewhat reticent to employ them, their chief argument being the importance they attach to reviewing their material themselves, repeatedly and in depth. All in all, there was a general distrust of just how effective automation could be in this regard. The editors held firm to the idea that the powers of human understanding and discernment outstrip those of any algorithm, and thus highlighted the value of personal interpretation and intuitive knowledge in their work.

A small number of the interviewees have begun to integrate advanced digital tools, such as Scenarist and Miro, for a more dynamic and accessible organization of their material. These tools replace traditional methods –the case of digital canvases supplanting whiteboards– facilitating a more efficient, adaptable organization.

At the heart of the field of film editing, a fundamental dichotomy is evident in relation to the integration of AI. Thus, while there is a certain disposition towards automated and semi-automated tools, at the same time there is also some reluctance to employ them, fuelled by the fear that automation will put human editors out of work. The tension that prevails is not contradictory, but rather reflects the complexity of incorporating new technologies in an area traditionally dominated by human skill.

Contrasting with the potential of AI, the editors express a marked preference for analogue tools and place considerable reliance on mnemonic devices. Editors rely on repetition and knowledge of their content for identifying valuable material. This trust in memory and intuition is crucial in a process where the ability to remember and connect different pieces of footage is essential in creating the narrative. The ability editors have to handle and remember content allows them to build a coherent, attractive story, underlining the importance of a deep, personal connection with their material.

One concern expressed was the risk of hyper classification that the use of automated tools might represent (see Figure 4). An excessive degree of organization could result in a

Against this backdrop, the reception of AI by professional editors is characterized by a duality of views. On the one hand, the promise of a more in-depth analysis and more efficient management of large volumes of data offers undeniable advantages, particularly in contexts in which the material is especially extensive. This analytical capability of AI promises not only to simplify the organization of material but also to reveal complex narrative connections and patterns, opening the door to previously unexplored creative approaches.

However, on the other hand, this positive view is contrasted with legitimate concerns about the possible erosion of the artistic essence of film montage. Intuition, sensitivity and a personal and emotional connection with the material are all qualities that it is feared might be diluted by the mechanical approach taken by algorithms. Yet, contrary to the perception that technology might supplant human creativity, the adoption of AI can also be envisioned as a collaborative tool in which creative capabilities are expanded without sacrificing human and artistic depth and authenticity.

The debate about the integration of AI has been extended to consider major ethical challenges, including those concerning AI bias and the implications for intellectual property and creative authenticity. Developing regulatory frameworks is therefore imperative to ensure that AI-powered film editing respects the fundamental values of cinema as an art form.

The analyses of the literature and the interviews undertaken here reveal an evolution in the development of innovative methodologies that include text-to-video editing and the automated organization of footage, all of which signals progress towards a more holistic understanding of audiovisual content. This evolution reflects a change from a focus on isolated tasks towards an integration that seeks to understand the material as a whole, responding to the need for tools that not only classify and organize but that also interpret and give new meaning to vast volumes of data.

Personalization has emerged as a central theme, evidencing a movement towards the adaptation of both content for specific audiences and editing interfaces to satisfy the individual preferences of each professional editor. This approach to customization promises to transform not only the efficiency of the editing process but also how cinematic narrative is conceived and experienced.

Throughout this discussion, it has become evident that the integration of AI into film editing should not be seen as a dilemma between technological innovation and artistic tradition, but rather as an opportunity to explore how technology can act as a catalyst for new forms of creative expression. Collaboration between film editors, AI developers and other professionals in the field is key to broadening the potential of these tools, heralding a future in which technology and human creativity combine in a way that enriches film editing.

6. Conclusions

This article has sought to provide a vision of contemporary film editing, highlighting the complexity and diversity of its practices. The qualitative analysis of the responses obtained from interviews conducted with 15 montage professionals has led to the identification of marked differences in the methods employed by editors depending on whether they are engaged in fiction or non-fiction projects. This is particularly evident as regards the organization, classification and handling of footage, with a clear preference being revealed for analogue methods and a marked dependence on mnemonic skills in the creative process. Despite their claims of being open to the adoption of automated technologies, the editors revealed a strong inclination for traditional approaches, which can be interpreted as evidence of the tension that persists between technological innovation and the preservation of established practices.

The study conducted here identifies AI's potential to be a complementary tool, especially in the management and organization of large volumes of footage, above all in non-fiction projects. Likewise, it reveals a growing interest in the development of more interactive, more

dynamic interfaces. This points to the future integration of advanced technologies that can improve workflows without compromising the artistic integrity of the montage.

This study, however, is not without its limitations. The sample of interviewees, although diverse, focuses on a limited number of professionals, and this could restrict any generalizations of the findings to the film industry as a whole. Furthermore, the research relies heavily on subjective perceptions and experiences, and hence having to interpret the film editors' personal opinions, without the possibility of undertaking an empirical analysis of the practical implementations of AI-powered montage.

These limitations highlight the need for future studies that can expand the scope of this current research. This would include working with a larger sample of professional editors and undertaking analyses of practical instances of AI integration in film projects. In this way, the theoretical and practical applications of AI in editing could be explored in greater depth and, importantly, the creative and ethical concerns associated with its use could be addressed.

In short, the study reported here suggests a path towards the development of AI tools designed specifically for film editing, tools that are intuitive and tailored to the individual needs of each project.

References

- Adobe Labs. *Project Blink*. Adobe Labs. Retrieved from <https://labs.adobe.com/projects/blink/>
- Adobe Research (2023) *Project blink: Creating the future of AI-powered video editing*, Adobe Research. Retrieved from <https://research.adobe.com/news/project-blink-creating-the-future-of-ai-powered-video-editing/>
- Arnold, T. & Tilton, L. (2019). Distant viewing: analyzing large visual corpora. *Digital Scholarship in the Humanities*, 34(1), i3–i16. <https://www.doi.org/10.1093/lc/fqz013>
- Barakat, M., Magdy, N., Elsisy, A., Youssef, R., Elattar, M. A. & Darweesh, M. S. (2021). Innovative deep learning-based video editing tool. *2021 17th International Computer Engineering Conference (ICENCO)*. <https://www.doi.org/10.1109/icenco49852.2021.9698919>
- Bieda, I. & Panchenko, T. (2022). A Systematic Mapping Study on Artificial Intelligence Tools Used in Video Editing. *IJCSNS International Journal of Computer Science and Network Security*, 22. <https://www.doi.org/10.22937/IJCSNS.2022.22.3.40>
- Bordwell, D. (2015). *Narration in the fiction film*. Oxford: Routledge, Taylor & Francis Group.
- Butler, J. (2014). Statistical analysis of television style: What can numbers tell us about TV editing? *Cinema Journal*, 54(1), 25–44. <https://www.doi.org/10.1353/cj.2014.0066>
- Casares, J., Long, A. C., Myers, B., Stevens, S. & Corbett, A. (2002). Simplifying Video Editing with Intelligent Interaction. *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. <https://www.doi.org/10.1145/506443.506539>
- Chambers, N. & Jurafsky, D. (2008). Unsupervised Learning of Narrative Event Chains. *Proceedings of ACL-08: HLT*, 789–797.
- Chow, P.-S. (2020). Ghost in the (Hollywood) machine: Emergent applications of artificial intelligence in the film industry. *NECSUS. European Journal of Media Studies*, 9(1), 193–214. <https://www.doi.org/10.25969/mediarep/14307>
- Christou, P. (2023). How to use Artificial Intelligence (AI) as a resource, methodological and analysis tool in qualitative research? *The Qualitative Report*. <https://www.doi.org/10.46743/2160-3715/2023.6406>
- Chávez Heras, D. (2023). *Creanalytics*: Automating the supercut as a form of critical technical practice. *Convergence: The International Journal of Research into New Media Technologies*, 135485652311745. <https://www.doi.org/10.1177/13548565231174592>
- Chen, B., Ziai, A., Tucker, R. & Xie, Y. (2022). “Match Cutting: Finding Cuts with Smooth Visual Transitions. *arXiv.org*, October 11, 2022. <https://www.doi.org/10.48550/arXiv.2210.05766>
- Cooper, A., Nascimento, F. & Francis, D. (2021). Exploring Film Language with a Digital Analysis Tool: the Case of Kinolab. *DHQ: Digital Humanities Quarterly*, 15(1).

- Cutting, J. E. (2016). The evolution of pace in popular movies. *Cognitive Research: Principles and Implications*, 1(1). <https://www.doi.org/10.1186/s41235-016-0029-0>
- Dancyger, K. (2014). *The Technique of Film and Video Editing*.
<https://www.doi.org/10.4324/9780240813967>
- Dengel, A., Gehrlein, R., Fernes, D., Görlich, S., Maurer, J., Pham, H. H., Großmann, G. & Eisermann, N. D.g. (2023). Qualitative research methods for large language models: Conducting semi-structured interviews with ChatGPT and BARD on computer science education. *Informatics*, 10(4), 78. <https://www.doi.org/10.3390/informatics10040078>
- Ding, D., Ma, Z., Chen, D., Chen, Q., Liu, Z. & Zhu, F. (2021). Advances in video compression system using Deep Neural Network: A review and case studies. *Proceedings of the IEEE*, 109(9), 1494–1520. <https://www.doi.org/10.1109/jproc.2021.3059994>
- Ding, D., Wang, J., Zhen, G., Mukherjee, D., Joshi, U. & Ma, Z. (2023). Neural adaptive loop filtering for video coding: Exploring multi-hypothesis sample refinement. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(10), 6057–6071.
<https://www.doi.org/10.1109/tcsvt.2023.3260266>
- Fensel, D., Šims ek, U., Angele, K., Huaman, E., Kärle, E., Panasiuk, O., ... & Wahler, A. (2020). Introduction. In *Knowledge graphs methodology, tools and selected use cases* (pp. 1–10). Cham: Springer.
- Ferguson, K. L. (2017). Digital Surrealism: Visualizing Walt Disney Animation Studios. *DHQ: Digital Humanities Quarterly*, 11(1).
- Girgensohn, A., Bly, S. A., Shipman, F. M., Boreczky, J. S. & Wilcox, L. (2001). *Home Video Editing Made Easy - Balancing Automation and User Control*. IFIP TC13 International Conference on Human-Computer Interaction.
- Gupta, A., Khan, F. F., Mukhopadhyay, R., Namboodiri, V. P. & Jawahar, C. V. (2021). Intelligent video editing. *Proceedings of the 12th Indian Conference on Computer Vision, Graphics and Image Processing*. <https://www.doi.org/10.1145/3490035.3490284>
- Heftberger, A. (2019). *Digital humanities and film studies visualising Dziga Vertov's work*. Cham: Springer.
- Hielscher, E. (2020). The Phenomenon of Interwar City Symphonies: A Combined Methodology of Digital Tools and Traditional Film Analysis Methods to Study Visual Motifs and Structural Patterns of Experimental-Documentary City Films. *DHQ: Digital Humanities Quarterly*, 14(4).
- Hu, Z., Rahimtoroghi, E., Munishkina, L., Swanson, R. & Walker, M. A. (2017). Unsupervised Induction of Contingent Event Pairs from Film Scenes. *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*. <https://www.doi.org/10.18653/v1/W17-5540>
- Kumano, M., Arikawa, Y., Amano, M. & Uehara, K. (2002). Video editing support system based on video grammar and content analysis. *2002 International Conference on Pattern Recognition*, 1051–1055. <https://www.doi.org/10.1109/icpr.2002.1048481>
- Leake, M., Davis, A., Truong, A. & Agrawala, M. (2017). Computational video editing for dialogue-driven scenes. *ACM Transactions on Graphics*, 36(4), 1–14.
<https://www.doi.org/10.1145/3072959.3073653>
- Lin, D. Ch.-E., Caba Heilbron, F., Lee, J.-Y., Wang, O. & Martelaro, N. (2022a). Videogenic: Video Highlights via Photogenic Moments. *arXiv.org*, November 22, 2022.
<https://arxiv.org/abs/2211.12493v1>
- Lin, D. Ch.-E., Caba Heilbron, F., Lee, J.-Y., Wang, O. & Martelaro, N. (2022b). VideoMap: Video Editing in Latent Space. *arXiv.org*, November 22, 2022. <https://arxiv.org/abs/2211.12492>
- Lino, C., Christie, M., Ranon, R. & Bares, W. (2011). The Director's Lens. *Proceedings of the 19th ACM International Conference on Multimedia*.
<https://www.doi.org/10.1145/2072298.2072341>
- Liu, D., Li, Y., Lin, J., Li, H. & Wu, F. (2020). Deep learning-based video coding. *ACM Computing Surveys*, 53(1), 1–35. <https://www.doi.org/10.1145/3368405>

- Lombardo, V. & Damiano, R. (2010). An intelligent tool for narrative-based video annotation and editing. *2010 International Conference on Complex, Intelligent and Software Intensive Systems*. <https://www.doi.org/10.1109/cisis.2010.181>
- Long, A., Myers, B., Casares, J. P., Stevens, S. & Corbett, A. (2002). *Video Editing Using Lenses and Semantic Zooming*. . Carnegie Mellon University. Journal contribution. <https://www.doi.org/10.1184/R1/6626477.v1>
- Manovich, L. (2000). *The language of new media*. Cambridge, MA: MIT Press.
- Manovich, L. (2020). *Cultural analytics*. Cambridge, MA: The MIT Press.
- Matsuo, Y., Amano, M. & Uehara, K. (2002). Mining video editing rules in Video streams. *Proceedings of the Tenth ACM International Conference on Multimedia*. <https://www.doi.org/10.1145/641007.641058>
- McDermott, S., Li, J. & Bares, W. (2002). Storyboard frame editing for cinematic composition. *Proceedings of the 7th International Conference on Intelligent User Interfaces*. <https://www.doi.org/10.1145/502716.502759>
- McInnes, L., Healy, J. & Astels, S. (2017). HDBSCAN: Hierarchical density-based clustering. *The Journal of Open-Source Software*, 2(11), 205. <https://www.doi.org/10.21105/joss.00205>
- Morante, M. F. (2017). Editing and montage: Definition and scope of practice. In *Editing and montage in international film and video: Theory and technique* (pp. 1–8). New York: Essay, Routledge.
- Murch, W. (2001). *In the Blink of an Eye*. Los Angeles: Silman-James Press.
- Myers, B. A., Casares, J. P., Stevens, S., Dabbish, L., Yocum, D. & Corbett, A. (2001). A multi-view intelligent editor for Digital Video Libraries. *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*. <https://www.doi.org/10.1145/379437.379461>
- Nack, F. & Parkes, A. (1997). The application of video semantics and theme representation in automated video editing. *Representation and Retrieval of Video Data in Multimedia Systems*, 57–83. https://www.doi.org/10.1007/978-0-585-31786-1_4
- Okun, J. A. & Zwerman, S. (2021). *The Ves Handbook of Visual Effects: Industry Standard VFX practices and procedures*. New York: Routledge.
- Outtagarts, A. & Mbodj, A. (2012). A cloud-based collaborative and automatic video editor. *2012 IEEE International Symposium on Multimedia*. <https://www.doi.org/10.1109/ism.2012.78>
- Pavel, A., Reed, C., Hartmann, B., & Agrawala, M. (2014). Video digests. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. <https://www.doi.org/10.1145/2642918.2647400>
- Pearlman, K., Brylla, C. & Kramer, M. (2018). Documentary editing and distributed cognition. In *Cognitive theory and documentary film* (pp. 303–319). London: Essay, Palgrave Macmillan.
- Podlesnyy, S. (2019). Towards data-driven automatic video editing. *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*, 361–368. https://www.doi.org/10.1007/978-3-030-32456-8_39
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G. *et al.* (2021). Learning Transferable Visual Models from Natural Language Supervision. *arXiv.org*, February 26, 2021. <https://www.doi.org/10.48550/arXiv.2103.00020>
- Reisz, K. (1957). *The technique of film editing*. London: Focal.
- Rohrbach, A., Rohrbach, M., Tandon, N. & Schiele, B. (2015). A dataset for movie description. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://www.doi.org/10.1109/cvpr.2015.7298940>
- Ronfard, R. (2017). Five Challenges for Intelligent Cinematography and Editing. Retrieved from <https://diglib.org/bitstream/handle/10.2312/wiced20171069/037-041.pdf>
- Rui, Y., He, L., Gupta, A. & Liu, Q. (2001). Building an intelligent camera management system. *Proceedings of the Ninth ACM International Conference on Multimedia*. <https://www.doi.org/10.1145/500141.500145>

- Salt, B. (1974). Statistical Style Analysis of Motion Pictures. *Film Quarterly*, 28(1), 13–22.
<https://www.doi.org/10.2307/1211438>
- Schörkhuber, D., Seitner, F., Salzbrunn, B., Gelautz, M. & Braun, G. (2017). Intelligent film assistant for personalized video creation on mobile devices. *Proceedings of the 15th International Conference on Advances in Mobile Computing & Multimedia-MoMM2017*.
<https://www.doi.org/10.1145/3151848.3151881>
- Singh, H., Kaur, K. & Singh, P. P. (2023). Artificial Intelligence as a facilitator for film production process. *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*. <https://www.doi.org/10.1109/aisc56616.2023.10085082>
- Smith, T. J. (2012). The attentional theory of cinematic continuity. *Projections*, 6(1), 1–27.
<https://www.doi.org/10.3167/proj.2012.060102>
- Soe, T. H. (2021). *AI Video Editing Tools. What Editors Want and How Far Is AI from Delivering?*
<https://www.doi.org/10.48550/arXiv.2109.07809>
- Suchan, J. & Bhatt, M. (2016). Semantic Question–Answering with Video and Eye–Tracking Data: AI Foundations for Human Visual Perception Driven Cognitive Film Studies. *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*, 2633–2639.
- Taskir, C. M., Pizlo, Z., Amir, A., Ponceleon, D. & Delp, E. J. (2006). Automated video program summarization using speech transcripts. *IEEE Transactions on Multimedia*, 8(4), 775–791.
<https://www.doi.org/10.1109/tmm.2006.876282>
- Tran, T. N., Drab, K. & Daszykowski, M. (2013). Revised DBSCAN algorithm to cluster data with dense adjacent clusters. *Chemometrics and Intelligent Laboratory Systems*, 120, 92–96.
<https://www.doi.org/10.1016/j.chemolab.2012.11.006>
- Tsivian, Y. (2009). Cinematics, part of the Humanities' Cyberinfrastructure. *Digital Tools in Media Studies*, 93–100. <https://www.doi.org/10.1515/9783839410233-007>
- Ueda, H., Miyatake, T., Sumino, S. & Nagasaka, A. (1993). Automatic structure visualization for video editing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems-CHI '93*. <https://www.doi.org/10.1145/169059.169118>
- Winer, D., Magliano, J., Clinton, J., Osterby, A., Ackerman, T. & Young, R. (2021). A specialized corpus for film understanding. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 13(2), 266–272.
<https://www.doi.org/10.1609/aiide.v13i2.12984>
- Wu, H.-Y., Santarra, T., Leece, M., Vargas, R. & Jhala, A. (2020). Joint attention for automated video editing. *ACM International Conference on Interactive Media Experiences*.
<https://www.doi.org/10.1145/3391614.3393656>
- Xie, H. & Sun, Y. (2022). An intelligent video editing automate framework using AI and Computer Vision. *Artificial Intelligence and Machine Learning*.
<https://www.doi.org/10.5121/csit.2022.121216>
- Zhang, X., Li, Y., Han, Y. & Wen, J. (2022). *AI Video Editing: A Survey*.
<https://www.doi.org/10.20944/preprints202201.0016.v1>
- Şen, M., Şen, Ş. N. & Şahin, T. G. (2023). A new era for data analysis in qualitative research: ChatGPT! *Shanlax International Journal of Education*, 11(S1–Oct), 1–15.
<https://www.doi.org/10.34293/education.v11i1s1-oct.6683>

Annex 1

Prompt for extraction of concepts:

prompt =

“““

I am an expert researcher specialized in qualitative analyses and the extraction of key concepts and their relationships from texts.

I always answer in Spanish.

My research interest is concerned with analysing interviews with professional editors about their film editing.

These interviews explore the practices, needs and perspectives of film editors as regards the adoption of artificial intelligence (AI) tools.

The themes covered include the differences in the film editing techniques of fiction and non-fiction, methods of organizing audiovisual material, the organization of film montage, the use of formal patterns, the applicability of automated tools, and preferences regarding software and operating systems.

My task is to analyse and synthesize each segment of the interviews into three key concepts, presented in a clear, structured JSON format.

When examining each fragment of text, I focus on identifying the most specific and relevant terms, considering their context and intrinsic meaning.

The output is a JSON list, where each element represents a paragraph with a maximum of three key concepts.

The structure of the response is as follows:

```
[  
  {  
    "text number": number,  
    "key concepts": ["concept 1", "concept 2," "concept 3"]  
  },  
  ...  
]
```

The response always maintains this structure, respecting JSON syntax

“““