Collaborative video editing

Pavel Okopnyi

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2024

UNIVERSITY OF BERGEN

Collaborative video editing

Pavel Okopnyi

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 26.01.2024

Year: 2024

Title: Collaborative video editing

Name: Pavel Okopnyi

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

This thesis was developed during my tenure as a PhD Research Fellow at the Department of Information Science and Media Studies at the University of Bergen. It forms part of the collaborative research project "Better Video Workflows via Real-Time Collaboration and AI-Techniques in TV and New Media," with a specific emphasis on the real-time collaboration aspect within the project. This project is a joint endeavour between the Interaction Research Group at the Department of Information Science and Media Studies, and Vizrt, a leading global broadcasting and media technology company based in Bergen.

The project has been generously funded by the Research Council of Norway under Grant No.: 269790.

Being an integral part of a larger initiative, I had the privilege of working under the guidance of two academic supervisors, whilst maintaining close ties with Vizrt. My thesis supervisor is Prof. Frody Guribye, and my co-supervisor is Prof. Oskar Juhlin. Their expertise in HCI and CSCW helped to navigate my research in these two fields. Throughout the course of my PhD, I have had several opportunities to present my research findings to the employees of Vizrt and involve them in the research process.

This collaborative project also encompasses explorations into innovative technologies and their applications, such as Vizrt's proprietary video editing software, Viz Story. While my research did not entail modifying Viz Story or creating new features, it focused on understanding and enhancing the collaborative aspects of video production and editing in particular. The findings of this study provide valuable insights that can guide the design and development of collaborative tools for video production, and thereby contribute to the overarching goal of the research project.

Acknowledgements

I'd like to start by thanking my main supervisor, Prof. Frode Guribye, and my co-supervisor, Prof. Oskar Juhlin. Their guidance, support and patience have been invaluable throughout this process.

I also want to thank Marija Slavkovik, the Head of Department, for her continuous support.

A big thanks goes out to Vizrt. Their team, particularly Roger Sætereng, Even Normann, and former employees Ronan Huggard, Brage Breivik and Valentina Caruso, have been immensely helpful. They've assisted in organising my research activities and have been instrumental in introducing me to TV professionals from various companies.

I would like to thank my colleagues at UiB who helped me throughout these years. The DiggUiB group, for their video production expertise, which helped kickstart this project immensely. My fellow PhD students, Oda Elise Nordberg, who helped organise some of the research activities; Than Htut Soe, who participated in the project studying video production workflows from the perspective of AI applications. Eivind Flobak and Anja Salzmann, for sharing their experiences of PhD life and work. Louise Sandal Løkeland, for her help with transcript translations. I must also thank Robert Gray, a poet and professor, for participating in the project.

I must also mention my friend, Ilya Musabirov, who was always there for good advice and support.

And, of course, I can't forget my family and friends. Their support and encouragement have been more helpful than I can express. So, thank you all.

Preface

When I began this research, I had limited knowledge of the TV, video, and film industries. However, there was a simple and straightforward question guiding me: *why is there no Google Docs for video?* Vizrt, the company I collaborated with, was also grappling with this question. Their software VizStory was based on a similar concept, but designing and developing a collaborative video-editing tool proved more complex than anticipated.

To find answers, I spoke to professionals from different parts of the world to understand how videos are made. What I discovered was a significant overlap between the TV and video industries. Many video professionals had a background in TV and vice versa. The work experience of the editors I interviewed varied widely, ranging from YouTube content creation to working on feature films and TV series. Most of them were excited about the research topic and considered it relevant.

As the world changed with the COVID-19 pandemic in 2020, this research became more relevant. Many organisations shifted to remote workflows, highlighting the pressing need for online collaborative video-production tools.

The research resulted in this article-based thesis. It includes three articles where I am the first author, which complies with the UiB guidelines for an article-based thesis.

The research was completed together with my supervisors, Frode Guribye and Oskar Juhlin. In the second study, Valentina Caruso, a UX researcher at Vizrt, also participated in data collection, analysis, and writing.

A Brief Overview of Research Contributions:

- **First article**: I executed all data collection activities, which consisted of 11 interviews with 13 participants and eight follow-up interactions via diverse platforms. I was the responsible coder during the data analysis process, while Frode and Oskar participated in collaborative coding sessions in which we refined the coding scheme and discussed the overarching themes and direction of the study.

- **Second article**: I was solely responsible for three interviews, while Valentina managed two. We jointly conducted the remaining seven. The data analysis began with my preliminary coding, which was refined further with input from Frode, Oskar, and Valentina.
- **Third article**: I planned, organised, and conducted the first workshop with assistance from Oda Elise Nordberg. The second workshop was directed by a group of HCI master's students at UiB as a part of their training, while I developed the workshop protocol and participated in its organisation. I planned and conducted workshops 3–5, and Roger Rebbestad Sæteren, an R&D Manager at Vizrt, helped organise these workshops. The analysis was once again a collective effort, with Frode and Oskar contributing significantly after my initial groundwork.

In each study, I produced the first draft of the paper and its overall structure. Other co-authors contributed by writing certain sections and refining the text, which was completed through an iterative process. Between iterations, we conducted online and offline meetings to discuss the text in progress and redistribute writing responsibilities among ourselves.

In addition to the three articles, this thesis includes an explanation of the context and motivation behind the research, the theoretical framework that underpins it, and the methodological approach adopted in conducting it. The thesis also articulates the common thread binding different facets of the research presented in the three articles, elaborates on the interconnections between the studies, and discusses the findings and contribution of the research.

Abstract

This thesis addresses the following question: how can collaboration be supported in video editing? In many domains, such as writing and design, collaborative tools have become common and widespread. However, video-editing software is still predominantly designed for solo users. Nevertheless, video editing is a social activity that, in a professional setting, often involves various people working together.

Based on interviews and design workshops, this thesis investigates the collaborative practices of video editors and explores the design space of collaborative video editing.

In three studies, this thesis looks at video editing from three angles. First, it investigates the collaborative practices of video editors and identifies the strategies and social mechanisms they employ to reach agreements with various parties involved in the video-production process. The first study identifies nine themes that characterise the ways video editors manage uncertainties and reach agreements, particularly through organisational mechanisms, documentation, and iconic referencing. The study also suggests three design paths to explore further.

Second, it examines video production from an organisational point of view, focusing on the recent shift towards remote work and its impact on video production. The second study delineates the short-term and long-term implications of adopting remote work in TV production organisations during the COVID-19 pandemic.

Third, it approaches collaborative video editing as a design problem and offers design ideas to enhance collaboration. Additionally, it uncovers challenges that might impede the adoption of new collaborative video-editing tools.

In synthesising the results of the three studies, as well as analysing previous research and existing video-editing tools, this thesis identifies three design approaches for supporting collaboration in video-editing software: *holistic*, *tailored*, and *configurable*. While discussed in the context of collaborative video editing, these approaches offer a broader analytical framework for considering the design of collaborative production tools.

Abstrakt

Samarbeid i videoredigering[1]

Denne avhandlingen tar opp følgende spørsmål: Hvordan kan vi støtte samarbeid i videoredigering? I ulike anvendelsesområder, som skriving og design, er bruk av samarbeidsverktøy utbredt. Likevel er programvare for videoredigering i hovedsak utviklet for individuell bruk. Videoredigering bør forstås som en sosial aktivitet og blir i profesjonelle sammenhenger ofte utført som et samarbeid mellom ulike aktører.

Basert på intervjuer og designverksteder, undersøker denne avhandlingen hvordan videoredigerere samarbeider og utforsker mulighetsrommet for å støtte samarbeid i videoredigering gjennom design av nye løsninger.

I tre studier undersøker denne avhandlingen videoredigering fra tre perspektiver. Først undersøker den samarbeidspraksiser blant profesjonelle videoredigerere og identifiserer ulike strategier og sosiale mekanismer som brukes for å oppnå enighet mellom aktørene som er involvert i videoproduksjon. Denne første studien identifiserer ni temaer som beskriver hvordan videoredigerere håndterer usikkerhet og oppnår enighet, spesielt gjennom organisatoriske mekanismer, dokumentasjon og ikoniske referanser. Studien foreslår også tre ulike retninger for design av nye løsninger for å støtte samarbeid i videoredigering.

Det andre studiet undersøker videoproduksjon fra et organisatorisk perspektiv, med fokus på en pågående overgang til distribuert arbeid og dets innvirkning på videoproduksjon. Den andre studien skisserer de kortsiktige og langsiktige implikasjonene av å innføre distribuerte arbeidsformer i TV-produksjonsorganisasjoner under COVID-19-pandemien.

Den siste studien ser på samarbeid i videoredigering som et designproblem og presenterer designideer for hvordan man kan støtte et slikt samarbeid. I tillegg peker denne studien på utfordringer som kan være til hinder for innføringen av nye videoredigeringsverktøy

[1]Title of the thesis in Norwegian

som skal støtte samarbeid.

Ved å sammenstille resultatene fra de tre studiene, samt analysere tidligere forskning og eksisterende videoredigeringsverktøy, identifiserer avhandlingen tre designtilnærminger for å støtte samarbeid i videoredigeringsprogramvare: holistisk, skreddersydd og konfigurerbar. Selv om disse tilnærmingene diskuteres med tanke på samarbeid i videoredigeringspraksiser, kan de tilby et bredere analytisk rammeverk for å vurdere utformingen av samarbeidsverktøy også for andre anvendelsesområder.

List of publications

1. Pavel Okopnyi, Oskar Juhlin, and Frode Guribye. *Unpacking Editorial Agreements in Collaborative Video Production*, ACM International Conference on Interactive Media Experiences (IMX '20) **17**, 6, 2020.

2. Pavel Okopnyi, Frode Guribye, Valentina Caruso & Oskar Juhlin, *Automation and redistribution of work: the impact of social distancing on live TV production*, Human–Computer Interaction **23**, 11, 2021.

3. Pavel Okopnyi, Oskar Juhlin, and Frode Guribye, *Designing for Collaborative Video Editing*, Nordic Human-Computer Interaction Conference (NordiCHI '22) **8**, 10, 2022.

Additional publications during the PhD

1. Okopny, P., Musabirov, I., & Bakhitova, A. (2018). *Designing for Code Sharing in a Data Science Course for non-STEM students.* Design & Learning-Centric Analytics, 23, 146.

2. Musabirov, I., Suvorova, A., Bulygin, D., & Okopnyi, P. (2019). *Co-aligning UX & development courses: the case of MSc in information systems and HCI.* In The Proceedings of EduCHI 2019 Symposium.

Contents

Chapter 1

Introduction

In today's digital age, video content has become ubiquitous, spreading across platforms ranging from traditional TV to modern social media sites and applications such as YouTube, Twitch, and TikTok [Bartolome and Niu, 2023]. This surge has led to a diverse array of individuals, from professionals to amateurs, engaging in video production. As Juhlin et al. [2014b] note, there has been a shift in forms of interaction "from mere consumption to a mix of production, consumption, and sharing" [Juhlin et al., 2014b, p.687], emphasizing that a previously passive audience now participates in video production and distribution.

This engagement is facilitated by the increased accessibility of production tools, which has created new user categories and trends. For example, the term "professional amateur" (pro-am), as introduced by Leadbeater and Miller [2004], describes an enthusiast who, while not professional, strives to achieve professional standards in various fields, including video [Juhlin et al., 2014a]. Similarly, Juhlin et al. [2014b] highlight three trends that conceptualise interactions with video: user production of video, the use of mobile devices in production, and interactive TV viewing.

The intricate and multidimensional nature of video production, which encompasses distinctions like live versus recorded content and the nuances of professional, pro-am, and amateur production, makes it a vast field. While the differentiation between live and recorded content presents its own set of challenges and complexities, this thesis does not delve deeply into that topic. Instead, it focuses on one facet of video production: video editing.

Video editing is a complex and multifaceted activity. In a professional setting, it typically involves a diverse team of collaborators who each bring unique skill sets and expertise to the process. Video workers managing tasks such as montage, colour correction, sound

levelling, and graphics production contribute to the creation of a video. Despite the collaborative nature of such work, existing video-editing software and tools are primarily designed for individual use, resulting in a discrepancy between the collaborative production setting and the functionality of editing software.

This Ph.D. thesis investigates collaborative video editing and focuses on suggesting design ideas for supporting collaborative workflows and remote collaboration. Drawing on interviews with professional video workers, collaborative design sessions, insights from the fields of human–computer interaction (HCI) and computer-supported cooperative work (CSCW), and an examination of existing collaborative tools for media production, this research contributes to an understanding of the complexities of collaborative video editing and the development of tools to support it.

This research aims to address the apparent disparity between individualised video-editing tools and the inherently collaborative nature of the video-editing process as well as to contribute to the understanding of designing for collaborative video editing.

This research focuses on professional video editing, exploring the challenges of collaboration between video editors and other contributors throughout the production process. The emphasis is primarily on team-based collaborative work. Specifically, the research addresses the design challenges associated with enabling collaboration between professionals and non-professionals. The main context of this research is professional video editing and TV production.

In the following, I introduce and elaborate on the research question of this thesis, summarise the articles that constitute it, and provide an outline of the subsequent text.

1.1 Research Question and Contribution

Remote collaboration is widespread in many different fields. People collaboratively write with Google Docs and design with Figma, Miro, and other applications [Feng et al., 2023]. Technological developments, such as broadband internet connection and cloud technologies, provide fertile ground for the development of such tools, and the technical feasibility of online collaborative video-editing software has been demonstrated by previous research [Klokmose et al., 2019; Takhirov and Duchateau, 2011]. There have been several notable efforts to support collaboration by developing new tools, for example, in a co-present setting by Bartindale et al. [2012] and in remote conditions by Pavel et al. [2016]. However, video-editing tools are still mostly designed with an individual user in mind. Existing collaborative video-editing tools (see 2.5) primarily target niche ap-

plications, such as editing videos for social media, and often offer a feature set that is relatively limited in comparison to more "professional" editing software such as Adobe Premier Pro.

As such, why are there so few professional collaborative video-editing tools? When considering the potential for improvement and innovation, this query evolves into a more structured and actionable research question:

> *How can collaboration be supported in video editing?*

There may be different approaches to designing collaborative video-editing software, and part of the contribution of this thesis is identifying and conceptualising them. To answer the research question, I first focus on gaining an understanding of how video editors organise their collaboration and, specifically, how they navigate disagreements in the editing process. Subsequently, I study the impact of a shift to remote work on video work. Finally, I address collaborative video editing as a design problem and suggest design approaches that may support it. In this way, the research question can be divided into three sub-questions:

1. *What tools, strategies, and mechanisms do video editors use to organise collaboration?*
2. *How does a shift to remote work reshape collaborative practices in video editing?*
3. *How can video editing software be designed for remote collaboration?*

By addressing these questions, this research contributes to the field of collaborative video editing by providing new insights in several key areas. First, it identifies the challenges associated with collaboration in video editing and describes the strategies and mechanisms video editors employ to achieve agreements. Second, it describes design paths and concepts for software tools that aim to support collaborative video editing. Third, it reveals challenges associated with collaborative video-editing software design. Fourth, it highlights how the adoption of remote workflows and automation changed video work in production organisations during the COVID-19 pandemic.

Finally, this research contributes by identifying three design approaches for supporting collaboration in video editing: *holistic*, *tailored*, and *configurable*. The holistic approach implies designing a comprehensive solution by integrating a broad array of functionalities into one tool. The tailored approach presumes a focus on specific tasks or user groups,

optimising software for particular needs or workflows. The configurable approach emphasises the flexibility of software, allowing users to shape the system to their requirements. These approaches, while discussed in the context of video-editing software, have broader implications for designing collaborative tools in other domains.

These contributions are discussed in more detail in chapter 4. The following section describes the three studies that constitute this research and their individual results.

1.2 Introduction to Articles

I address the research question in three studies. First, I investigate social practices in video editing, focusing specifically on the reviewing process, which has been highlighted by previous research [Pavel et al., 2016]. Based on the results, the study suggests design directions to support collaboration. In the second study, I investigate the impact of social-distancing regulations on video work in TV production companies and how they were addressed by adopting remote workflows. The third study explores the design space for collaborative video editing, identifies design challenges, and suggests concepts to address them.

Thus, the first article scrutinises a specific practice of reviewing to untangle its collaborative nature, and the second article examines collaborative video production from an organisational perspective. Finally, the third article considers collaborative video production as a design problem. In the following, I summarise the articles.

1.2.1 Article I: Unpacking Editorial Agreements in Collaborative Video Production

The first article [Okopnyi et al., 2020] is dedicated to investigating practices surrounding the reviewing process.

The video-production process is often marked by uncertainty. Specifically, various parties who participate in the production process, such as editors, directors, and domain experts, often face flexibility of interpretation (i.e., "interpretive flexibility", see [Collins, 1981; Pinch and Bijker, 1984]) concerning different aspects of video, including colours, story, acting, and pacing.

The article analyses strategies editors employ to reduce uncertainty and achieve agreements with other parties. In particular, the article identifies four ways interpretive flexi-

bility manifests itself in video work, such as context-related uncertainty, and five ways of reaching editorial agreements, including the use of documentation or iconic referencing.

The article suggests three general design paths for supporting collaboration in video editing. The first path, "scaffolding" [Mercer, 1994; Wood et al., 1976], addresses the problem of interpretive flexibility by providing a preset structure that supports collaboration. Like the approach reported by Luther et al. [2015], the article proposes integrating such structures to accommodate participants' level of knowledge of video production.

The second path, "iconic referencing", aims to augment communication between reviewers and video editors by providing examples that reflect collaborators' ideas for the video in production. The term "iconic referencing" alludes to semiotic theory [Peirce and Buchler, 1902], and the mechanism is intended to bolster the practice of indexical referencing present in the reviewing process.

The third path, "suggestive editing," focuses on including various parties in the editing process. The article suggests developing a lean video-editing software suite with limited functionality designed to be used by non-editors. Thus, participants without video editing training would be able to perform certain manipulations that a video editor could accept or decline.

The article suggests that the presented design paths should be investigated in research through design studies and presents several design ideas that can be further developed.

1.2.2 Article II: Automation and Redistribution of Work: the Impact of Social Distancing on Live TV Production

The second article [Okopnyi et al., 2021] analyses how TV production companies adapted to restrictions imposed by the COVID-19 pandemic.

Due to social-distancing regulations, many companies were forced to implement remote workflows in video production. The article investigates how the TV industry relied on new workflows and internet technologies to continue producing video content when TV studios had been mostly vacated and TV workers were required to work remotely. The article identifies two conceptual themes that describe socio-technical changes in TV production: automation and redistribution of work.

The notion of redistribution is central to this article. The term "redistribution" extends beyond the spatial reconfiguration inherent in remote work to include the transformation of work practices, the reallocation of responsibilities, and the reskilling of various actors

involved in the production process. For instance, journalists and talk-show guests needed to acquire technical skills, such as assembling a camera or managing noise levels, which traditionally have been the realm of specialised TV workers like camera operators.

Moreover, redistribution also indicates the incorporation of automated systems and internet tools designed for the general public, such as Skype and Zoom, into the production process. These tools have enabled video communication and production in ways that have been thought to require the use of specialised equipment and setups. With automation, traditional production workflows change and require significantly more work at the preproduction stage, such as developing scripts for automation systems and establishing backup systems and procedures to manage breakdowns.

The forced shift towards a new work model could indicate a significant transformation in the TV and wider video-production industries. The adaptations introduced during the pandemic may not be temporary measures but could represent a longer-term change towards an increasingly democratised and collaborative mode of production where the line between professional and amateur is increasingly blurred. It also indicates a trend of tools and capabilities empowering individuals to contribute to processes from which they were previously excluded, thus reshaping the landscape of the TV industry in profound ways and creating new working environments where video production is more accessible to non-professionals and smaller organisations.

1.2.3 Article III: Designing for Collaborative Video Editing

The third article [Okopnyi et al., 2022] explores the design space for remote collaboration in video editing. Based on a series of workshops, the article presents 10 design concepts demonstrating how collaboration can be supported in contemporary video-editing software. The workshop participants were tasked with envisioning collaborative features that could be implemented in non-linear editing (NLE) software to support concurrent editing by multiple collaborators working in a shared workspace.

The design concepts represent different combinations of features or approaches that can be adopted when designing software for remote collaboration in video production. Each concept highlights different possibilities and compromises within the design space, addressing various aspects of the collaborative editing process. Through an analysis of the workshop results (i.e., transcripts and sketches), the article summarises these aspects into three overarching themes: focus and distraction, new workflows and tools, and working with the temporal dimension.

Maintaining focus on work and avoiding distraction appeared to be a general and serious concern expressed by workshop participants in their discussion of design concepts. This concern was raised especially when discussing "generic" collaborative elements that are present in other collaborative tools, such as colour-coded avatars. The introduction of these elements might create sources of distraction for video editors. Thus, adding such elements might be challenging. The results also suggest that introducing additional structural mechanisms might help address distraction by separating shared workspaces through functional and time-slot separation.

Several design concepts involve supporting various activities that usually occur during the video editing process but are often overlooked by NLE software. For example, reviewing often requires using external tools, from written notes to third-party software. These tools typically have no or limited integration with NLE software and require additional time and effort when sharing the video with the reviewer and collecting and processing feedback.

Finally, the article addresses the challenge of designing for time-based media. Traditionally, the temporal dimension is represented in the NLE software as a line, that is, a *timeline*. The article suggests that alternative representations, such as a storyboard [Goldman et al., 2006], might help facilitate collaboration.

1.3 Outline of the Thesis

This thesis is structured into five main chapters, including an introduction, to provide an understanding of the challenges and opportunities presented by collaborative video editing.

Chapter 2 provides essential information on the field of video production and introduces contemporary video production technologies and technological trends to the reader. It also summarises previous research dedicated to video production and design for collaboration in creative industries.

Chapter 3 details the research methodology employed in the study, including design workshops and interviews with video editing professionals as well as data analysis and synthesis.

Chapter 4 presents key findings from the design workshops and interviews, examines the challenges of collaborative video editing, explores opportunities for enhancing collaboration, and discusses implications for the design of future collaborative video-editing

tools.

Chapter 5 concludes this thesis, summarising the findings and suggesting directions for further research.

Chapter 2

Background

This chapter lays the foundation for understanding and contextualising the findings from the studies conducted within this research project.

Initially, it presents an overview of existing research in the human–computer interaction (HCI) and computer-supported cooperative work (CSCW) fields, focusing on collaboration in media production. Subsequently, it explores the video-production process with an emphasis on the video-editing stage, which is the focus of this research.

The chapter proceeds by discussing the representation of video in non-linear editing (NLE) software, providing essential background information for comprehending the design aspects of such tools. This is followed by an exploration of current collaborative systems employed in video production and a summary of research on collaborative video editing.

The chapter then outlines various strategies adopted by collaborators to facilitate communication before discussing scientific frameworks for designing collaborative systems, with special attention given to the concept of workspace awareness.

Last, the chapter briefly summarises the technological trends shaping the video-production and TV industries, including the growing accessibility of video-production tools, the rise of remote work, and the increasing role of automation.

2.1 Collaborative Media Production in Research

Collaborative media production has been an important topic in the CSCW and HCI fields. Numerous studies approach this subject from different perspectives, investigating

strategies for supporting and enhancing collaboration, current practices and case studies, and technical challenges of collaborative media production.

Some studies explore how existing general-purpose tools, such as mobile phones, can be used in media production [Feinberg et al., 2016; Salzmann et al., 2020] and facilitate creative collaboration [McGrath et al., 2016]. For instance, Feinberg et al. [2016] highlight the possibilities and challenges of using mobile phones as a platform for video editing, noting the need to keep interfaces uncluttered and ensure large zones for finger touches due to the small screen size.

Others investigate ways to enhance and optimise remote collaboration through technological advancements, such as video-mediated collaborative systems [Kim et al., 2018], virtual and mixed-reality techniques [Lee et al., 2017; Nguyen et al., 2017], and novel interactive devices and approaches [Cabral and Correia, 2012; Taylor et al., 2007]. Moreover, McGrath et al. [2016] and O'Leary et al. [2018] emphasise the necessity for tools that support not only creative work like editing but also the broader production process. This includes planning, progress-tracking, and version control, which refers to the management of multiple iterations of a project to keep track of changes and enable rollback to previous states if needed.

These new means of facilitating collaboration in media production have been found to improve co-presence and increase mutual understanding between remote participants [Kim et al., 2018; Lee et al., 2018]. Kim et al. [2018] noted that collaboration in such settings occurs in one of two modes: between peers who work together as equals or between an expert who guides and instructs a less experienced collaborator [Kim et al., 2018]. Remote activities, such as remote sketching, can aid in guidance; however, their use can also result in a higher task load [Huang et al., 2019].

Some studies have investigated specific practices, such as feedback and reviewing, which are found to facilitate individual adaptation and are a resource for achieving creative outcomes [Bartindale et al., 2016a; de Stobbeleir et al., 2011; Pavel et al., 2016]. For example, de Stobbeleir et al. [2011] emphasise two ways creative workers seek feedback: active inquiry and monitoring the environment for indirect feedback.

Similarly, Chua et al. [2017] assess the challenges of referencing during discussions in media production. They highlight that referencing can sometimes involve multiple objects, part of an object, or a whole document, thereby necessitating a proper instrument for facilitating such complex references.

Finally, some studies investigate the implementation of software systems for remote collaborative media production work and how various technical challenges might hin-

der it. Such challenges include the organisation of version control and conflict resolution [Novikov and Proskurnin, 2003; Zünd et al., 2017], data compression and transfer [Buysschaert et al., 2020], and systems architecture and practical implementation of complex software for media work [De Roure et al., 2018; Klokmose et al., 2015, 2019; Tang and Minneman, 1991; Thalmann et al., 2016].

2.2 Video Production

Video production might seem like a simple process. However, depending on the nature of a project, it can require multiple professionals with various skill sets to work together in an organised manner. Owens and Millerson [2012] identify two methods of approaching video production: *empirical* and *planned*. The empirical method is guided by instinct and opportunity; producers capture video footage based on their ideas and later create a video from what they have found. This approach can result in fresh uninhibited content but risks being haphazard and lacking cohesion. The planned method involves video producers determining the exact form of the video in advance and creating it accordingly. This approach offers systematic purpose and coordination but may be limited by the organisation of the production process, resulting in a potentially unoriginal video [Owens and Millerson, 2012].

Similarly, in their study of the use of video technologies by home movie makers, Kirk et al. [2007] identify two types of video work, *lightweight* and *heavyweight*. Lightweight video work involves the spontaneous capturing of video during events, primarily for in-the-moment sharing and enjoyment by participants. Heavyweight video work focuses on intentionally capturing significant events, requiring greater effort and attention to quality [Kirk et al., 2007]. This thesis focuses primarily on planned, heavyweight video work.

The video-production process is traditionally divided into three stages: pre-production, production, and post-production [Ascher and Pincus, 2007; Diefenbach, 2009].

The pre-production stage focuses on the planning of the filming process, and various creative and logistical decisions are made. This stage involves brainstorming ideas, writing a script, creating storyboards, scouting locations, casting actors, and organising a production crew.

The production stage involves filming the video. This stage includes recording the audio, directing the actors, and capturing the footage.

Finally, the post-production stage involves steps such as editing the footage, performing colour grading, adding visual effects, and mixing the audio. At this stage, the video is compiled, polished, and refined before it is distributed to the audience.

2.3 Video Editing

Defining the exact boundaries of video editing and separating it from other post-production activities might be challenging. Ascher and Pincus [2007] note that video editors often take "someone else's job", such as sound mastering, visual effects, graphics, and colour correction and grading [Ascher and Pincus, 2007, p.648]. Multiple handbooks and guides on video production include such activities in their description of the video editing process [Diefenbach, 2009; Owens and Millerson, 2012].

Video editing is highly dependent on previous stages of video production. Owens and Millerson [2012] note that the key shots of the video and their sequence are decided early, such as during the process of storyboard development [Owens and Millerson, 2012, p.34]. They also suggest that, during the filming process, the operator should consider the editing which will occur after filming to allow some leeway for the editor. For example, the operator may use longer shots, film background, film the same scene with different paces and angles, and preserve failed takes which might be useful during editing [Owens and Millerson, 2012, p.155]. Notably, the collaborative component of video editing, including reviewing, is somewhat overlooked in the professional video production literature. For example, Owens and Millerson [2012] presents the video-production process as linear, without iterations, and does not mention the reviewing in the context of editing. Similarly, Ascher and Pincus [2007] discuss video editing from a mostly technical perspective, focusing on software, equipment and file formats. Both of these books mention reviewing primarily in the context of the filming process, emphasizing the need to review raw footage to ensure its quality.

Traditionally, video editing was performed physically by cutting and joining film fragments in a so-called "linear editing" process. Currently, a digital approach, NLE, is common [Ascher and Pincus, 2007]. NLE is performed with software suites such as Final Cut, Vegas Pro, DaVinci Resolve, Filmora, and Adobe Premier Pro. These suites feature tools and instruments that encompass many aspects of post-production, such as sound mastering, visual effects application, graphics, and colour correction. They are also typically designed for a single user and provide limited support for collaboration.

When the editing is completed, the editor must render the video, creating a new video

file from the elements of raw footage and applying editor-defined effects. This process is computationally intensive and time-consuming. Thus, video editing often occurs in iterations: the editor creates a draft edit, which is a version of a movie that is rendered and shown to other parties, such as directors and producers, who are then able to review and discuss the intermediate result of editing and provide feedback to the editor [Pavel et al., 2016].

2.4 Representing Video in NLE Software

There are many ways to represent a video. Due to the complexity of the medium, a video clip or movie can be seen from different perspectives: narrative, technical, and user interface (UI) design.

From a narrative viewpoint, video may be approached as scenes, story beats, and dialogue lines. These elements are usually created before the filming process in the form of screenplays and storyboards, which are artefacts that represent sequences of textual descriptions and images of what is intended to occur on the screen and what the viewer is supposed to see [Murch, 2001; Owens and Millerson, 2012]. Such artefacts guide both the filming process and subsequent editing [Bartindale et al., 2012, 2016a,b; Goldman et al., 2006; Mackay and Pagani, 1994]. In their prototype application Video Mosaic, Mackay and Pagani [1994] presented a hybrid approach to video editing that involved a physical (paper) storyboard and online software. Storyboard elements represented the clips from the raw footage, which were printed user-selected frames from the clips. To edit a video, the user manipulated storyboard elements, moving them in a two-dimensional space.

From a technical perspective, video is a set of media files in various formats, such as MPEG-4 and H.264 [Rao et al., 2014]. With NLE software, the editor does not manipulate the footage itself, only its abstract representation in the interface. Thus, from a technical standpoint, the editing of a video implies the editing of a video project file, usually in an XML format, which describes parts of the footage that should be arranged together and in what order [Novikov and Proskurnin, 2003].

From the UI design perspective, a video can be seen as a set of various elements representing pieces of media, such as clips, graphics, visual effects, soundtracks, and subtitles, which are organised along a timeline that is an abstract line representing the temporal dimension. [Mackay and Pagani, 1994]. The timeline is a common way of organising the elements of video in NLE [Fonseca and Carrapatoso, 2006, 1999] and other review-

ing software [Chatti et al., 2016; Pavel et al., 2016]. It is usually separated into several tracks representing various types of media, including video, sound, and graphics. In NLE software, the editor arranges media pieces along the timeline, defining cuts, transitions, and various effects.

In their tangible prototype, Zigelbaum et al. [2007] represented a video with plastic tokens for each clip. Every token contained a screen, allowing the users to preview the clip. The users could physically join the tokens in a line to form a movie. At the joint, the users could place a physical representation of a transition between two clips. Similarly, Merz et al. [2018] represented clips as wooden blocks placed on an interactive tabletop. The users could connect the blocks into a linear sequence to edit the video. Likewise, Terrenghi et al. [2008], in their tabletop application, presented video clips as virtual film strips that users could organise into sequences. Taylor et al. [2007] employed plastic tokens to represent clips, transitions, and video effects. In their prototype, users could attach tokens together in a sequence with "transition tokens" placed between clips and "effects tokens" attached to the sides of the "clip tokens". In each of these prototypes, the users could freely reorganise the workspace, whether physical table or interactive screen space, to place clips and other representations and move them around.

In their toolkit "Videostrates", Klokmose et al. [2019] represented video in the form of an HTML file. Based on the "Webstrates" platform [Klokmose et al., 2015], the paper suggested the development of various interfaces to work with video, each tailored to a specific task, such as editing video structure or subtitles. The idea of personalised interfaces, where each user sees a representation of complex media tailored to their needs and tasks, resonates with some current practices in NLE software development. For example, DaVinci Resolve employs a "pages" metaphor[1]: each page represents a video project tailored to the needs of a specific role, such as a video editor or a colour corrector. Thus, different users see and work with different representations of the same video while employing different tools.

A number of video editing applications, such as SimonSays[2], Reduct[3] and Descript[4], represent video as text. These applications provide automatic video transcription and allow users to edit the transcripts by cutting and rearranging their parts. When parts of the text are being edited, the application edits the corresponding parts of the video, creating a rough draft focused on narrative.

[1]https://www.blackmagicdesign.com/no/products/davinciresolve/
[2]https://www.simonsaysai.com/
[3]https://reduct.video/
[4]https://www.descript.com

2.5 Existing Collaborative Video Tools

There is significant interest in supporting collaboration in the video industry. This interest is indicated by numerous applications and services that aim to enable various collaborative practices. Specifically, many instruments focus on the reviewing process, in which various stakeholders like editors, producers, and clients provide feedback and suggest changes to the video content.

One prominent industrial collaborative tool is *Frame.io*[5], which allows video editors to share videos with reviewers and collect feedback. Frame.io is a standalone web-based service, meaning it is not part of any existing NLE software, though it has limited integration with existing editing tools. This integration allows video editors to import reviewers' commentaries into NLE software and view them attached to the timeline. However, Frame.io nonetheless requires video editors to render the video and upload it.

There are several services with similar functionality to Frame.io, such as *Vimeo*, *Wipster*, *ReviewStudio*, *screenlight.tv*, and *LookAt.io*. Some of these services, including Wipster and screenlight.tv, do not limit themselves to video but allow reviewing of various types of digital artefacts, such as images and audio.

At the time of writing, *DaVinci Resolve* seems to be the only professional post-production NLE tool that supports online collaboration[6]. The software allows multiple collaborators to work on the same project. Communication is managed via a built-in chat, and collaboration is enabled with a bin- and timeline-locking mechanism that locks parts of the video project when edited, meaning only one person at a time can work with them.

There are several web-based online collaborative video-editing services. For example, the service *Motionbox*[7] provides a simplified web-based collaborative NLE software that allows multiple participants to edit videos together in a shared workspace. The UI of the service is similar to other NLE software and includes a timeline and preview window, but it is very limited in its functionality. At the time of writing, the service itself remains under development. Unlike all-in-one software suites like DaVinci Resolve, which could be considered "universal", Motionbox is tailored towards specific tasks and needs, such as creating video content for social networks and music visualisation. This is true for many other similar online services, including *Ozone*[8], *Scalar*[9], *Veed.io*[10], *Scenery*[11], and

[5] https://frame.io
[6] https://www.blackmagicdesign.com/no/products/davinciresolve/collaboration
[7] https://motionbox.io/
[8] https://ozone.tech/
[9] https://scalar.video
[10] https://www.veed.io
[11] https://scenery.video

Kapwing[12]. At the time of writing, some, including Ozone and Scalar, remain under development. Several services, such as Kapwing, offer users highly specialised tools, such as "meme-generator" and "censor-video" tools, that provide users with tailored UI while employing the same video-editing software system at their core. Incidentally, most of these services either appeared recently or remain in development, which signifies interest within the industry for collaborative video-production software.

2.6 Collaborative Video Editing in Research

Co-present collaborative video editing has been investigated in design explorations by various researchers. Taylor et al. [2007] developed a system called "VideoPlay", a tabletop application that employed Microsoft Surface, a high-precision multi-touch display, and presented video clips as physical tiles that users could manipulate manually. The system was intended to provide playful interaction and make the video editing process more engaging. Zigelbaum et al. [2007] created the "Tangible Video Editor" (TVE), a physical system that presented video clips as plastic tokens with built-in screens. The evaluation of the prototype suggested that users who worked with TVE were more engaged in the creative process. However, some users complained about the lack of functionality of TVE.

Inspired by both VideoPlay and TVE, Merz et al. [2018] developed a prototype editor called "ClipWorks". The system employed an interactive tabletop and physical blocks that served as the elements of the UI. The system was designed for schoolchildren and had a limited number of functions. Similarly, Terrenghi et al. [2008] designed an interactive tabletop application to enable spontaneous video editing. The application provided limited editing functionality and a physical tool with an interactive screen.

Bartindale et al. [2012] designed the tabletop system "StoryCrate". The prototype was deployed on a set during the filming process and allowed users to playback recently filmed clips and perform editing. The system supported production workflows by allowing participants to review recently recorded video and audio and informing them about the current state of the filming process. In a follow-up study, Bartindale et al. [2016a] presented a tabletop collaborative video-editing system "TryFilm", that was also deployed and evaluated on a filming set. The system was used by the cast and the filming crew to review actors' performances and reflect on the video, as well as to perform draft editing of marketing material.

[12]https://www.kapwing.com

However, the specific topic of remote collaborative video editing has been somewhat overlooked in research. There have been limited attempts to address this topic, and they approach it from a primarily technical perspective, addressing the feasibility and viability of online collaborative video-editing software.

Novikov and Proskurnin [2003] presented a data model called "concurrent video" that provided means to organise and manipulate video footage by multiple video editors simultaneously. The technical details of this model were further investigated, highlighting the possibility of creating a collaborative video-editing software [Proskurnin, 2005].

Similarly, Fonseca and Carrapatoso [1999] presented a software architecture for collaborative video editing that aimed to allow multiple remote co-editors to work together on the same video. They suggested that such software systems would be economically attractive for the broadcasting and entertainment industries. They emphasised that remote collaborative video editing involves two groups of challenges: technical, which are related to networking issues, security, and reliability, and social, as co-editors would need to follow specific social protocols when working together in a shared workspace [Fonseca and Carrapatoso, 1999]. Subsequently, Fonseca and Carrapatoso [2006] demonstrated and evaluated the prototype of a collaborative video-editing application called "Coview". The application performed "quite well" in the evaluation [Fonseca and Carrapatoso, 2006, p.489]. It may become a valuable and powerful tool for professional journalists, specifically when a reporter who sends video material to a TV studio from a distant location wants to be involved in the editing process.

Takhirov and Duchateau [2011] presented a cloud-based collaborative video production tool "Creaza VideoCloud." Their research focused on the technical implementation of collaborative video tools. Schenk et al. [2016] demonstrated how video could be presented in a text-based format, thus enabling collaborative video editing via text-based tools such as version control systems. The Videostrates toolkit by Klokmose et al. [2019] allowed users to create collaborative video applications tailored to their needs by configuring different features through editing HTML code. That study provided proof-of-concept technology and emphasised the need for further design research in this domain.

2.7 Addressing Ambiguity in Collaboration

A considerable body of research has emphasised the inherent complexity of communication within media production contexts [O'Leary et al., 2018; Phalip et al., 2007; Retelny and Hinds, 2016], including video [Pavel et al., 2016], highlighting the ambiguity and

inaccuracies that hinder such communication. This ambiguity can emerge due to the lack of a shared professional language [Phalip et al., 2007], the lack of a precisely framed scope for a discussion, and other challenges in communication [Phalip et al., 2009]. The problem intensifies in remote settings, where communication is mediated through technologically limited channels such as text chat, emails, and video calls [Pavel et al., 2016; Phalip et al., 2009].

This ambiguity resembles a central concept of social constructivism in science and technology studies: "interpretive flexibility" [Collins, 1981; Doherty et al., 2006; Meyer and Schulz-Schaeffer, 2006]. Interpretive flexibility posits that scientific knowledge and technological artefacts are always open to interpretation, which is shaped by social and cultural factors such as interests, values, and the larger socio-political context. To achieve consensus on knowledge claims and ensure that scientific knowledge can be effectively communicated and used, interpretive flexibility must be limited [Collin, 2011; Collins, 1981]. As such, possible interpretations by different parties must be limited through so-called closure mechanisms, including social closure mechanisms and the use of "rhetorical and presentational devices" [Collins, 1981, p.5] such as metaphors, analogies, and visualisations. Similarly, the interpretive flexibility of media must be addressed to alleviate communicational challenges in its production. For example, to facilitate discussions, creative workers can produce and present tailored versions of their work that highlight certain elements or aspects of the product, thus limiting possible interpretations [Feng et al., 2023; O'Leary et al., 2018; Retelny and Hinds, 2016].

In face-to-face settings, collaborators can supplement verbal communication with non-verbal cues, such as gestures, to help coordinate collaborators' actions and develop a shared understanding of media [Deacon et al., 2019]. Collaboration may also benefit from externalising creative ideas, which can be facilitated by interactive systems such as StoryCrate [Bartindale et al., 2012]. As noted above, when deployed on a filming set, it allowed participants to edit footage on the set and continuously visualise and share creative ideas with collaborators.

To address the lack of non-verbal cues in remote settings, Pavel et al. [2016] proposed a solution in their prototype VidCrit [Pavel et al., 2016]. They enhanced the reviewing process by allowing reviewers to record video feedback, thereby enabling the incorporation of non-verbal elements such as tone of voice and gestures. This offered editors a more nuanced understanding of the feedback, capturing not only the content of the reviewer's message but also the intent and sentiment behind it [Pavel et al., 2016]. Additionally, this prototype allowed reviewers to highlight elements of the video by drawing on the video player, providing spatial references akin to indexical referencing in a face-to-face setting. This idea has also been implemented in applications such as Frame.io,

which permits reviewers to draw on video players and leave comments attached to certain points of the timeline.

Reviewing processes in a remote setting can also be improved through guiding and scaffolding, which implies using temporary frameworks or guidelines to help individuals accomplish tasks [Bartindale et al., 2016b; Deacon et al., 2022]. Equipped with predefined principles and rubrics, even novice reviewers, who do not have education or experience in the field, are able to produce valuable feedback comparable to that produced by experts [Luther et al., 2015; Yuan et al., 2016]. Similarly, non-professional creators involved in film production can benefit from the utilisation of instructions, templates, and guides that embody the crucial information required for production. As such, they can scaffold creative, logistical, and technical requirements to ensure the quality of the product [Bartindale et al., 2016b].

2.8 Designing for Collaborative Remote Work

The early years of computer-supported work were marked by an endeavour to expedite and streamline traditional office tasks through technological interventions: office automation. The focus of office automation was primarily on the effectiveness and efficiency of computer systems to address business needs. As Hammer and Sirbu [1989] noted, "[t]he real payoffs of an office automation system must be measurable in business terms" [Hammer and Sirbu, 1989, p.40].

The subsequent shift towards CSCW marked a new understanding of the dynamics of work processes. The fields of CSCW and HCI have long been focused on the study of collaboration, identifying various modes of communication, such as synchronous and asynchronous [Lee, 2017] as well as co-present and remote participation [Neumayr et al., 2018], and exploring the design of collaborative software or groupware [Schmidt and Bannon, 1992].

The term "groupware" was introduced by Johnson-Lenz and Johnson-Lenz [1998] in the late 1970s to early 1980s and refers to a class of software applications specifically designed to facilitate collaboration and communication among groups of people working together on a shared task or project. Today, terms like "collaborative software" and "team collaboration tools" are often used interchangeably with "groupware" and encompass a wide range of collaborative tools and technologies [Bolstad and Endsley, 2003]. Unlike single-user software, groupware introduces new work for users, as it requires the coordination of workflows with collaborators [Grudin, 1988, 1994; Schmidt and Simone,

1999] and the management of such systems [Feng et al., 2023]. In other words, workers must invest additional effort to perform their tasks.

More than three decades ago, Schmidt and Bannon [1992] suggested focusing on "cooperative work," a term they used to describe mutually dependent work requiring coordination and collaboration [Schmidt and Bannon, 1992]. Their understanding of "cooperative work" was derived from a sociological perspective, considering cooperative work not only in terms of the tasks to be completed but also considering the social relationships, norms, and structures that shape how people work together. This perspective underlined support requirements for cooperative work arrangements, which comprise the technical and organisational tools, systems and resources, and cultural norms needed to facilitate cooperative work.

In recent years, the conversation has broadened to include other concepts, such as the "ecology of tools" [Bødker and Klokmose, 2012; Guribye and Nyre, 2017], which describes the interconnected relationship between the tools used in a particular field and their users. This perspective acknowledges these tool ecologies' constant evolution as professionals introduce and adapt new technologies.

CSCW scholars have developed various scientific frameworks that address the challenge of supporting remote collaboration in software systems by enabling document and file sharing, implementing information exchange, and providing awareness. In the following, I describe the groupware design frameworks most relevant to this project as well as the concept of awareness.

2.8.1 Groupware Design Frameworks

In a study of reasons behind the failure of many collaborative systems, Grudin [1988] identified three contributing factors that appear to be common in such systems. First, there is a disparity between those who perform the work, specifically the additional work, and those who benefit from it. For example, if the system requires a worker to record notes to keep track of executed tasks, the beneficiary of such a system is the manager who needs such recordings, not the worker. Second, the designer who develops collaborative systems might be unable to empathise with various users, misunderstanding their needs and neglecting their perspectives on workflows and processes. Third, the difficulty of evaluating collaborative systems might be underestimated when compared to single-user applications. This underestimation results in insufficient user tests, field observation hours, and other evaluation activities, thus failing to capture the intricacies of collaboration [Grudin, 1988].

Similarly, Ellis et al. [1991] noted that groupware interface design should consider social factors, such as group dynamics and organisational structure. They suggested that social scientists and users should play a role in developing such systems. Moreover, as collaborators in different roles use the same software differently, such software must be "flexible and accommodate a variety of team behaviours and tasks" [Ellis et al., 1991, p. 45].

Furthermore, Cockburn and Jones [1995] analysed the problems that cause collaborative systems to fail and formulated four principles of groupware design. Specifically, they emphasised that collaboration implies additional effort imposed by multiple factors in a remote collaborative environment. For example, the technical limitations of information and communications technology tools can hinder communication between remote parties. Thus, information transferred from one participant to another might be inaccurate and incomplete [Hollan and Stornetta, 1992]. Like [Ellis et al., 1991], Cockburn and Jones [1995] note that the lack of flexibility in software systems introduces constraints to users, who might seek ways to work around those system-imposed restrictions, such as by employing paper notes when they cannot take notes in the software. The lack of integration between various tools also introduces additional effort to users who must switch between single-user and multi-user tools. Additionally, the need to learn (and relearn) different tools and their interfaces might discourage users [Cockburn and Jones, 1995].

To address such challenges, Cockburn and Jones [1995] proposed four design principles. The first principle, "maximise personal acceptance", suggests encouraging specific users to adopt collaborative systems by addressing UI design issues through various strategies. For example, designers can solve the problem of learning and remembering interfaces of individual systems by exploiting similarities in various software tools and using design guidelines.

The "minimise requirements" principle suggests reducing the personal costs of additional actions that a user needs to perform in a collaborative environment, such as actions related to communication and coordination.

Similarly, the "minimise constraints" principle aims to avoid inflexible and constraining ways of using the software. This principle suggests allowing groups of collaborators to develop protocols governing their collaboration. As noted by Neuwirth et al. [1990], the "premature" definition of users' roles in a collaborative system might lead to various problems in the process of use. Thus, the system should not be rigid and enforce a specific work model or workflow but rather provide space for users to develop the workflows most suited to them.

The last principle, "external integration," suggests that designers should consider the

system's role in a broader work environment. Maximising integration with other systems helps users by reducing their transitions between tools in the course of work [Cockburn and Jones, 1995].

Gutwin and Greenberg [2000] introduced a conceptual framework known as the "mechanics of collaboration", which delineates the fundamental low-level actions that collaborators must undertake to successfully complete tasks within a shared workspace. These mechanics, identified through an analysis of previous research and literature, encompass explicit communication, consequential communication, coordination of action, planning, monitoring, assistance, and protection. The authors posited that many usability issues encountered in collaborative software systems are often a result of inadequate support for these essential mechanics. They proposed evaluating these mechanics against three criteria: effectiveness, efficiency, and satisfaction. To facilitate this evaluation, the authors recommended the use of discount usability techniques, such as heuristic evaluation and task-based walkthroughs, which are quick, low-cost methods for identifying usability issues in a UI [Gutwin and Greenberg, 2000].

It is important to note that, to a significant extent, these studies address the organisational context of collaborative work and focus on the acceptance and adoption of collaborative software within organisations. However, it is difficult to draw specific design implications from these studies (i.e., what specific features, UI elements, and functionality should be implemented in the software to support collaboration).

2.8.2 Awareness

Awareness is a prominent concept in the CSCW field and refers to the understanding that individuals and groups have about their environment, tasks, and activities within a collaborative setting. Awareness is crucial for designing systems and tools that support effective communication, decision-making, and coordination among group members [Bolstad and Endsley, 2003]. It provides a context for users' activities [Dourish and Bellotti, 1992] and enables them to manage tasks effectively, make informed decisions, and maintain a sense of group cohesion, even when working remotely or asynchronously.

Dourish and Bellotti [1992] suggested that awareness of collaborator's work should be not only active, that is, requiring users to perform actions to inform each other of the work they are performing, but also passive. The passive approach presumes that information regarding changes in the shared workspace is gathered automatically and presented to a user when needed.

In the context of the HCI and CSCW fields, two specific types of awareness are often discussed: *situational awareness* and *workspace awareness*.

Situational awareness originates from the aviation and military operations fields and has been adopted in other domains, including HCI and CSCW. In addition to focusing on the user's understanding of their current situation, it necessitates the ability to anticipate future system states and to account for the user's goals [Endsley, 1995].

Workspace awareness focuses on an "up-to-the-moment understanding of another person's interaction with the shared workspace" [Gutwin and Greenberg, 2002, p. 417]. As such, Gutwin and Greenberg [2002] developed an extensive workspace-awareness framework that addresses collaboration in real-time distributed systems. The framework suggests that the information that comprises workspace awareness can be presented as a set of categorised elements, each answering a specific question about the state of the shared workspace. For example, "*Who* is present?", "*What* are they doing?", and "*Where* are they working?". The framework also suggests including awareness of the past, or historical awareness, that addresses questions such as "*How* did this operation happen?", "*When* did it happen?", and "*Who* was here and *when*?" [Gutwin and Greenberg, 2002].

By providing this information to the users, a collaborative system enhances collaborators' coupling, or the degree to which people are working together [Salvador et al., 1996], simplifies communication, helps to coordinate actions, and allows collaborators to predict each other's actions. Consequently, supporting workspace awareness significantly enhances the usability of collaborative software. This augmentation can streamline communications efficiency and foster a more effective collaborative environment [Gutwin and Greenberg, 1999].

Workspace awareness elements can be implemented in software design in the form of features, widgets, and various UI elements, such as "telepointers" and "telecarets", or the mouse and text cursors of remote participants [**?**]. These elements are commonly found in web-based collaborative software such as Google Docs and some of the collaborative NLE applications discussed above. For example, Motionbox demonstrates the presence of other users in the shared workspace with avatars and indicates the actions of other users with telepointers and "artefact marking", which highlights selected clips and timeline elements remotely.

2.9 Trends in Video Production Technology

Video production, as a craft or vocation, has a rich history spanning over a century [Buckingham et al., 2007]. Throughout this period, the tools and technology employed in the field have experienced significant transformations, continually shaping how practitioners approach their work. Today, video production encompasses a diverse array of individuals, ranging from amateurs exploring the medium as a hobby to highly skilled professionals who have dedicated their careers to mastering the intricacies of the craft. This diversity is reflected in different production approaches [Kirk et al., 2007; Owens and Millerson, 2012] and an assortment of video production tools and instruments [Guribye and Nyre, 2017].

Video production has become more accessible due to technological advancements [Juhlin et al., 2014b; Perry et al., 2009; Yarosh et al., 2016], such as mobile and internet technologies and the adoption of automation. These technological trends have significantly affected the work practices of both professional video workers and hobbyists. The mitigation of technological challenges allows video amateurs to shift their focus from technical aspects of a video, such as camera movement, to its content, including story and narrative [Weilenmann et al., 2014]. Thus, the boundary between professional and amateur video creators has become increasingly blurred [Webb et al., 2016].

2.9.1 Increasing Accessibility of Production Tools

Over the past two decades, camera phones, which are mobile phones with cameras capable of filming a video of decent quality, have become widespread. Additionally, video-editing software has become available for mobile phones and tablets [Feinberg et al., 2016]. This trend of video devices becoming more accessible and portable is not new. Since the 1920s, amateur and semi-professional videographers and filmmakers have existed alongside highly skilled professional video workers. The introduction of camcorders — portable video cameras — created opportunities for amateurs to participate in professional video production. For example, an amateur with a video camera could capture an ongoing event and provide footage suitable for breaking news coverage on TV [Buckingham et al., 2007]. Nonetheless, there are notable differences between professionals and amateurs in terms of video quality. For example, Guo et al. [2013] identified several predictors that help distinguish amateur and professional videos, such as camera motion and structure.

For professionals in the TV and video industries, such accessibility has also created

changes. For example, journalists who used to focus on writing stories and reporting are now also expected to film footage with the use of new tools and applications [Phillips et al., 2009]. In this new practice of "mobile journalism" (*MoJo*) [Salzmann et al., 2020; Westlund and Quinn, 2018], journalists are acquiring new video skills such as filming, editing, and live-streaming, and they are responsible for performing such skills in their organisations [Kumar and Haneef, 2018; Nygren, 2014].

2.9.2 Remote Work

Concurrently, the same period witnessed a surge in the prevalence and acceptance of remote work [Baker et al., 1999; Cousins et al., 2007; Koehne et al., 2012]. However, while some organisations embraced remote work and the "work-from-home" approach [Karis et al., 2016], others, such as TV production companies, largely have not adopted remote work until recently [Coche and Lynn, 2020]. Baker et al. [1999] identified various factors that influence the adoption of remote collaboration, such as security concerns, productivity, and cost considerations. Among the advantages of remote work, they emphasised the ability to work with distant companies and talents and to compete on a global scale.

However, remote work requires significant coordination between parties. For example, Baker et al. [1999] noted that this obligation could overwhelm collaboration to the point of it being cancelled. Likewise, Koehne et al. [2012] argued that remote work is inherently social, with individual remote workers developing various coping strategies to manage tasks and maintain visibility within their teams. These strategies help ensure that remote team members can effectively collaborate and contribute to the shared goals of the organisation, even when working from a distance.

In 2020–2022, the adoption of remote work increased due to social-distancing regulations associated with the COVID-19 pandemic. The pandemic, in general, catalysed innovation across various sectors, driving the development of new technologies, business models, and collaborative approaches [Heinonen and Strandvik, 2020]. It impacted both media consumption [Casero-Ripolles, 2020] and production, prompting media production companies to adopt remote workflows [Coche and Lynn, 2020].

2.9.3 Automation in Video Production

One way of mitigating the technical challenges of video production is delegating tasks to computer systems and algorithms, or automation. In academia and industry, automation

is suggested for use in multiple specialised systems, which are often powered by artificial intelligence (AI) technologies. Such systems address various aspects, stages, and parts of video production, such as automatic story writing [Sirén-Heikel et al., 2019], directing and capturing [Adams and Venkatesh, 2005], camera operation [Herrmann et al., 2020], live TV production [Barlow, 1972, 1979; Freilich and Meyer, 1963], editing [Chi et al., 2013; Girgensohn et al., 2000; HUA et al., 2003], creating summaries and highlights [Shipman et al., 2003, 2008], subtitling [Soe et al., 2021], and broadcasting [Barlow, 1979; Richards, 2010].

From the HCI and CSCW perspective, automation involves utilising computer systems to perform routine work typically handled by humans, thus reducing their workload and enabling users to focus on more creative, complex, and rewarding activities. However, the replacement of human workers with automated systems has several implications. Automation requires formalising work procedures and practices [Gerson and Star, 1986], leading to an interaction with automation systems akin to end-user programming, as suggested by Fröhlich et al. [2020]. This interaction becomes less continuous, allowing users to dedicate their time to solving less trivial tasks. Nonetheless, Bainbridge [1983] highlights an irony: operators lacking regular practice of the automated task might lose the skill needed to intervene and correct an error in case of an automation system breakdown.

The "ironies of automation" [Bainbridge, 1983] further illustrate potential issues from implementing automated systems in video production or other domains. These include increased complexity, workload shifts, diminished creativity and innovation due to the formalisation of tasks, loss of situational awareness, and ethical considerations.

Additionally, employees might view automation as a threat to their job security and consequently resist its implementation and adoption [Lindén, 2020]. However, research suggests that people frequently overestimate the adverse consequences of automation and underestimate its positive impacts [Autor, 2015]. Empirical studies indicate that automation often does not replace human labour; instead, it complements and augments it [Arntz et al., 2016; Lindén, 2020].

Thus, adopting automation requires that potential benefits and drawbacks be balanced, which highlights the importance of understanding the nuanced relationship between automation and human labour in various domains, including video production.

In summary, this chapter laid the foundation for the research presented below by providing an overview of various topics pertinent to collaborative video editing and detailing current tools utilised in this field. It highlighted that the realm of collaborative video editing is not only multifaceted but also an intricate confluence of different factors,

including technology, creativity, coordination, and human interaction. These interconnected and mutually impactful aspects contribute to the domain's complexity. Having established the nuanced landscape of collaborative video editing, the research methodology employed in this thesis will now be discussed.

Chapter 3

Research Approach and Methods

Collaboration in video production occurs in various forms, such as face-to-face conversations in the field, closed meetings, phone calls, emails, messages on social media, and handwritten documents. To study and design for collaboration, I employed a user-centred design approach [Abras et al., 2004; Norman, 2013] that presumes the involvement of end-users, that is, professional video workers. Through interviews and design workshops, we consulted the users about their needs and allowed them to participate in the design process.

The most promising method to study collaboration seemed to be semi-structured interviews, which would allow me to investigate the domain without spending large amounts of time observing collaboration. Additionally, participation or observation would be largely impossible for two reasons. First, video workers and other parties, such as their clients, often consider the work sensitive and private and do not welcome external participants. Second, a significant amount of collaboration occurs digitally through messages and emails. Such collaboration is difficult to observe. Thus, I had to rely on the retrospective accounts of participants.

The second data collection method I used in this research was design workshops. During the workshops, invited video professionals and designers created various concepts to support collaboration in the video editing process.

Study	Number of Participants	Collection and Analysis Methods
Article I	13	Interviews, Thematic Analysis
Article II	12	Interviews (online and offline), Thematic Analysis
Article III	20	Workshops (online and offline), Content Analysis

Table 3.1: Methods Summary

For data analysis, I used thematic and content analysis methods. Both methods aimed to extract meaning from texts such as interview transcripts and produce high-level categories that could be used to identify and address various aspects of collaboration in video production. The summary of methods is presented in table 3.1.

In the following, I explain the data gathering and analysis processes in more detail as well as the challenges that I had to overcome while conducting this research.

3.1 Sampling and Participants

The overall goal of the sampling was to find people who had significant experience in video editing or other forms of post-production of any type of video, from short clips to feature films, and who had been working in teams. In this search for participants, I needed to rely on various sources and both personal and professional connections.

In this research project, I mostly employed convenience sampling [Marshall, 1996] with some elements of snowball sampling. For the three studies, my colleagues and I used a total of five sources to recruit participants:

1. Professional connections through Vizrt. Vizrt is an international software producer for TV studios and video production companies with global clients. Thus, they were able to invite some of their clients to participate in interviews and design workshops. Participants were recruited through Vizrt in all three studies. In the third study, we also recruited some Vizrt employees, specifically UX designers, to participate in design workshops.

2. Professional connections through the University of Bergen. At the university, there are professional and amateur video workers who produce educational video material, such as documentaries and video lectures. I recruited these participants at the beginning of the project, and their insight helped me understand the complexity of the video domain. In the third study, we also recruited three master's students who were studying HCI and CSCW to participate in a pilot workshop. Later, these students helped to organise and run the second workshop.

3. Social media. In the third study, five participants were recruited through social media posts for a workshop. I also attempted (and failed) to recruit participants through social media for the first study.

4. Freelance websites. In the third study, my colleagues at Vizrt and I recruited six

freelancers via *upwork.com*. The participants were from various European countries and had diverse backgrounds.

5. Personal networks. During the project, I recruited a total of five participants through personal connections. In two cases, participants also helped me recruit their former and current colleagues. Three participants were also recruited through personal connections by my co-authors.

The sampling was not random: we targeted people with specific experiences, roles, and backgrounds. We sought those with experience in video production in teams or organisations. We also attempted to focus on professional video workers, such as TV workers and video-editing crews.

In total, 44 participants were involved in the research. One person participated in two studies, the first and the second. Overall, participants had significantly diverse experiences and backgrounds. Except for three students who participated in the pilot workshop in the third study, all participants had some experience in video production, which ranged from YouTube videos and short commercials to feature films and TV shows. Four participants in the first study were not professional video editors. However, their experience in video production was connected to their occupation. Most participants had experience acting in multiple roles in the video-production process, such as writing, editing, camera operation, directing, and producing. In many cases, participants had experience in both the TV and video industries. Our participants were largely from Europe, although one was from the Middle East, one was from Asia, and two were from the Americas.

3.2 Data Collection

In this section, I describe the data-gathering process and its important aspects: working with a professional language, sensitive topics, and the challenges and opportunities the COVID-19 pandemic introduced into the project.

3.2.1 Process

In Study 1, data were gathered through semi-structured interviews and follow-up interviews via face-to-face meetings, email, and social media. I conducted nine interviews with 11 participants. Two interviews were with pairs of participants. Two interviews

were contextual: I conducted them in video-editing studios, and participants demonstrated many of the minute activities they referred to. Three interviews were followed up by subsequent interviews in face-to-face meetings, and four interviews were followed up with conversations over email and social media.

The first three interviews were focused on the topic of collaboration during video production. Specifically, I asked about the modes and forms of collaboration: does it occur in person? Does it occur online? What kind of questions are being resolved by a collaborative effort? Beginning with the fourth interview, it became apparent that there is a specific topic relevant to different kinds of video production: reviewing. Consequently, I rearranged the interview guide and reformulated some questions to accommodate this new focus of the study. Previous interviewees were revisited through personal meetings and emails to address new questions about reviewing practices.

In Study 2, data were gathered through face-to-face and online semi-structured interviews and informal communications. The initial aim of the study was to investigate how video production professionals adapted to the challenges of the COVID-19 pandemic. After the first two interviews, the preliminary results suggested that, among video professionals, those who were affected the most worked for TV companies. Thus, the topic of the study was refocused and narrowed to the TV industry. In the interviews, we asked about the recent changes that had occurred to workplace practises due to social-distancing policies. We focused on how the workflows and tools used in video production had changed. We also discussed with participants how these changes had affected various aspects of their work and the future of the TV and video industries.

In Study 3, data were gathered through five design workshops, both online and in-person. The workshops were organised in two rounds. In the first round, two workshops were held at the University of Bergen, and these were partially videotaped and transcribed. The participants sketched various design ideas on paper, and the resulting sheets were collected. These sketches were used in the second round of workshops as inspiration points for the participants. In the second round, three workshops were held online via Zoom[1]. The workshops were video recorded and transcribed. During the workshops, the participants used the balsamiq.cloud[2] application, which allows design ideas to be sketched online. After the workshops, the sketches were exported. During the analysis, some sketches were redrawn to better present the design concepts.

[1]https://zoom.us/
[2]https://balsamiq.cloud/

3.2.2 Challenges

The process of gathering data did not always go smoothly, and it was not without hurdles. The following paragraphs briefly describe the challenges of this research.

First, it was not easy to gain access to the field. The primary challenge when involving professionals in the research was their availability. Specifically, not all the potential participants could dedicate time to an interview or participate in a workshop. In some instances, my initial inquiry regarding participation was met with no response, which could be attributed to a variety of factors, including workload, time constraints, or a lack of interest in the study. In other instances, while potential participants expressed interest in the study, their busy schedules and pressing deadlines prevented them from committing time for participation. I found this to be a reminder of the demanding nature of the field of video editing, where time is often scarce.

Another obvious challenge was the use of professional language: video workers often employ terms that may be misinterpreted by others. In the first study, we presented quotes and described situations that reflected similar cases of misunderstandings between video professionals and their clients who are less video literate. During this research, I needed to familiarise myself with professional terminology by reading books, articles, and blog posts on related topics and by speaking to professional video editors who were kind enough to explain certain terminology during the interviews and informal talks.

In some cases, participants considered the discussed topics to be sensitive. Specifically, in the first study, some participants hesitated to discuss conflicts between video professionals and clients which they had observed or participated in. The unease stemmed from the implication that admitting to such conflicts might reflect poorly on the professional conduct of those involved. This encompassed issues such as the use of inappropriate language in professional settings and overall unprofessional conduct. Navigating these conversations required a delicate balance between respect for participants' comfort levels and the need for honesty to attain an accurate understanding of the phenomenon. I underscored the importance of these discussions for the research, reassuring participants that the focus was on gaining insights into the realities of the video production industry and not on judging individuals or their conduct.

Similarly, during the interviews for the second study, some participants expressed concerns about their and their colleagues' job security. They were apprehensive that our research findings might lead TV companies to downsize their workforce and eliminate certain jobs. These concerns were so pronounced that one participant continually returned to this topic throughout the interview, revealing a deep-seated anxiety about

potential negative consequences stemming from our research.

Thus, we anonymized the interview transcripts as soon as they were produced and eliminated or obfuscated any information that could potentially identify the participant, their specific role, the company they worked for, and even their country of residence. The latter appeared to be an appropriate measure as there are very few TV companies in smaller countries.

This experience highlighted the importance of trust and transparent communication between a researcher and participants as well as the necessity of approaching sensitive topics respectfully and clearly. It also underscored the need for researchers to remain flexible and responsive to the evolving needs and concerns of participants. Our commitment to participant anonymity and privacy was not only a legal obligation but also a critical element that facilitated the honest and open dialogues that significantly contributed to this research.

During the COVID-19 pandemic, the data collection process had to be changed. Instead of face-to-face meetings, we needed to rely on services like Skype and Zoom to conduct fieldwork. This situation proved somewhat beneficial as digital communication became normal [Arshad, 2020], and we could employ remote participants worldwide. This geographical freedom enriched our data pool by bringing in diverse perspectives and experiences.

The pandemic also allowed us to study the active transformation of some work practices. We treated the pandemic as a world-scale breaching experiment [Garfinkel, 1967], a research method designed to disrupt social norms to study the resulting behaviours, that distorted the normal order in the video and TV industries. It forced organisations to adapt to new challenges by adopting "new" technologies and workflows, where "new" refers to mobile devices and tools like Skype and Zoom, already widely used in other domains but rather novel in a professional video and TV production context. This shift provided a unique opportunity to study work practises in a state of flux, offering unprecedented insight into the processes and challenges of adapting to sudden changes.

In retrospect, while the pandemic unquestionably presented substantial obstacles, it also acted as a catalyst, enriching our study and shedding light on unexpected aspects of collaborative video editing in a remote context.

Furthermore, the pandemic increased the relevance of our research not only from an academic perspective but also in the eyes of our study participants. The sudden transition to remote work meant that the issues we were investigating were relevant to the daily challenges these professionals faced. This allowed the participants to perceive our study

as relevant and valuable to their current situations.

3.3 Analysis

Overall, the collected data can be categorised into four groups:

1. Interview audio recordings and notes. The audio recordings were transcribed and anonymised. The transcripts served as a foundation for analysis in the first and the second studies.
2. Emails and private communications via social media. This data contained answers to follow-up questions and was attached to the transcripts for analysis in the first and the second studies.
3. Sketches of design ideas. During the workshops in the third study, the participants created sketches both digitally and physically. Sketches on paper were photographed or scanned.
4. Video recording of workshops. These recordings were partially transcribed and used in analysis together with the sketches in the third study.

To identify core patterns and extract meanings from the data, we used qualitative content analysis[Patton, 2002, p. 452] and thematic analysis [Braun and Clarke, 2006] methods.

3.3.1 Thematic Analysis

Braun and Clarke [2006] identified six steps of thematic analysis, which we followed. First, we familiarised ourselves with the data by reading the transcripts and searching for patterns in the text. Second, we generated initial codes. Third, we searched for possible themes that would consistently connect the codes. Fourth, we reviewed the themes by returning to the transcripts and searching for the context in which the themes occurred. Fifth, we defined and named the themes. Sixth, we documented the results in an organised manner by writing the papers.

The analysis was conducted in iterative collaborative sessions. As the responsible author and the topic owner of this research, I acted as a responsible coder, conducted the initial coding, and presented the intermediary results to my colleagues. Together, we refined the coding, defined the themes, and identified the most interesting and significant findings.

The iterative nature of the analysis process allowed us to rearticulate the meanings we extracted from the empirical data set. In this process, specific research questions and answers arose in the interaction between the corpus and previous research [Krippendorff, 2018]. As such, we could refocus our efforts and formulate follow-up questions if needed. In such cases, we attached the new transcripts to the corpus and refined the coding scheme.

3.3.2 Content Analysis

In the third study, we gathered data of two types: the transcripts of discussions between participants in the workshops and the sketches of design concepts. We employed content analysis to answer the research questions of why it is difficult to design for remote collaborative video editing and how to achieve it. Content analysis is a research technique that aims to make valid inferences from data to their context [Krippendorff, 2018]. Thus, we aimed to analyse not only "what" was said but also "who" said it and "why". In addition to analysing text, we considered design concepts expressed in both physical and digital sketches.

We sought patterns in texts that would explain the reasoning behind design concepts. More specifically, we aimed to answer the following questions: why is this concept beneficial? What are the drawbacks of this concept? How should this concept work?

We coded the data in three steps. First, we searched for the most apparent aspects of the design concepts mentioned by the workshop participants. Second, we aimed to identify motivations behind the design concepts and possible challenges that might emerge in their implementation. Third, we sought patterns and overarching themes that described such challenges and motivations. We inferred the importance that workshop participants assigned to ideas and identified patterns based on the "volume" at which a subject was discussed (i.e., the number of mentions and length of discussions). Ultimately, we produced a list of high-level categories that place design concepts in the context of their implementation and use and address various aspects of collaboration in video editing.

By analysing both verbal discussions and sketched design concepts, we developed a holistic understanding of collaborative video editing. The next chapter elaborates on and discusses the findings.

Chapter 4

Findings and Discussion

To answer the research question of this thesis, "*How can remote collaboration be supported in video editing?*", three studies were conducted. The first study investigated current collaborative practices in video editing and scrutinised reviewing processes and interpersonal communication between collaborators. The second study examined video work from an organisational perspective. It investigated how organisations adapted to social-distancing regulations and explored relations between technology trends and innovations and video production work. The third study focused on supporting collaboration in video editing with design concepts and investigated challenges designers must overcome.

The first section summarises and discusses the key findings of the studies. The second section briefly describes challenges in collaborative video work and examines the strategies and mechanisms workers employ to overcome them. A more detailed and structured description can be found in the first paper [Okopnyi et al., 2020]. The third section outlines design paths and concepts to support collaboration in video-editing software. The fourth section identifies and discusses the challenges designers face when developing collaborative video-editing software. The fifth section considers changes in video work caused by the adoption of remote work and automation and speculates on the future of video-production practices and workflows. The sixth section identifies three potential approaches for designing collaborative tools, which were synthesised from the results of this study, previous research, and existing non-linear editing (NLE) software. The final section discusses the limitations of this research.

4.1 Results and Contributions

In total, this research project encompassed 23 interviews with 25 respondents and five design workshops with a total of 20 participants.

In the first study, we closely examined practices in collaborative video work. We focused on negotiations between video editors and other parties, such as producers, directors, and clients, as well as reviewing practices, wherein the editor seeks feedback and approval from collaborators. Reviewing has been addressed in previous design studies [Pavel et al., 2016]. In our study, however, we unpacked this practice to understand its social aspects. Specifically, we considered mechanisms and strategies that help reduce uncertainty and achieve agreements between parties involved in the post-production process.

We identified four categories that characterise editorial uncertainty. Specifically, we demonstrated that it is difficult to articulate an idea for a video in words; the difference between editing and reviewing contexts can introduce confusion, and it might persist in a co-present environment. It can also be difficult to refer to various video elements and time. To mitigate such uncertainty, editors employ various mechanisms, including references to written statements, documents and icons; closure by rhetorics of practicalities; and organisational closure mechanisms.

Based on our findings, we proposed three design paths for video-editing software development. First, we suggested using *scaffolding* [Luther et al., 2015; Wood et al., 1976; Yuan et al., 2016], a preset structure that supports participant collaboration. Second, we recommended adding tools and features for *iconic referencing* [Peirce and Buchler, 1902] in video-editing software. Third, we propounded *suggestive editing*, an inclusion of reviewers into the editing process via tailored video-editing tools that allow not only comments but also suggestions for specific edits to the video.

The second study focused on how social-distancing regulations changed the work of TV companies during the COVID-19 pandemic. In this study, we focused on adaptations made on an organisational level. We conducted 12 interviews with TV and video workers to determine how they and their organisations personally adapted to social-distancing regulations. We identified two main themes, automation and redistribution, when conceptualising the socio-technical changes in TV production.

To comply with social-distancing regulations, TV companies were forced to adopt remote workflows. Such workflows involved using general-purpose applications, services, and devices, such as Skype and mobile phones; relying on internet infrastructure; and finding ways of making productions leaner and less dependent on traditional work processes

and organisational structures. Some of these changes in production practice might be irreversible and denote a major trend in the TV and video production industries. As some respondents observed, commercial organisations will not be eager to return to the "old ways" of conducting work because the "new ways" are less expensive, and such a claim has been supported by previous research [Albiniak, 2022]. The implications of such changes include a shift in the definition of broadcast quality and the transformation of TV literacy, or a shift from the hyper-specialisation of TV workers to multiskilling.

In the third study, we conducted a series of design workshops to develop design ideas and concepts for supporting collaboration in video editing. We limited and focused the ideation process during the workshops by introducing both the concept of workspace awareness [Gutwin and Greenberg, 2002] and schematic NLE representation as starting points. The workshops resulted in multiple concepts created with pen and paper or digital sketches and prototypes. We developed some of the sketches further to present them in the article. During the analysis, we also identified specific categories describing advantages and drawbacks of the concepts from the participants' perspectives.

The first category highlights the challenging nature of video editing, which requires intense focus and avoiding distraction. The second category suggests extending design efforts beyond the scope of NLE software and including various activities in addition to video editing, such as filming and reviewing processes. The third category addresses the time-based nature of video and time representation in NLE software.

Our findings mirror the shift in the computer-supported cooperative work (CSCW) field from focusing on efficiency-oriented considerations to a holistic understanding of the interplay between tools and people as well as the subtle complexities of their interactions.

In [Okopnyi et al., 2021], we showcased how normally passive participants evolved during the co-creation of video content. In our analysis, we employed the notion of "redistribution of work, control, and skills" to characterise what is traditionally seen as an example of the "division of labour". Importantly, this redistribution encapsulates more than simply a shift in work: it also involves a transformation in roles and responsibilities. The implications of such redistribution also extend to the "ecologies of tools". They do not merely support existing work arrangements and team structures but can also instigate the emergence of new team compositions and workflows. For instance, a journalist might assume the role of an editor or the interview participant might become an active player as a camera operator.

In [Okopnyi et al., 2021], we showcased how normally passive participants evolved into the co-creation of video content. During our analysis, we employed the notion of "redistribution of work, control, and skills" to characterise what would traditionally be seen

as an example of the "division of labour." Importantly, this redistribution encapsulates more than just a shift in work. It also involves a transformation in roles and responsibilities. The implications of such redistribution extend to the "ecologies of tools" as well. They do not merely support existing work arrangements and team structures but can also instigate new team compositions and workflows to emerge, for instance, in situations where a journalist assumes the role of an editor or when the interview participant becomes an active player as a camera operator.

Last, during our design workshops [Okopnyi et al., 2022], participants tasked with conceiving collaborative video-editing tools frequently voiced concerns beyond efficiency. Reanalysing the results illuminated various topics such as the sense of ownership and control over work; the privacy and confidentiality of work; and the convenience and comfort of the editing process, evident in time-saving measures, the elimination of redundant tasks, the ability to maintain focus, and the importance of working with familiar tools. Interestingly, the reanalysis also affirmed that even individuals with remarkably similar backgrounds could hold conflicting views on the functionality and features of collaborative tools, with some participants expressing concerns regarding design concepts which others welcomed.

What do these findings imply for designs supporting collaboration? To a significant extent, classic CSCW papers emphasise constraints, requirements, personal acceptance [Cockburn and Jones, 1995], adoption and benefit [Grudin, 1988], and awareness [Dourish and Bellotti, 1992]. We add to this knowledge by emphasising the specific needs of users, such as the sense of ownership of work and workspace, as we heard arguments for both allowing and forbidding collaborators to view, review, or participate in ongoing video editing. In sum, enabling collaboration can be seen as a double-edged sword. On the one hand, it promises enhanced productivity. On the other hand, it introduces challenges in coordination and potential conflicts over creative control. Consequently, software designers must navigate these opposing user needs. Traditional CSCW principles provide a foundation for such designs, but the nuanced needs of video editors add layers of complexity. Evidently, simply implementing collaboration tools is insufficient, and it is necessary to understand and acknowledge the unique challenges of collaborative video work in design.

4.2 Challenges of Collaborative Video Work

Effective collaboration among co-workers is a complex and challenging task that requires overcoming various communication barriers. While previous research has addressed more

general issues of designing for collaboration, such as acknowledging an organisational structure in software [Ellis et al., 1991], our research has identified specific challenges that arise when collaborating on video production.

In the analysis, we drew inspiration from science and technology studies and the concept of "interpretive flexibility" [Collins, 1981], which refers to how different actors or groups interpret and use an object, technology, or idea. The complexity of video is inherent in the media due to its richness: video can convey multiple layers of information through visual elements, motion, sound, timing, and other elements. Editors find such complexity challenging, and we used "interpretive flexibility" to conceptualise this challenge. Following the steps described by Collins [1981], we first demonstrated how interpretive flexibility manifests itself in collaborative video editing. Subsequently, we identified various closure mechanisms that video editors employ to "[limit] the potentially endless debate about possible interpretation[s]" [Collins, 1981, p.7] and achieve agreements.

The communication of ideas and concepts often requires additional context beyond what language can provide. As one of the informants noted, "our language is insufficient" [Okopnyi et al., 2020, p.120]. The professional language of video editing is complex, utilising unique terms, jargon, and assumptions. During pre-production, the idea for a video is often conveyed through a textual task description, presenting a challenge for participants who must briefly and artistically describe the desired look and feel [Bartindale et al., 2016b]. Thus, subsequent editing is marred with the uncertainty and vagueness of communication between editors and other participants, who may struggle to articulate their desires from the beginning of the video-production process.

Consequently, even when collaborating parties are well-versed in relevant jargon, they may nonetheless struggle to convey their thoughts effectively. Providing references is crucial for video reviewing, but it is often challenging to accurately identify where an issue is located in a video. Two common referencing approaches are referencing by time and referencing by objects, but both have associated ambiguities.

According to our interviewees, video editors generally prefer time-coded comments if other tools are lacking as these allow collaborators to provide feedback with references to a timeline. However, ambiguity can arise even with specific instructions, such as providing feedback at a certain time code. For example, a comment referencing a specific time code may not be clear if the video has multiple frames within that interval, resulting in confusion for the editor and potentially requiring further clarification from the reviewer [Okopnyi et al., 2020].

Alternatively, collaborators may reference objects or specific scenes in the footage to position their reviewing comments. However, even seemingly concrete and detailed de-

scriptions can be insufficient and lead to confusion. This is because objects or scenes may appear multiple times in the footage, making it difficult to identify the intended reference.

Furthermore, effective collaboration often demands extended discussions, which can be challenging when participants come from diverse backgrounds and lack familiarity with each other's professional language. This challenge persists even in co-present settings. During face-to-face interactions, collaborators can discuss changes and brainstorm solutions when disagreements arise. In this setting, editors can observe the reviewer's body language and tone of voice, which provide insight into their reactions and emotions towards the video draft. However, when communication is limited to text comments, editors can attempt to infer the reviewer's feedback by reading "between the lines" [Pavel et al., 2016, p.519] as it can be difficult to discern whether a video is satisfactory for the client based on a list of critiques received via email.

Video workers can ask collaborators to use exemplars, such as videos they like or find suitable. These exemplars manage expectations, reduce the open-endedness of the editing process, and offer producers a chance to reuse existing elements, such as fonts, effects, colours, and filters, instead of designing them from scratch.

Another factor that poses a challenge to collaborative video work is a lack of contextual information regarding the production process. Collaborators may not be aware of the current production stage or the procedures and practices that constitute video work at a given moment. Those unfamiliar with the video-production process may also be unaware of what work has been completed and what will be performed at a specific stage. Furthermore, reviewers may focus on small and irrelevant details when asked to review the work in progress and thus provide non-constructive feedback.

Editors use local practical constraints and rhetorical devices, such as establishing "deadlines" and the expected number of "iterations", to delimit collaborators. These devices draw on local circumstances (e.g., the availability of the editor) to reduce interpretive flexibility, and although such circumstances are always present, they are not always evoked.

In pre-production, artefacts such as contracts, meeting notes, storyboards, and screenplays can be used to specify content, style, and narrative, as well as time constraints and other demands, thereby limiting the scope of the production process. Other parties, such as clients, may attempt to extend the task during production, but documentation can be used to avoid unnecessary costs. Limiting the number of editing iterations and defining them in a contract can also prevent the process from being prolonged.

Organisational closure mechanisms restrict options for collaborative activity. When working with external parties, such as clients, the video production team can establish a certain organisational structure in which a single producer acts as a negotiator between the production team and the client, thus limiting the client's opportunity to participate in the project. Another strategy is to limit communication to a single person on the client side who has the mandate to make decisions.

Uncertainty also arises in the need to comprehend the potential viewer's context, as exemplified by a case where the client reported an absence of sound, attributing it to the poor quality of the video. The discrepancy occurred due to the editor's assumption that the client would use high-quality sound equipment or headphones while the client was using low-quality equipment [Okopnyi et al., 2020].

Additionally, rendering and sharing a video with collaborators can be an issue. The typical process involves doing so via a file-sharing service. However, existing reviewing software, such as Frame.io, does not integrate well with NLE software, which creates an additional step for video editors who must render and share a video with others for feedback. This process can be time-consuming and may result in poor video quality, which could impact the feedback received.

In exceptional circumstances, collaboration is impeded by participants' incapacity to work together due to external factors, such as the social-distancing restrictions enforced during the COVID-19 pandemic. In this period, video workers who were previously co-located were compelled to work remotely, limiting their communication to telephone and videoconferencing tools. Video workers also needed to delegate some work to other parties or redistribute work among themselves. For example, journalists were required to assume the roles of camera operators. Similarly, some TV guests independently managed their interviews, handled lighting, and operated cameras on their phones, among other tasks normally performed by TV-studio personnel [Okopnyi et al., 2021]. This redistribution of work was not limited to people. Some tasks, such as camera operation, were delegated to machines, that is, supported by automation.

These challenges and the strategies that may be used to address them provide a background for designing to support collaborative work.

4.3 Supporting Collaborative Video Editing

To support collaborative video editing, we suggested three design paths [Okopnyi et al., 2020] and several design concepts [Okopnyi et al., 2022]. The design paths relate to iden-

tified closure mechanisms that video editors employ to achieve agreements. They demonstrate possible high-level directions for collaborative editing software design, specifically NLE software. The concepts present concrete ideas, some of which can be implemented as features in existing software. The essence of these design ideas is two-fold. First, they aim to provide means of communication between collaborators. Second, they allow the redistribution of work and responsibilities.

4.3.1 Supporting Communication

Here, communication implies two things. First, collaborators convey information to each other directly via verbal and non-verbal methods, such as text messages and hand gestures. Second, communication implies support for different kinds of awareness.

In both academia and industry, there have been explicit attempts to design for the former. For example, the "VidCrit" application developed by Pavel et al. [2016] is a specialised software that aims to support communication between reviewers and video editors [Pavel et al., 2016]. Our research proposes several ways of supporting such communication between different parties. For instance, the *integrated reviewing tool* concept [Okopnyi et al., 2022] suggests implementing existing additional services (e.g., Frame.io) as an integral feature of NLE software. Such integration can reduce collaborators' workloads and personal costs [Cockburn and Jones, 1995] because users would perform fewer actions when asking for feedback or providing reviews. Specifically, with such an instrument in place, they would avoid rendering and uploading the video to a dedicated reviewing service. The acquisition of Frame.io by Adobe in 2021[1] further validates this design idea, signalling industry recognition of the importance of integrated reviewing tools within NLE software.

In addition to verbal communication, such as text messages, various means of conveying non-verbal information are needed. One such approach is *iconic referencing* [Okopnyi et al., 2020], where collaborators attach representative artefacts of their vision for different aspects of the video production, such as colour schemes, pacing, and montage style. This idea aligns with the findings of Pavel et al. [2016], which suggest that verbal communication alone is insufficient for effective collaboration. Therefore, incorporating non-verbal communication methods can aid in creating a more comprehensive understanding among collaborators.

Scholars have acknowledged the importance of supporting awareness in collaborative software for decades [Dourish and Bellotti, 1992; Gutwin and Greenberg, 1999, 2002].

[1] https://techcrunch.com/2021/08/19/adobe-buying-frame-io-in-1-28b-deal/

However, when considering collaborative video work and video editing, there are very few examples of such considerations in previous research. For instance, Bartindale et al. [2016a] suggested facilitating situational awareness on the filming set by introducing an interactive editing system that the filming crew could use to experiment with the newly shot video footage. In our research, we suggest supporting additional types of awareness. For example, awareness of current progress [Bartindale et al., 2012] can be supported through scaffolding mechanisms such as the *documentational scaffolding* design path [Okopnyi et al., 2020] and the *annotations* concept [Okopnyi et al., 2022]. We also suggest enabling workspace awareness through various UI elements and features. Some elements are specific to NLE software, such as *playheads*, while others are generic and can be found in other collaborative software, such as *colour-coded avatars* [Okopnyi et al., 2022] and telepointers [Heinrich et al., 2013].

4.3.2 Redistribution of Work and Responsibilities

Video production teams are often formed on a project basis and depend on its scope, type, theme, and other parameters. For example, some interviewees in [Okopnyi et al., 2020] had experience in projects ranging from small YouTube clips and advertisement videos to TV shows and feature films. Like Klokmose et al. [2019], we suggest that an approach which considers specific workflows and tasks needs to address such volatility. This mirrors the idea expressed by Ellis et al. [1991] that collaborative software must consider the social factors of teamwork, adapt to a variety of tasks, and be flexible in its design. This approach not only accommodates a variety of organisational structures but also defines flexible workflows and role definitions. The importance of such flexibility has been highlighted by previous research [Cockburn and Jones, 1995; Neuwirth et al., 1990].

Considering this volatility and the need to support different collaboration styles, such as a mutual collaboration between peers and an unequal collaboration between experts and non-experts [Kim et al., 2018], we proposed design concepts that allow users to define their workflows, roles, and responsibilities dynamically.

For instance, the *time-slot separation* concept suggests supporting mutual collaboration between video editors working concurrently on a video project. The concept enables editors to define their "territory", a specific section of the timeline they are working on, and restrict access to it, thereby creating a sense of ownership and privacy. This approach promotes territorial functioning [Larsen-Ledet and Korsgaard, 2019] as each video editor can work independently on their designated section, increasing their focus and efficiency while reducing the risk of conflicts arising from overlapping tasks. Other design con-

cepts, such as the *integrated reviewing tool* and *master styles and functional separation* [Okopnyi et al., 2022], illustrate how non-mutual collaboration in video production can be supported.

The structure of a video is often predefined during pre-production in the form of scripts, storyboards, and other documents [Bartindale et al., 2016b]. Thus, such artefacts are often needed in the video editing process [Okopnyi et al., 2020]. We acknowledged this need by suggesting extending the scope of NLE software beyond the post-production stage and the video editing process per se. Some of the aforementioned design concepts, such as the *integrated reviewing tool*, illustrate this idea. Another example is the *scaffolding* design concept [Okopnyi et al., 2022]. This concept suggests creating a mobile application to select, highlight, and annotate video footage fragments immediately after filming. Although this concept resembles the "TryFilm" system presented by Bartindale et al. [2016a], it aims not to support the filming process but to inform subsequent editing and reduce the workload of video editors. It does so by redistributing some of the work, specifically selecting "good" and "bad" takes, from video editors to directors or other participants who are present on set.

Another type of redistribution involves including previously excluded participants and allowing them to actively participate in the editing process. For instance, the *suggestive editing* design path [Okopnyi et al., 2020] would allow reviewers, such as clients, to propose small and specific changes to video professionals by performing them rather than by describing such changes in annotations or emails. This design path suggests developing a lean and tailored NLE software or a dedicated UI as a part of existing software, allowing non-professionals to perform limited actions, such as moving parts of the video along the timeline. Thus, video editors would share responsibility for the final video with those who normally do not directly participate in the editing process.

Finally, an important type of redistribution occurs when work previously performed manually by humans is carried out by software systems, in other words, adding automation [Bainbridge, 1983]. For instance, the *master styles and functional separation* design concept implies that a video editor does not need to manually apply the templates for graphics created by a graphics producer. Another example is the *integrated reviewing tool* design concept, which shifts the burden of rendering and uploading the video file from users to the video-editing software system. This redistribution of work from manual labour to automatic systems is desirable as such systems would save time otherwise spent on tedious tasks and allow greater focus on creative aspects of video work [Okopnyi et al., 2021].

4.4 Challenges in Design for Collaboration

Designing for collaborative video editing and implementing relevant design concepts can be challenging due to various factors. In [Okopnyi et al., 2022], we demonstrated a need to consider potential distractions introduced by new UI elements. Additionally, video editors work with an abstract representation of media, the timeline, which adds complexity to a virtual workspace beyond that of other shared workspace applications like collaborative writing tools. Furthermore, video work itself is a complex process involving various interconnected activities often not supported by NLE software.

Contemporary NLE software supports various post-production activities, including montage, colour grading, graphics production, and sound mastering. Thus, commercial suites like DaVinci Resolve are complex and contain many parts and applications. The user interface (UI) of such software systems reflects this complexity by providing many features and functions that are readily available to users. Introducing new features, such as elements of workspace awareness and new means of communication, adds to this complexity and may make the software less appealing to end users. Thus, "generic" design ideas, such as adding a text chat to allow collaborators to exchange messages, might fail.

It is important to reiterate that video is a complex medium, and the abstract line-based representation of a video, the timeline, usually consists of multiple tracks, each containing multiple video and audio clips alongside other elements [Casares et al., 2002]. Video editors must comprehend the connection between this representation and the video [Goldman et al., 2006]. Thus, during editing, they are required to shift back and forth between editing this representation and reviewing the result. We argue this need imposes certain limitations on designers who attempt to devise collaborative video-editing software. To address this challenge, in [Okopnyi et al., 2022] we suggested employing alternative abstract representations of a video. For example, storyboards, which divide video into scenes, are often employed to represent a video in its early production stages [Owens and Millerson, 2012]. Previous research has demonstrated that storyboards could also represent video during post-production [Mackay and Pagani, 1994; Myers et al., 2001]. The *time slot separation* concept [Okopnyi et al., 2022] suggests adding dynamically defined sections on top of a timeline, providing a way to represent a video as a sequence of scenes, similar to a storyboard. Rather than simply reflecting the technical structure of the video, as with clips and tracks, this concept would introduce an additional abstract representation that could mirror the narrative structure of the video, offering a potentially more intuitive and meaningful way for collaborators to engage with the video-editing process. However, it might also introduce more complexity to video-editing software. Therefore, it requires user testing and proper evaluation,

which may also be challenging [Grudin, 1988].

Due to the volatility of video workflows and the dynamic nature of video-production teams, it is difficult to define the boundaries of the video-editing process [Ascher and Pincus, 2007]. Thus, the design concepts and paths proposed here often extend beyond the scope of video editing and address other activities, such as reviewing. Though Frame.io and other applications support these activities, they often lack integration with NLE software. This integration, upon implementation, might dramatically impact the editing process. However, as we have shown in our research, video editors sometimes limit communication with their collaborators to ensure the quality of the feedback [Okopnyi et al., 2020] and maintain privacy [Okopnyi et al., 2022].

To address the aforementioned challenges, designers must find a balance between introducing new features that support collaboration and maintaining a software's usability and appeal to end users, which requires a comprehensive understanding of the video-editing process and user needs. One way of addressing these challenges may be identifying relevant aspects of creative practices and using them to inform design rather than focusing on specific workflows [Bartindale et al., 2012]. Another strategy is developing flexible software that can be reconfigured by a user with limited technical skills according to their needs [Klokmose et al., 2019].

4.5 Changes in Video Work

When designing for improvements in work processes, it is important to consider potential resistance from professionals who might perceive changes to their work as threatening their job security [Autor, 2015]. For example, introducing automation might be met with scepticism as some workers fear it will make their skills obsolete. However, while some work may become redundant, automation can result in a redistribution of work, allowing professionals to focus less on mundane tasks and more on creative aspects of their work [Linden, 2017]. Nevertheless, workers may need to further educate themselves to operate new instruments and program automatic systems [Bainbridge, 1983; Nygren, 2014].

Remote workflows allow professionals from different locations to collaborate, providing greater opportunities for job seekers to work with companies worldwide. This is especially important for video production, where a team's location is no longer a limiting factor in determining who can participate in a project. However, the subsequent increase in dependency on general internet infrastructure can lead to a degradation in video quality

due to network latency and bandwidth limitations.

Work may also be redistributed from professional video workers to people with other backgrounds equipped with mobile devices [Salzmann et al., 2020]. For example, individuals might be asked to film themselves answering interview questions, potentially leading to lower-quality videos. However, the increasing quality of mobile cameras and the use of AI could help mitigate some of these issues [Bertini et al., 2022; Soe, 2021; Soe et al., 2021]. Furthermore, our research has shown that the quality of a video may not be as important as the story it tells because viewers may be more interested in the narrative than the production quality [Okopnyi et al., 2021]. Nonetheless, in such a setup, work is redistributed from a professional worker to a combination of human and machine, and the issue of interaction between humans and machines emerges [Soe et al., 2021].

Another issue is the contradiction between formalisation and flexibility. Video-production automation requires the production process to be formalised to a certain degree. However, the inclusion of amateurs and pro-ams in the production process requires more flexible workflows. This tension between formalisation and flexibility must be carefully navigated to ensure efficiency and creativity. Bartindale et al. [2016b] raised a similar concern when discussing scaffolding in the collaborative creative process. On the one hand, scaffolding, such as templates and guides, benefit the creators by simplifying the production process. On the other hand, it impedes creative freedom. Nevertheless, the authors suggested that professional production companies would benefit from employing scaffolding that guides the work performed by non-professional contributors, such as amateur filming crews [Bartindale et al., 2016b].

In the context of professional video work, such as TV production, trends towards automation and work redistribution suggest that video workers will be required to possess versatile skills and familiarity with various internet technologies, general-purpose applications, and mobile devices. This shift will likely lower the barrier for new entrants, and legacy TV companies may begin to seek talent outside the industry. Additionally, the adoption of practices and tools from video-based social-media platforms, like Twitch and YouTube, could become more common in the TV industry, resulting in leaner production processes and increased interaction with live audiences. In terms of Kirk et al. [2007], "lightweight" video work, such as the spontaneous capturing of events with tools at hand (e.g., smartphones), is being incorporated into "heavyweight" intentional video work. This shift may result in a conflict between different video production practices, approaches, technologies, and workflows.

4.6 Approaches to Design for Collaboration

A major barrier to including new participants in the video-production process, specifically video editing, is the complexity of contemporary video production software, such as DaVinci Resolve and Adobe Premiere Pro. These software packages contain multiple tools and instruments, making them difficult for inexperienced users. As a result, it may be challenging to include amateurs, such as clients who order a video production, in the process. Previous research has suggested that video-production software should be more flexible and re-configurable to facilitate the inclusion of new collaborators [Klokmose et al., 2019]. Additionally, specialised systems have been designed and developed to focus on specific activities and practices in production, such as spontaneous on-set storyboarding [Bartindale et al., 2012] and video reviewing [Pavel et al., 2016].

On the one hand, a holistic video-production system that allows for a wide range of tasks is beneficial. On the other hand, modular systems with loosely connected components that can be tailored to the needs of a specific user or project may be more effective in facilitating the inclusion of less-skilled users. A discussion akin to the Tanenbaum–Torvalds debate [Tanenbaum, 1999] in computer science may be necessary to explore trade-offs between holistic, specialised, and modular systems in the context of video production. That debate revolved around trade-offs between monolithic (holistic) and microkernel (modular) operating-system (OS) design, focusing on performance, reliability, and adaptability. In the monolithic model, defended by Linus Torvalds, all components are tightly integrated, providing efficiency and a unified experience but potentially leading to complexity and inflexibility. The microkernel model, favoured by Andrew S. Tanenbaum, offers flexibility, fault isolation, and adaptability to specific needs, but it may sacrifice some performance and cohesive integration. In the context of video-editing software design, at least three design approaches can compete in a similar discussion, each reflecting different philosophies and practical considerations.

The first approach, “holistic,” involves developing comprehensive all-in-one tools that offer a wide range of features to cater to diverse user needs. This approach, evident in software like DaVinci Resolve, provides users with an extensive suite of functionalities. However, these tools often come with a steep learning curve, potentially limiting their accessibility to users with lower proficiency.

A second approach, “tailored”, is reminiscent of tools such as Frame.io, Kapwing, and StoryCrate [Bartindale et al., 2012]. This approach revolves around creating specialised tools for specific use cases. These tools may have a narrower range of functionalities, but their simplified interfaces and targeted feature sets can make them more intuitive and user-friendly. The trade-off, however, is that their applicability may be limited to

certain contexts or user groups. Simultaneously, some of these specialised tools, such as Frame.io, are built not to replace but rather to augment existing NLE software systems.

The third approach, "configurable", involves designing configurable modular software systems that can be customised according to a user's needs and preferences. This approach, demonstrated in systems like Videostrates [Klokmose et al., 2019], offers a promising means of balancing functionality and usability. By allowing users to add or remove features according to their specific requirements, these systems can potentially cater to a broader user base without overwhelming users with unnecessary functionalities. However, this approach might require users to possess technical expertise unrelated to video production, such as knowledge of HTML.

In this framework, the design concepts proposed in [Okopnyi et al., 2022] can be viewed as a trade-off between holistic and tailored approaches. The concepts originated from an attempt to utilise the holistic approach and envisioned adding collaborative features to existing professional NLE software systems. However, in the discussions, workshop participants drifted towards adopting a tailored approach expressed in the form of special modes (see *spectating mode* and *integrated reviewing tool* concepts) and dedicated applications (see *scaffolding* concept). Incidentally, the configurable approach was never explored in the workshops.

Modern NLE software increasingly represents an intricate balance between holistic and tailored approaches, striving to combine extensive functionality with a user-friendly design. Such balancing is evident in the evolution of software like Adobe Premiere Pro and its offshoot, Adobe Premiere Rush. Adobe Premiere Pro represents a holistic approach and offers a comprehensive suite of video-editing functionality that can manage complex projects and professional needs. Its extensive range of features supports diverse video-production tasks, but this complexity can be daunting for novice users or those with specific streamlined requirements.

Conversely, Adobe Premiere Rush[2] represents a tailored approach. It is designed explicitly for social-media content creation and offers a reduced set of functionalities, making it more accessible for users who may not require a full range of video-editing tools. It is also designed to be used on mobile devices, allowing video to be captured and edited on the go.

According to the classifications of Kirk et al. [2007] and Owens and Millerson [2012], Adobe Premiere Pro is typically associated with planned heavyweight video work, while Adobe Premiere Rush is oriented towards empirical lightweight video work. It is im-

[2]https://www.adobe.com/products/premiere-rush.html

portant to clarify that this association does not necessarily establish a fixed correlation between types of video work and corresponding NLE software design approaches. Previous research, as indicated by Bartindale et al. [2012, 2016a], has shown that tailored systems can be designed for and employed in planned video work. Moreover, the configurable approach, presents itself as an adaptable paradigm, transcending these predefined categories. It provides a flexible framework that can be customised to align with diverse video-work requirements ranging from structured heavyweight projects to more agile lightweight operations.

Notably, specialised or niche NLE software, such as Motionbox and Kapwing, seems to evolve more rapidly in terms of incorporating and refining collaborative features. This phenomenon may be attributed to the inherently targeted nature of these tools, which enable them to adapt more swiftly to emerging needs for remote collaboration. With simpler and more focused functionality, these platforms may find it easier to innovate and integrate collaborative features. In contrast, more comprehensive holistic systems might encounter greater difficulties when integrating new collaborative functionalities given their complexity and potential for clashes with existing features. The rapid evolution of specialised tools embracing collaborative features underscores the potential strengths of a tailored approach, especially in rapidly changing fields like remote collaboration, where agility in responding to user needs and technological shifts may be paramount.

The delineation of three design approaches — holistic, tailored, and configurable — extends beyond the specific domain of collaborative video editing and contributes to the broader CSCW field. In the following, I qualify these design approaches based on how they might address the design challenges identified in the third study.

The primary challenge of video editing lies in its demanding nature, which necessitates a focus on work while avoiding distractions. A holistic NLE system, with its vast array of tools, seeks to fulfil every possible editing need, yet this very abundance might become overwhelming and make it challenging to focus. Implementing collaborative features in a holistic system increases this complexity. However, a tailored system can help streamline the editing process and minimise distractions by addressing specific tasks. Nonetheless, this specificity might mean that collaborators will need to adopt the workflow imposed by the software.

Configurable systems are effective because of their flexibility. Users can adapt them to their workflow, ideally minimising any elements that might divert their attention. However, establishing such systems to reach this point might be challenging and, if done incorrectly, counterproductive due to possible breakdowns that would divert users' attention from video work to software maintenance.

Another challenge concerns looking beyond the immediate scope of video editing. In this case, the holistic approach has an advantage because it potentially offers tools that cater not only to editing but also to various related activities, such as reviewing. Tailored systems, given their niche focus, may not always consider these broader aspects unless explicitly designed to. Configurable systems, however, provide a middle ground. They allow users the freedom to integrate functionalities beyond the conventional NLE space, though the effectiveness of this process is largely dependent on a user's awareness and the range of options provided.

Finally, given the time-based nature of video, the representation of time emerges as a pivotal design issue. Holistic systems, due to their comprehensive nature, may provide varied means to depict the temporal dimension of video. However, as previously mentioned, the "timeline" metaphor seems to prevail. Tailored systems can offer alternative ways to represent time if that is their focus. Configurable systems are effective in this regard as well. In principle, such systems grant users the power to select or even design their own temporal representation tool. Nonetheless, this adaptability relies heavily on user expertise in delineating their needs.

The three approaches discussed here are derived from previous research, existing tool analysis, and the studies conducted for this thesis. While broad, they have analytical potential beyond the realm of collaborative video editing. By identifying these approaches, this thesis offers a conceptual framework that may guide and inform the design of collaborative systems across various domains. Whether applied to healthcare, education, or creative industries, these design philosophies encapsulate essential considerations for balancing functionality, usability, and adaptability. Identifying these approaches as potential design strategies provides additional perspectives that may inform future research and practice within CSCW.

4.7 Limitations

While this research is situated in the realm of professional video editing and TV production, its findings may not necessarily apply to other forms of media production.

Implementing remote workflows in video editing, as discussed in this study, might vary based on the specific requirements and preferences of different production teams and organisations. External factors, such as budget, time constraints, and legal or ethical considerations, may also play a pivotal role in shaping these workflows.

Solo video producers, including YouTubers and Twitch live-streamers, are not covered

in this research. This thesis intentionally does not investigate the practices of such solo video producers.

This research spanned several years, and during this period, the video production landscape has likely changed significantly. Notably, the onset of the COVID-19 pandemic pushed many organisations and individuals towards remote workflows. The long-term consequences and deeper impacts of such drastic shifts remain beyond the scope of this thesis. Additionally, technological progress might have ushered in new tools and platforms not accounted for in this thesis.

The data upon which this research is founded is qualitative in nature. As with all qualitative data, inherent biases and limitations are present. While every effort has been made to acknowledge and mitigate these issues throughout the research process, they remain an intrinsic part of the study's structure.

Chapter 5

Conclusion

This thesis has explored collaborative practices in video editing and identified opportunities for and challenges of designing tools that support collaboration. The research produced several findings and contributions that have implications for both the computer-supported cooperative work (CSCW) field and the practical domain of media-production software design.

5.1 Contribution

The iterative nature of the video-editing process and the critical role of reviewing in shaping the final output were highlighted in the study. The research identified nine relevant themes for understanding reviewing in video editing, describing how uncertainty manifests in collaborative video editing and how editors reach agreements through various strategies and social mechanisms, such as the use of documentation and iconic referencing. Understanding these practices is crucial when addressing the inherent complexities and dynamics of video-editing collaborations and can serve as a foundation for developing collaborative video-editing tools.

The research also reported on the implications of adopting remote practices in TV production organisations. These implications include the increasing role of general public Internet technologies and mobile devices in professional video production and the associated reskilling of video workers. The research identified two important concepts that qualify these implications: automation and redistribution.

Several design ideas that encompass both broad design paths and specific concepts were proposed. Design paths describe general directions, which we suggest to designers who

intend to create new features and tools that support collaboration in video editing. For example, we recommend exploring possible ways to support iconic referencing in video-editing software. Design concepts present specific ideas, some of which may be implemented in existing non-linear editing (NLE) software, such as time-slot separation, and others that involve the development of new applications.

These ideas demonstrate technological advancements' potential to facilitate and enhance collaboration. However, they also underscore the necessity of a careful nuanced approach to design that considers the complex nature of video-editing tasks and the diversity of user needs.

Additionally, this research uncovered challenges in designing collaborative video-production tools. For example, implementing collaborative features in existing NLE tools may introduce unwanted distractions, impeding the editing process rather than facilitating it. The complexity of these tools, combined with the need to accommodate various use cases and user preferences, renders design a complex task.

Building on the insights gathered throughout this research, I have identified three potential approaches for designing collaborative video-editing software. These approaches represent a synthesis of strategies observed in previous research, existing software, and the results of this thesis.

The first approach is holistic and aims to create a complex system that supports a wide range of tasks and incorporates collaboration features. This approach implies a steep learning curve for users, as is the case with existing systems like Adobe Premiere Pro. The second approach involves tailored applications designed for specific practices [Bartindale et al., 2012]. However, the challenge of this approach is the difficulty in developing an application for every unique use case. The third approach focuses on flexible systems that can be reconfigured by their users to suit individual needs, but it requires users to possess some technical expertise, such as basic web-development skills [Klokmose et al., 2019].

These three approaches provide potential pathways for the design of collaborative video-editing tools. However, selecting the appropriate approach depends heavily on the specific context of use, the user's technical proficiency, and the nature of the collaborative tasks involved.

The rapidly evolving nature of video production demands a continuous examination of how technology can facilitate collaboration in this field. While previous research has made significant strides in understanding and supporting remote collaboration in various domains, the area of video editing remains relatively uncharted. Unlike online

collaborative tools for writing, such as Google Docs, there is no "gold standard" tool for online collaborative video editing. This gap signals a promising direction for future research and technological innovation.

As media production continues to evolve, so too should the tools that support collaboration in this domain. As such, the design ideas and challenges identified in this thesis provide a roadmap for future development in this area. However, these ideas and challenges also highlight the need for a discussion about how to facilitate collaboration in complex tools like video-editing software. Given the intricacies of these tools and the risk of feature overload [Zhang and Padmanabhan, 2011], the question of how to enable collaboration without overwhelming users remains unanswered.

In conclusion, this thesis contributes to the understanding of collaboration in video production and offers a starting point for designing tools to facilitate such collaboration.

5.2 Implications for future research

The findings of this thesis have identified several directions for future research which have implications for both video production and collaborative software design.

First, a deeper exploration of emerging collaborative practices in video production is necessary. As technology evolves and the video-production landscape continues to change, new forms of collaboration are likely to arise. Future studies could focus on identifying and understanding these practices, particularly in remote and asynchronous collaboration. For instance, investigating the role of communication tools such as video conferencing and instant messaging in the video-production process may provide valuable insights.

Second, the design ideas presented in this thesis could be further developed and tested. The general design paths and specific design concepts suggested offer promising starting points, but they must be refined and evaluated through user testing and iterative design processes. A potential direction for future research might be the exploration of alternative representations of time within user interfaces to enhance both individual user experience and collaborative processes.

Third, the organisational challenges that arise when including less-skilled participants in the video-production process must be addressed. Redistributing work and responsibilities to accommodate diverse skill levels is a significant challenge. This shift also requires investigating the potential of automated tools, including generative AI tools,

in aiding users and streamlining collaborative video-editing processes while maintaining high-quality output. However, it is equally important to explore strategies for seamlessly integrating these innovative tools and workflows into existing production processes. For example, new workers with experience in video production for social-media sites might influence production workflows in organisations that traditionally employ "heavyweight" video work [Kirk et al., 2007].

Fourth, the challenge of designing to support collaboration in video editing requires further attention. The complexity of video-production software is a significant barrier to collaboration, especially for less-skilled users. Future research should investigate ways of making these tools more accessible without sacrificing efficiency.

Finally, the broader question of how to enable collaboration for complex professional software, such as computer-aided design (CAD) systems and music-production software, is ready for further exploration. As technology continues to evolve, the demand for online collaborative tools will likely increase. However, the primary challenge lies in balancing the need for complex functionalities with the desire for user-friendly interfaces. Future research could explore innovative ways to attain this balance. Additionally, examining and qualifying the identified design approaches may offer valuable insights into the design of collaborative software.

The research directions proposed above offer exciting opportunities for future studies to build on the findings of this thesis and contribute to the ongoing development of collaborative video-production tools.

Bibliography

C. Abras, D. Maloney-Krichmar, and J. Preece. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37 (4):445–456, 2004.

B. Adams and S. Venkatesh. Situated Event Bootstrapping and Capture Guidance for Automated Home Movie Authoring. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, MULTIMEDIA '05, pages 754–763, New York, NY, USA, 2005. ACM. ISBN 978-1-59593-044-6. doi: 10.1145/1101149.1101312. URL http://doi.acm.org/10.1145/1101149.1101312.

P. Albiniak. How the Pandemic Has Permanently Changed TV, Jan. 2022. URL https://www.nexttv.com/news/how-the-pandemic-has-permanently-changed-tv.

M. Arntz, T. Gregory, and U. Zierahn. The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. May 2016. doi: https://doi.org/10.1787/5jlz9h56dvq7-en. URL https://www.oecd-ilibrary.org/social-issues-migration-health/the-risk-of-automation-for-jobs-in-oecd-countries_5jlz9h56dvq7-en. Publisher: OECD.

M. Arshad. COVID-19: It's time to be Thankful to our ICT Professionals. *Information Technology*, 9(2):9, 2020.

S. Ascher and E. Pincus. *The filmmaker's handbook: A comprehensive guide for the digital age*. Penguin, 3rd edition, 2007.

D. H. Autor. Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3):3–30, Sept. 2015. ISSN 0895-3309. doi: 10.1257/jep.29.3.3. URL https://www.aeaweb.org/articles?id=10.1257%2Fjep.29.3.3&source=post_page----------.

L. Bainbridge. Ironies of automation. *Automatica*, 19(6):775–779, Nov. 1983. ISSN 0005-1098. doi: 10.1016/0005-1098(83)90046-8. URL http://www.sciencedirect.com/science/article/pii/0005109883900468.

E. Baker, J. Geirland, T. Fisher, and A. Chandler. Media Production: Towards Creative Collaboration Using Communication Networks. *Computer Supported Cooperative Work (CSCW)*, 8(4):303–332, Dec. 1999. ISSN 1573-7551. doi: 10.1023/A:1008616002814. URL https://doi.org/10.1023/A:1008616002814.

M. W. S. Barlow. Computer Television Broadcast Automation. *Journal of the SMPTE*, 81(3):172–172, Mar. 1972. ISSN 0361-4573. doi: 10.5594/J05628. Conference Name: Journal of the SMPTE.

M. W. S. Barlow. The Automation of Large Program Routing Switchers in the CBC. *SMPTE Journal*, 88(7):493–497, July 1979. ISSN 0036-1682. doi: 10.5594/J06643. Conference Name: SMPTE Journal.

T. Bartindale, A. Sheikh, N. Taylor, P. Wright, and P. Olivier. StoryCrate: Tabletop Storyboarding for Live Film Production. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 169–178, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi: 10.1145/2207676.2207700. URL http://doi.acm.org/10.1145/2207676.2207700. event-place: Austin, Texas, USA.

T. Bartindale, G. Schofield, C. Crivellaro, and P. Wright. TryFilm: Situated Support for Interactive Media Productions. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1412–1422, New York, NY, USA, 2016a. ACM. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819929. URL http://doi.acm.org/10.1145/2818048.2819929. event-place: San Francisco, California, USA.

T. Bartindale, G. Schofield, and P. Wright. Scaffolding Community Documentary Film Making using Commissioning Templates. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2705–2716, New York, NY, USA, May 2016b. Association for Computing Machinery. ISBN 978-1-4503-3362-7. doi: 10.1145/2858036.2858102. URL https://doi.org/10.1145/2858036.2858102.

A. Bartolome and S. Niu. A Literature Review of Video-Sharing Platform Research in HCI. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–20, New York, NY, USA, Apr. 2023. Association for Computing Machinery. ISBN 978-1-4503-9421-5. doi: 10.1145/3544548.3581107. URL https://dl.acm.org/doi/10.1145/3544548.3581107.

M. Bertini, L. Galteri, L. Seidenari, T. Uricchio, and A. D. Bimbo. Fast and effective AI approaches for video quality improvement. In *Proceedings of the 1st Mile-High Video Conference*, MHV '22, pages 77–78, New York, NY, USA, Mar. 2022. Association for Computing Machinery. ISBN 978-1-4503-9222-8. doi: 10.1145/3510450.3517270. URL https://dl.acm.org/doi/10.1145/3510450.3517270.

C. A. Bolstad and M. R. Endsley. Tools for supporting team collaboration. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 47, pages 374–378. SAGE Publications Sage CA: Los Angeles, CA, 2003. Issue: 3.

V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, Jan. 2006. ISSN 1478-0887. doi: 10.1191/1478088706qp063oa. URL `https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa`. Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

D. Buckingham, M. Pini, and R. Willett. 'Take back the tube!': The discursive construction of amateur film and video making. *Journal of Media Practice*, 8(2):183–201, Jan. 2007. ISSN 1468-2753. doi: 10.1386/jmpr.8.2.183_1. URL `https://doi.org/10.1386/jmpr.8.2.183_1`. Publisher: Routledge _eprint: https://doi.org/10.1386/jmpr.8.2.183_1.

C. Buysschaert, A. Descampe, and J.-B. Lorent. Creating Bandwidth-Efficient Workflows with JPEG XS and SMPTE ST 2110. *SMPTE Motion Imaging Journal*, 129(7):33–40, Aug. 2020. ISSN 2160-2492. doi: 10.5594/JMI.2020.3001465. Conference Name: SMPTE Motion Imaging Journal.

S. Bødker and C. N. Klokmose. Dynamics in artifact ecologies. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, pages 448–457, 2012.

D. Cabral and N. Correia. Videoink: A Pen-based Approach for Video Editing. In *Adjunct Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST Adjunct Proceedings '12, pages 67–68, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1582-1. doi: 10.1145/2380296.2380325. URL `http://doi.acm.org/10.1145/2380296.2380325`.

J. Casares, A. C. Long, B. A. Myers, R. Bhatnagar, S. M. Stevens, L. Dabbish, D. Yocum, and A. Corbett. Simplifying Video Editing Using Metadata. In *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '02, pages 157–166, New York, NY, USA, 2002. ACM. ISBN 978-1-58113-515-2. doi: 10.1145/778712.778737. URL `http://doi.acm.org/10.1145/778712.778737`.

A. Casero-Ripolles. Impact of COVID-19 on the Media System. Communicative and Democratic Consequences of News Consumption during the Outbreak. SSRN Scholarly Paper ID 3594133, Social Science Research Network, Rochester, NY, Apr. 2020. URL `https://papers.ssrn.com/abstract=3594133`.

M. A. Chatti, M. Marinov, O. Sabov, R. Laksono, Z. Sofyan, A. M. Fahmy Yousef, and U. Schroeder. Video annotation and analytics in CourseMapper. *Smart Learning Environments*, 3(1):10, July 2016. ISSN 2196-7091. doi: 10.1186/s40561-016-0035-1. URL https://doi.org/10.1186/s40561-016-0035-1.

P.-Y. Chi, J. Liu, J. Linder, M. Dontcheva, W. Li, and B. Hartmann. DemoCut: generating concise instructional videos for physical demonstrations. pages 141–150. ACM, Oct. 2013. ISBN 978-1-4503-2268-3. doi: 10.1145/2501988.2502052. URL http://dl.acm.org/citation.cfm?id=2501988.2502052.

S. H. Chua, T.-J. K. P. Monserrat, D. Yoon, J. Kim, and S. Zhao. Korero: Facilitating Complex Referencing of Visual Materials in Asynchronous Discussion Interface. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW):34:1–34:19, Dec. 2017. ISSN 2573-0142. doi: 10.1145/3134669. URL http://doi.acm.org/10.1145/3134669.

R. Coche and B. J. Lynn. Behind the Scenes: COVID-19 Consequences on Broadcast Sports Production. *International Journal of Sport Communication*, 13(3):484–493, Aug. 2020. ISSN 1936-3915, 1936-3907. doi: 10.1123/ijsc.2020-0231. URL https://journals.humankinetics.com/view/journals/ijsc/13/3/article-p484.xml. Publisher: Human Kinetics Section: International Journal of Sport Communication.

A. Cockburn and S. Jones. Four principles of groupware design. *Interacting with Computers*, 7(2):195–210, 1995.

F. Collin. Harry Collins and the Empirical Programme of Relativism. In F. Collin, editor, *Science Studies as Naturalized Philosophy*, Synthese Library, pages 83–108. Springer Netherlands, Dordrecht, 2011. ISBN 978-90-481-9741-5. doi: 10.1007/978-90-481-9741-5_5. URL https://doi.org/10.1007/978-90-481-9741-5_5.

H. Collins. Stages in the Empirical Programme of Relativism. *Social Studies of Science*, 11(1):3–10, Feb. 1981. ISSN 0306-3127. doi: 10.1177/030631278101100101. URL https://doi.org/10.1177/030631278101100101.

K. C. Cousins, D. Robey, and I. Zigurs. Managing strategic contradictions in hybrid teams. *European Journal of Information Systems*, 16(4):460–478, Aug. 2007. ISSN 0960-085X. doi: 10.1057/palgrave.ejis.3000692. URL https://orsociety.tandfonline.com/doi/full/10.1057/palgrave.ejis.3000692. Publisher: Taylor & Francis.

D. De Roure, G. Klyne, J. Pybus, D. M. Weigl, and K. Page. Music SOFA: An architecture for semantically informed recomposition of Digital Music Objects. In *Proceedings of the 1st International Workshop on Semantic Applications for Audio and*

Music, SAAM '18, pages 33–41, New York, NY, USA, Oct. 2018. Association for Computing Machinery. ISBN 978-1-4503-6495-9. doi: 10.1145/3243907.3243912. URL https://doi.org/10.1145/3243907.3243912.

K. E. M. de Stobbeleir, S. J. Ashford, and D. Buyens. Self-Regulation of Creativity at Work: The Role of Feedback-Seeking Behavior in Creative Performance. *Academy of Management Journal*, 54(4):811–831, Aug. 2011. ISSN 0001-4273. doi: 10.5465/amj.2011.64870144. URL https://journals.aom.org/doi/abs/10.5465/AMJ.2011.64870144.

T. Deacon, N. Bryan-Kinns, P. G. Healey, and M. Barthet. Shaping Sounds: The Role of Gesture in Collaborative Spatial Music Composition. In *Proceedings of the 2019 on Creativity and Cognition*, C&C '19, pages 121–132, New York, NY, USA, June 2019. Association for Computing Machinery. ISBN 978-1-4503-5917-7. doi: 10.1145/3325480.3325493. URL https://doi.org/10.1145/3325480.3325493.

T. Deacon, P. Healey, and M. Barthet. "It's cleaner, definitely": Collaborative Process in Audio Production. *Computer Supported Cooperative Work (CSCW)*, Nov. 2022. ISSN 1573-7551. doi: 10.1007/s10606-022-09448-1. URL https://doi.org/10.1007/s10606-022-09448-1.

D. L. Diefenbach. *Video production techniques: Theory and practice from concept to screen*. Routledge, 2009.

N. F. Doherty, C. R. Coombs, and J. Loan-Clarke. A re-conceptualization of the interpretive flexibility of information technologies: redressing the balance between the social and the technical. *European Journal of Information Systems*, 15(6): 569–582, Dec. 2006. ISSN 0960-085X. doi: 10.1057/palgrave.ejis.3000653. URL https://doi.org/10.1057/palgrave.ejis.3000653. Publisher: Taylor & Francis _eprint: https://doi.org/10.1057/palgrave.ejis.3000653.

P. Dourish and V. Bellotti. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 107–114. ACM, 1992.

C. A. Ellis, S. J. Gibbs, and G. Rein. Groupware: some issues and experiences. *Communications of the ACM*, 34(1):39–58, Jan. 1991. ISSN 0001-0782. doi: 10.1145/99977.99987. URL https://doi.org/10.1145/99977.99987.

M. R. Endsley. Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1):32–64, Mar. 1995. ISSN 0018-7208. doi: 10.1518/001872095779049543. URL https://doi.org/10.1518/001872095779049543. Publisher: SAGE Publications Inc.

M. A. Feinberg, K.-B. Song, and I.-T. Lim. KineMaster: Pro Video Editing on Android. In *ACM SIGGRAPH 2016 Appy Hour*, SIGGRAPH '16, pages 7:1–7:1, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4376-3. doi: 10.1145/2936744.2956677. URL http://doi.acm.org/10.1145/2936744.2956677.

K. J. K. Feng, T. W. Li, and A. X. Zhang. Understanding Collaborative Practices and Tools of Professional UX Practitioners in Software Organizations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–20, New York, NY, USA, Apr. 2023. Association for Computing Machinery. ISBN 978-1-4503-9421-5. doi: 10.1145/3544548.3581273. URL https://dl.acm.org/doi/10.1145/3544548.3581273.

B. Fonseca and E. Carrapatoso. Coview: A Cooperative Architecture for Digital Video Editing. *SMPTE Motion Imaging Journal*, 115(11-12):482–491, Nov. 2006. ISSN 2160-2492. doi: 10.5594/J16132. Conference Name: SMPTE Motion Imaging Journal.

J. B. Fonseca and E. M. Carrapatoso. Collaborative editing of digital video. In *IEE European Workshop Distributed Imaging (Ref. No. 1999/109)*, pages 20–1. IET, 1999.

A. Freilich and S. Meyer. The "STEP" System A Unique, Low-Cost TV Automation System. *IEEE Transactions on Broadcasting*, BC-9(1):16–25, Feb. 1963. ISSN 1557-9611. doi: 10.1109/TBC.1963.265906. Conference Name: IEEE Transactions on Broadcasting.

P. Fröhlich, M. Baldauf, T. Meneweger, M. Tscheligi, B. de Ruyter, and F. Paternó. Everyday automation experience: a research agenda. *Personal and Ubiquitous Computing*, 24(6):725–734, Dec. 2020. ISSN 1617-4917. doi: 10.1007/s00779-020-01450-y. URL https://doi.org/10.1007/s00779-020-01450-y.

H. Garfinkel. Studies in ethnomethodology. *Englewood Cliffs, N*, 1967.

E. M. Gerson and S. L. Star. Analyzing Due Process in the Workplace. *ACM Trans. Inf. Syst.*, 4(3):257–270, July 1986. ISSN 1046-8188. doi: 10.1145/214427.214431. URL http://doi.acm.org/10.1145/214427.214431.

A. Girgensohn, J. Boreczky, P. Chiu, J. Doherty, J. Foote, G. Golovchinsky, S. Uchihashi, and L. Wilcox. A Semi-automatic Approach to Home Video Editing. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology*, UIST '00, pages 81–89, New York, NY, USA, 2000. ACM. ISBN 978-1-58113-212-0. doi: 10.1145/354401.354415. URL http://doi.acm.org/10.1145/354401.354415.

D. B. Goldman, B. Curless, D. Salesin, and S. M. Seitz. Schematic Storyboarding for Video Visualization and Editing. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH

'06, pages 862–871, New York, NY, USA, 2006. ACM. ISBN 978-1-59593-364-5. doi: 10.1145/1179352.1141967. URL http://doi.acm.org/10.1145/1179352.1141967.

J. Grudin. Why CSCW applications fail: problems in the design and evaluationof organizational interfaces. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*, pages 85–93, 1988.

J. Grudin. Groupware and social dynamics: eight challenges for developers. *Communications of the ACM*, 37(1):92–105, Jan. 1994. ISSN 0001-0782, 1557-7317. doi: 10.1145/175222.175230. URL https://dl.acm.org/doi/10.1145/175222.175230.

J. Guo, C. Gurrin, and S. Lao. Who produced this video, amateur or professional? In *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*, ICMR '13, pages 271–278, New York, NY, USA, Apr. 2013. Association for Computing Machinery. ISBN 978-1-4503-2033-7. doi: 10.1145/2461466.2461509. URL https://doi.org/10.1145/2461466.2461509.

F. Guribye and L. Nyre. The Changing Ecology of Tools for Live News Reporting. *Journalism Practice*, 11(10):1216–1230, Nov. 2017. ISSN 1751-2786. doi: 10.1080/17512786.2016.1259011. URL https://doi.org/10.1080/17512786.2016.1259011.

C. Gutwin and S. Greenberg. The effects of workspace awareness support on the usability of real-time distributed groupware. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 6(3):243–281, 1999. Publisher: ACM New York, NY, USA.

C. Gutwin and S. Greenberg. The mechanics of collaboration: developing low cost usability evaluation methods for shared workspaces. In *Proceedings IEEE 9th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2000)*, pages 98–103, June 2000. doi: 10.1109/ENABL.2000.883711. ISSN: 1080-1383.

C. Gutwin and S. Greenberg. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):411–446, Sept. 2002. ISSN 0925-9724, 1573-7551. doi: 10.1023/A:1021271517844. URL https://link.springer.com/article/10.1023/A:1021271517844.

M. Hammer and M. Sirbu. What is office automation? (1980). In *Perspectives on the computer revolution*, pages 351–368. Ablex Publishing Corp., USA, June 1989. ISBN 978-0-89391-369-4.

K. Heinonen and T. Strandvik. Reframing service innovation: COVID-19 as a catalyst for imposed service innovation. *Journal of Service Management*, 32(1):101–112, Jan. 2020. ISSN 1757-5818. doi: 10.1108/JOSM-05-2020-0161. URL https://doi.org/10.1108/JOSM-05-2020-0161. Publisher: Emerald Publishing Limited.

M. Heinrich, F. J. Grüneberger, T. Springer, P. Hauer, and M. Gaedke. GAwI: A Comprehensive Workspace Awareness Library for Collaborative Web Applications. In F. Daniel, P. Dolog, and Q. Li, editors, *Web Engineering*, Lecture Notes in Computer Science, pages 482–485, Berlin, Heidelberg, 2013. Springer. ISBN 978-3-642-39200-9. doi: 10.1007/978-3-642-39200-9_40.

C. Herrmann, R. S. Bowen, N. Wadhwa, R. Garg, Q. He, J. T. Barron, and R. Zabih. Learning to Autofocus. pages 2230–2239, 2020. URL `https://openaccess.thecvf.com/content_CVPR_2020/html/Herrmann_Learning_to_Autofocus_CVPR_2020_paper.html`.

J. Hollan and S. Stornetta. Beyond being there. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 119–125, 1992.

X.-S. HUA, L. LU, and H.-J. ZHANG. AVE: Automated Home Video Editing. In *Proceedings of the Eleventh ACM International Conference on Multimedia*, MULTIMEDIA '03, pages 490–497, New York, NY, USA, 2003. ACM. ISBN 978-1-58113-722-4. doi: 10.1145/957013.957121. URL `http://doi.acm.org/10.1145/957013.957121`.

W. Huang, S. Kim, M. Billinghurst, and L. Alem. Sharing hand gesture and sketch cues in remote collaboration. *Journal of Visual Communication and Image Representation*, 58:428–438, Jan. 2019. ISSN 1047-3203. doi: 10.1016/j.jvcir.2018.12.010. URL `https://www.sciencedirect.com/science/article/pii/S1047320318303365`.

P. Johnson-Lenz and T. Johnson-Lenz. Groupware: coining and defining it. *ACM SIGGROUP Bulletin*, 19(2):34, Aug. 1998. ISSN 2372-7403, 2372-739X. doi: 10.1145/290575.290585. URL `https://dl.acm.org/doi/10.1145/290575.290585`.

O. Juhlin, A. Engström, and E. Önnevall. Long Tail TV Revisited: From Ordinary Camera Phone Use to Pro-am Video Production. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 1325–1334, New York, NY, USA, 2014a. ACM. ISBN 978-1-4503-2473-1. doi: 10.1145/2556288.2557315. URL `http://doi.acm.org/10.1145/2556288.2557315`.

O. Juhlin, G. Zoric, A. Engström, and E. Reponen. Video interaction: a research agenda. *Personal and Ubiquitous Computing*, 18(3):685–692, Mar. 2014b. ISSN 16174909. doi: http://dx.doi.org/10.1007/s00779-013-0705-8. URL `https://www.proquest.com/docview/1498211149/abstract/11D25A2D6084E7CPQ/1`. Num Pages: 685-692 Place: London, Netherlands Publisher: Springer Nature B.V.

D. Karis, D. Wildman, and A. Mané. Improving Remote Collaboration With Video Conferencing and Video Portals. *Human–Computer Interaction*, 31(1):1–58, Jan. 2016. ISSN 0737-0024. doi: 10.1080/07370024.2014.921506. URL `https://`

doi.org/10.1080/07370024.2014.921506. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370024.2014.921506.

S. Kim, M. Billinghurst, and G. Lee. The Effect of Collaboration Styles and View Independence on Video-Mediated Remote Collaboration. *Computer Supported Cooperative Work (CSCW)*, 27(3):569–607, Dec. 2018. ISSN 1573-7551. doi: 10.1007/s10606-018-9324-2. URL https://doi.org/10.1007/s10606-018-9324-2.

D. Kirk, A. Sellen, R. Harper, and K. Wood. Understanding videowork. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 61–70. ACM, 2007.

C. N. Klokmose, J. R. Eagan, S. Baader, W. Mackay, and M. Beaudouin-Lafon. Webstrates: Shareable Dynamic Media. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pages 280–290, New York, NY, USA, Nov. 2015. Association for Computing Machinery. ISBN 978-1-4503-3779-3. doi: 10.1145/2807442.2807446. URL https://doi.org/10.1145/2807442.2807446.

C. N. Klokmose, C. Remy, J. B. Kristensen, R. Bagge, M. Beaudouin-Lafon, and W. Mackay. Videostrates: Collaborative, Distributed and Programmable Video Manipulation. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, pages 233–247, New York, NY, USA, Oct. 2019. Association for Computing Machinery. ISBN 978-1-4503-6816-2. doi: 10.1145/3332165.3347912. URL https://doi.org/10.1145/3332165.3347912.

B. Koehne, P. C. Shih, and J. S. Olson. Remote and alone: coping with being the remote member on the team. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, pages 1257–1266, New York, NY, USA, Feb. 2012. Association for Computing Machinery. ISBN 978-1-4503-1086-4. doi: 10.1145/2145204.2145393. URL https://doi.org/10.1145/2145204.2145393.

K. Krippendorff. *Content analysis: An introduction to its methodology*. Sage publications, 2018.

A. Kumar and M. S. M. Haneef. Is Mojo (En)De-Skilling? *Journalism Practice*, 12(10): 1292–1310, Nov. 2018. ISSN 1751-2786. doi: 10.1080/17512786.2017.1389291. URL https://doi.org/10.1080/17512786.2017.1389291. Publisher: Routledge _eprint: https://doi.org/10.1080/17512786.2017.1389291.

I. Larsen-Ledet and H. Korsgaard. Territorial Functioning in Collaborative Writing. *Computer Supported Cooperative Work (CSCW)*, 28(3):391–433, June 2019. ISSN 1573-7551. doi: 10.1007/s10606-019-09359-8. URL https://doi.org/10.1007/s10606-019-09359-8.

C. Leadbeater and P. Miller. *The pro-am revolution: How enthusiasts are changing our society and economy*. Demos London, 2004.

G. A. Lee, T. Teo, S. Kim, and M. Billinghurst. Mixed reality collaboration through sharing a live panorama. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*, SA '17, pages 1–4, New York, NY, USA, Nov. 2017. Association for Computing Machinery. ISBN 978-1-4503-5410-3. doi: 10.1145/3132787.3139203. URL https://doi.org/10.1145/3132787.3139203.

G. A. Lee, T. Teo, S. Kim, and M. Billinghurst. A User Study on MR Remote Collaboration Using Live 360 Video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 153–164, Oct. 2018. doi: 10.1109/ISMAR.2018.00051. ISSN: 1554-7868.

S. W. Lee. Hybrid Use of Asynchronous and Synchronous Interaction for Collaborative Creation. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pages 95–98, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5419-6. doi: 10.1145/3131785.3131841. URL http://doi.acm.org/10.1145/3131785.3131841.

C.-G. Linden. Decades of Automation in the Newsroom. *Digital Journalism*, 5(2): 123–140, Feb. 2017. ISSN 2167-0811. doi: 10.1080/21670811.2016.1160791. URL https://doi.org/10.1080/21670811.2016.1160791. Publisher: Routledge _eprint: https://doi.org/10.1080/21670811.2016.1160791.

R. b. C.-G. Lindén. Automating the News: How Algorithms are Rewriting the Media. *Mass Communication and Society*, 23(6):968–970, Nov. 2020. ISSN 1520-5436. doi: 10.1080/15205436.2020.1783887. URL https://doi.org/10.1080/15205436.2020.1783887. Publisher: Routledge _eprint: https://doi.org/10.1080/15205436.2020.1783887.

K. Luther, J.-L. Tolentino, W. Wu, A. Pavel, B. P. Bailey, M. Agrawala, B. Hartmann, and S. P. Dow. Structuring, Aggregating, and Evaluating Crowdsourced Design Critique. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, pages 473–485, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-2922-4. doi: 10.1145/2675133.2675283. URL http://doi.acm.org/10.1145/2675133.2675283. event-place: Vancouver, BC, Canada.

W. Mackay and D. Pagani. Video Mosaic: Laying out time in a physical space. In *Proceedings of the second ACM international conference on Multimedia*, pages 165–172, 1994.

M. N. Marshall. Sampling for qualitative research. *Family practice*, 13(6):522–526, 1996. Publisher: Oxford University Press.

S. McGrath, A. Chamberlain, and S. Benford. Making Music Together: An Exploration of Amateur and Pro-Am Grime Music Production. pages 186–193. ACM Press, 2016. ISBN 978-1-4503-4822-5. doi: 10.1145/2986416.2986432. URL `http://dl.acm.org/citation.cfm?doid=2986416.2986432`.

N. Mercer. Neo-Vygotskian theory and classroom education. *Language, literacy and learning in educational practice*, pages 92–110, 1994. Publisher: Multilingual Matters Clevedon, England.

A. Merz, A. Hu, and T. Lin. ClipWorks: A Tangible Interface for Collaborative Video Editing. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*, IDC '18, pages 497–500, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5152-2. doi: 10.1145/3202185.3210758. URL `http://doi.acm.org/10.1145/3202185.3210758`.

U. Meyer and I. Schulz-Schaeffer. Three forms of interpretative flexibility. *Science, Technology & Innovation Studies*, 2, 2006.

W. Murch. *In the Blink of an Eye*, volume 995. Silman-James Press Los Angeles, 2001.

B. A. Myers, J. P. Casares, S. Stevens, L. Dabbish, D. Yocum, and A. Corbett. A Multi-view Intelligent Editor for Digital Video Libraries. In *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '01, pages 106–115, New York, NY, USA, 2001. ACM. ISBN 978-1-58113-345-5. doi: 10.1145/379437.379461. URL `http://doi.acm.org/10.1145/379437.379461`.

T. Neumayr, H.-C. Jetter, M. Augstein, J. Friedl, and T. Luger. Domino: A Descriptive Framework for Hybrid Collaboration and Coupling Styles in Partially Distributed Teams. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):128:1–128:24, Nov. 2018. doi: 10.1145/3274397. URL `https://doi.org/10.1145/3274397`.

C. M. Neuwirth, D. S. Kaufer, R. Chandhok, and J. H. Morris. Issues in the design of computer support for co-authoring and commenting. In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*, pages 183–195, 1990.

C. Nguyen, S. DiVerdi, A. Hertzmann, and F. Liu. Vremiere: In-Headset Virtual Reality Video Editing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5428–5438, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4655-9. doi: 10.1145/3025453.3025675. URL `http://doi.acm.org/10.1145/3025453.3025675`.

D. Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

B. Novikov and O. Proskurnin. Towards Collaborative Video Authoring. In L. Kalinichenko, R. Manthey, B. Thalheim, and U. Wloka, editors, *Advances in Databases and Information Systems*, Lecture Notes in Computer Science, pages 370–384. Springer Berlin Heidelberg, 2003. ISBN 978-3-540-39403-7.

G. Nygren. Multiskilling in the newsroom–de-skilling or re-skilling of journalistic work? *The Journal of Media Innovations*, 1(2):75–96, 2014.

P. Okopnyi, O. Juhlin, and F. Guribye. Unpacking Editorial Agreements in Collaborative Video Production. In *ACM International Conference on Interactive Media Experiences*, IMX '20, pages 117–126, New York, NY, USA, June 2020. Association for Computing Machinery. ISBN 978-1-4503-7976-2. doi: 10.1145/3391614.3393652. URL https://doi.org/10.1145/3391614.3393652.

P. Okopnyi, F. Guribye, V. Caruso, and O. Juhlin. Automation and redistribution of work: the impact of social distancing on live TV production. *Human–Computer Interaction*, 0(0):1–23, Nov. 2021. ISSN 0737-0024. doi: 10.1080/07370024.2021.1984917. URL https://doi.org/10.1080/07370024.2021.1984917. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370024.2021.1984917.

P. Okopnyi, O. Juhlin, and F. Guribye. Designing for Collaborative Video Editing. In *Nordic Human-Computer Interaction Conference*, NordiCHI '22, pages 1–11, New York, NY, USA, Oct. 2022. Association for Computing Machinery. ISBN 978-1-4503-9699-8. doi: 10.1145/3546155.3546664. URL https://doi.org/10.1145/3546155.3546664.

J. O'Leary, H. Winnemöller, W. Li, M. Dontcheva, and M. Dixon. Charrette: Supporting In-Person Discussions around Iterations in User Interface Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–11, Montreal QC, Canada, Apr. 2018. Association for Computing Machinery. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3174109. URL https://doi.org/10.1145/3173574.3174109.

J. Owens and G. Millerson. *Video production handbook*. Routledge, 2012.

M. Q. Patton. *Qualitative research & evaluation methods*. sage, 2002.

A. Pavel, D. B. Goldman, B. Hartmann, and M. Agrawala. VidCrit: Video-based Asynchronous Video Review. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 517–528, New York, NY,

USA, 2016. ACM. ISBN 978-1-4503-4189-9. doi: 10.1145/2984511.2984552. URL http://doi.acm.org/10.1145/2984511.2984552.

C. S. Peirce and J. Buchler. Logic as semiotic: The theory of signs, 1902.

M. Perry, O. Juhlin, M. Esbjörnsson, and A. Engström. Lean collaboration through video gestures: co-ordinating the production of live televised sport. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2279–2288. ACM, 2009.

J. Phalip, M. Morphett, and E. Edmonds. Alleviating communication challenges in film scoring: an interaction design approach. In *Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces*, pages 9–16. ACM, 2007.

J. Phalip, E. A. Edmonds, and D. Jean. Supporting Remote Creative Collaboration in Film Scoring. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, C&C '09, pages 211–220, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-865-0. doi: 10.1145/1640233.1640266. URL http://doi.acm.org/10.1145/1640233.1640266. event-place: Berkeley, California, USA.

A. Phillips, J. B. Singer, T. Vlad, and L. B. Becker. Implications of Technological Change For Journalists' Tasks and Skills. *Journal of Media Business Studies*, 6(1):61–85, Mar. 2009. ISSN 1652-2354. doi: 10.1080/16522354.2009.11073479. URL https://doi.org/10.1080/16522354.2009.11073479. Publisher: Routledge _eprint: https://doi.org/10.1080/16522354.2009.11073479.

T. J. Pinch and W. E. Bijker. The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science*, 14(3):399–441, Aug. 1984. ISSN 0306-3127. doi: 10.1177/030631284014003004. URL https://doi.org/10.1177/030631284014003004.

O. Proskurnin. Concurent video: Operational extensions. In *J. Barzdins and A. Caplinskas (ed.), Databases and Information Systems, Selected Papers from the Sixth International Baltic Conference DB&IS 2004*, pages 78–87, 2005.

K. R. Rao, D. N. Kim, and J. J. Hwang. Video Coding Standards and Video Formats. In K. Rao, D. N. Kim, and J. J. Hwang, editors, *Video coding standards: AVS China, H.264/MPEG-4 PART 10, HEVC, VP6, DIRAC and VC-1*, Signals and Communication Technology, pages 37–50. Springer Netherlands, Dordrecht, 2014. ISBN 978-94-007-6742-3. doi: 10.1007/978-94-007-6742-3_2. URL https://doi.org/10.1007/978-94-007-6742-3_2.

D. Retelny and P. Hinds. Embedding Intentions in Drawings: How Architects Craft and Curate Drawings to Achieve Their Goals. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1310–1322, San Francisco, California, USA, Feb. 2016. Association for Computing Machinery. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819932. URL https://doi.org/10.1145/2818048.2819932.

G. Richards. Here is the news [Automation Broadcasting]. *Engineering Technology*, 5 (9):36–38, June 2010. ISSN 1750-9637. Conference Name: Engineering Technology.

T. Salvador, J. Scholtz, and J. Larson. The Denver model for groupware design. *ACM SIGCHI Bulletin*, 28(1):52–58, Jan. 1996. ISSN 0736-6906. doi: 10.1145/249170.249185. URL https://doi.org/10.1145/249170.249185.

A. Salzmann, F. Guribye, and A. Gynnild. "We in the Mojo Community" – Exploring a Global Network of Mobile Journalists. *Journalism Practice*, 0(0):1–18, Apr. 2020. ISSN 1751-2786. doi: 10.1080/17512786.2020.1742772. URL https://doi.org/10.1080/17512786.2020.1742772. Publisher: Routledge _eprint: https://doi.org/10.1080/17512786.2020.1742772.

C. Schenk, S. Schimmler, and U. M. Borghoff. Interoperable Access to Video Content as a Basis for Collaborative Video Editing. In *2016 International Conference on Collaboration Technologies and Systems (CTS)*, pages 233–240, Oct. 2016. doi: 10.1109/CTS.2016.0054.

K. Schmidt and L. Bannon. Taking CSCW seriously. *Computer Supported Cooperative Work (CSCW)*, 1(1):7–40, Mar. 1992. ISSN 1573-7551. doi: 10.1007/BF00752449. URL https://doi.org/10.1007/BF00752449.

K. Schmidt and C. Simone. Coordination mechanisms: towards a conceptual foundation of CSCW systems design. Computer Support Coop Work. J Collaborative Comput 5(2/3):155-200. *Computer Supported Cooperative Work*, 5, Nov. 1999. doi: 10.1007/BF00133655.

F. Shipman, A. Girgensohn, and L. Wilcox. Generation of Interactive Multi-level Video Summaries. In *Proceedings of the Eleventh ACM International Conference on Multimedia*, MULTIMEDIA '03, pages 392–401, New York, NY, USA, 2003. ACM. ISBN 978-1-58113-722-4. doi: 10.1145/957013.957096. URL http://doi.acm.org/10.1145/957013.957096.

F. Shipman, A. Girgensohn, and L. Wilcox. Authoring, Viewing, and Generating Hypervideo: An Overview of Hyper-Hitchcock. *ACM Trans. Multimedia Comput. Commun.*

Appl., 5(2):15:1–15:19, Nov. 2008. ISSN 1551-6857. doi: 10.1145/1413862.1413868. URL http://doi.acm.org/10.1145/1413862.1413868.

S. Sirén-Heikel, L. Leppänen, C.-G. Lindén, and A. Bäck. Unboxing news automation. *Nordic Journal of Media Studies*, 1(1):47–66, 2019.

T. H. Soe. AI video editing tools. What editors want and how far is AI from delivering? *arXiv preprint arXiv:2109.07809*, 2021.

T. H. Soe, F. Guribye, and M. Slavkovik. Evaluating AI assisted subtitling. In *ACM International Conference on Interactive Media Experiences*, IMX '21, pages 96–107, New York, NY, USA, June 2021. Association for Computing Machinery. ISBN 978-1-4503-8389-9. doi: 10.1145/3452918.3458792. URL https://doi.org/10.1145/3452918.3458792.

N. Takhirov and F. Duchateau. A cloud-based and social authoring tool for video. In *Proceedings of the 11th ACM symposium on Document engineering*, DocEng '11, pages 65–68, New York, NY, USA, Sept. 2011. Association for Computing Machinery. ISBN 978-1-4503-0863-2. doi: 10.1145/2034691.2034705. URL https://doi.org/10.1145/2034691.2034705.

A. Tanenbaum. The Tanenbaum-Torvalds debate, 1999.

J. C. Tang and S. L. Minneman. VideoDraw: a video interface for collaborative drawing. *ACM Transactions on Information Systems (TOIS)*, 9(2):170–184, 1991.

S. Taylor, S. Izadi, K. Ozenc, and R. Harper. VideoPlay: Playful and Social Editing of Video using Tangible Objects and Multi-touch Interaction. Oct. 2007. URL https://www.microsoft.com/en-us/research/publication/videoplay-playful-and-social-editing-of-video-using-tangible-objects-and-multi-tou

L. Terrenghi, T. Fritsche, and A. Butz. Designing environments for collaborative video editing. 2008.

F. Thalmann, G. Fazekas, G. A. Wiggins, and M. B. Sandler. Creating, Visualizing, and Analyzing Dynamic Music Objects in the Browser with the Dymo Designer. In *Proceedings of the Audio Mostly 2016*, AM '16, pages 39–46, New York, NY, USA, Oct. 2016. Association for Computing Machinery. ISBN 978-1-4503-4822-5. doi: 10.1145/2986416.2986445. URL https://doi.org/10.1145/2986416.2986445.

A. M. Webb, C. Wang, A. Kerne, and P. Cesar. Distributed liveness: Understanding how new technologies transform performance experiences. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 432–437, 2016.

A. Weilenmann, R. Säljö, and A. Engström. Mobile video literacy: negotiating the use of a new visual technology. *Personal and Ubiquitous Computing*, 18(3):737–752, Mar. 2014. ISSN 1617-4909. doi: 10.1007/s00779-013-0703-x. URL https://doi.org/10.1007/s00779-013-0703-x.

O. Westlund and S. Quinn. Mobile Journalism and MoJos, Aug. 2018. URL https://oxfordre.com/communication/communication/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-841. ISBN: 9780190228613.

D. Wood, J. S. Bruner, and G. Ross. The role of tutoring in problem solving. *Child Psychology & Psychiatry & Allied Disciplines*, 1976. Publisher: Pergamon Press.

S. Yarosh, E. Bonsignore, S. McRoberts, and T. Peyton. YouthTube: Youth Video Authorship on YouTube and Vine. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1423–1437, New York, NY, USA, Feb. 2016. Association for Computing Machinery. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819961. URL https://doi.org/10.1145/2818048.2819961.

A. Yuan, K. Luther, M. Krause, S. I. Vennix, S. P. Dow, and B. Hartmann. Almost an Expert: The Effects of Rubrics and Expertise on Perceived Value of Crowdsourced Design Critiques. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1005–1017, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819953. URL http://doi.acm.org/10.1145/2818048.2819953. event-place: San Francisco, California, USA.

K. Zhang and P. Padmanabhan. Feature Overload, June 2011. URL https://papers.ssrn.com/abstract=1856439.

J. Zigelbaum, M. S. Horn, O. Shaer, and R. J. Jacob. The tangible video editor: collaborative video editing with active tokens. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 43–46. ACM, 2007.

F. Zünd, S. Poulakos, M. Kapadia, and R. W. Sumner. Story Version Control and Graphical Visualization for Collaborative Story Authoring. In *Proceedings of the 14th European Conference on Visual Media Production (CVMP 2017)*, CVMP 2017, pages 1–10, New York, NY, USA, Dec. 2017. Association for Computing Machinery. ISBN 978-1-4503-5329-8. doi: 10.1145/3150165.3150175. URL https://doi.org/10.1145/3150165.3150175.

Chapter 6

Article I: Unpacking Editorial Agreements in Collaborative Video Production

Okopnyi, Pavel, Oskar Juhlin, Frode Guribye

ACM International Conference on Interactive Media Experiences (IMX'20), 117-126 (2020)

Chapter 7

Article II: Automation and redistribution of work: the impact of social distancing on live TV production

Okopnyi, Pavel, Frode Guribye, Valentina Caruso, Oskar Juhlin

Human–Computer Interaction, **2023/1**, 1-24 (2023). Published online: 23 Nov 2021.

HUMAN–COMPUTER INTERACTION
2023, VOL. 38, NO. 1, 1–24
https://doi.org/10.1080/07370024.2021.1984917

Taylor & Francis
Taylor & Francis Group

Automation and redistribution of work: the impact of social distancing on live TV production

Pavel Okopnyi [a], Frode Guribye [a], Valentina Caruso[b], and Oskar Juhlin[a,c]

[a]Department of Information Science and Media Studies, University of Bergen, Bergen, Norway; [b]Vizrt, Bergen, Norway; [c]Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden

ABSTRACT

The TV industry has long been under pressure to adapt its workflows to use advanced Internet technologies. It also must face competition from social media, video blogs, and livestreaming platforms, which are enabled by lightweight production tools and new distribution channels. The social-distancing regulations introduced due to the COVID-19 pandemic added to the list of challenging adaptations. One of the remaining bastions of legacy TV production is the live broadcast of sporting events and news. These production practices rely on tight collaboration in small spaces, such as control rooms and outside broadcast vans. This paper focuses on current socio-technical changes, especially those changes and adaptations in collaborative practices and workflows in TV production. Some changes necessary during the pandemic may be imposed, temporary adjustments to the ongoing situation, but some might induce permanent changes in key work practices in TV production. Further, these imposed changes are aligned with already ongoing changes in the industry, which are now being accelerated. We characterize the changes along two main dimensions: redistribution of work and automation.

ARTICLE HISTORY
Received 8 February 2021
Revised 20 September 2021
Accepted 21 September 2021

KEYWORDS
Live TV production; automation; remote work; distributed work; ecologies of artifacts; COVID-19 implications; TV literacy

1. Introduction

Live TV broadcasting is a traditional media format that continues to attract attention from mass audiences. Although the number of viewers always extensively exceeds the number of producers, production still depends on collaboration among a large number of colocated people with various skills and tasks. The interaction depends on fine-tuned audio communication and visual gestures (Engström et al., 2010; Perry et al., 2009).

This industry has been under pressure to adapt its workflows to make use of advances in Internet technologies and user practices (Juhlin, Zoric et al., 2014). The Internet provides additional broadcast platforms, such as mobile media, and it competes with TV in the form of user-generated social media, video blogs, and livestreaming platforms. Both the new technologies for viewing TV and the new ways of producing video are enabled by lightweight and low-cost production tools and distribution channels.

The way the Internet relates to TV media has received increased attention in computer science fields such as Human-Computer Interaction (HCI) and Computer-Supported Collaborative Work (CSCW), as well as from the area of journalism studies. The interest has been both to unpack traditional practices and to study emergent consumption and production practices between the two areas.

At the same time, the transformation of the TV industry has been quite slow. For example, the transition from production conducted by functionally separated teams, including staff with specific

CONTACT Pavel Okopnyi pavel.okopnyi@uib.no Department of Information Science and Media Studies, University of Bergen, Fosswinckelsgt. 6, 5020 Bergen, Norway.

skills, to production conducted by multiskilled individuals, such as in social media, has been slower than expected (Kumar & Haneef, 2018). It has been argued that the latter workflow lacks production quality.

With the advent of the COVID-19 pandemic and concomitant global regulations on increased social distancing, the TV industry faces a challenge that requires reinterpretation of its relationship to Internet technologies and user practices. On one hand, the pandemic has increased the demand for live broadcast TV News (Castriota et al., 2020), and TV consumption has grown during the pandemic as people are trying to acquire information from reliable, authoritative sources more than from social media sites and influencers (Casero-Ripolles, 2020). On the other hand, the social-distancing requirements impede on the ability to work in large colocated teams in studios, causing the number of staff members in the studio to be reduced. In short, they are asked to do the same or more with less of the resources that they usually have at hand (Castriota et al., 2020).

In this situation, we ask how the TV industry addresses this challenge, especially how TV companies relate to emerging Internet technologies and socio-technical practices. There are indications that these companies have adapted to changes and embraced remote workflows during the first months of the pandemic (Andueza-López & López-Plaza, 2020). Further, they became less hesitant in using public network infrastructures for production and distribution (Coche & Lynn, 2020). TV companies also started adopting hybrid production methods, accepting footage filmed with mobile phones and the use of mobile devices in general in the workflow (Túñez-López et al., 2020). Some changes necessary during the pandemic can be seen as temporary adjustments to the ongoing situation, but others might become permanent in key TV production work practices. Heinonen and Strandvik (2020), in an analysis of the implications of the pandemic on service innovation, make a distinction between how service innovation has usually been driven by discretionary activities aimed at securing their position in a market and *imposed innovation* that is driven by resilience and renewal. In this paper, we focus on how the changes can be imposed and temporary, but still be aligned with previously existing currents of change in live TV production.

On a general level, we see the implications and consequences of social distancing as being part of socio-technical change (Bijker & Law, 1992). Such a perspective stipulates that socio-technical changes are not taking place in a vacuum and are emergent and heterogeneous; they involve a number of different actors, artifacts, and social and organizational arrangements. The changes taking place in TV production occur in a social, technological, and organizational context, where all these aspects are intertwined. In this context, the implications can be seen as relating to ongoing changes and available tools. TV production relies on advanced and specialized equipment and tools, and the way these are used depends on the various workflows and practices in different organizations. These tools can be seen as an ecology of artifacts, following Bødker and Klokmose (2012). The ecology of tools in TV production has been described by Guribye and Nyre (2017) as "the way new and old tools fit together, exist side by side, hook into and supplement each other, and how they compete for the same territory, are redundant, and support more or less the same tasks. A new tool will typically find a niche in the overall work practice and the ecology of tools" (p. 1218). As part of the socio-technical changes in response to COVID-19, there is a need to adopt new artifacts alongside adjusting practices and workflows – tools that have been available in the organizational context, but could be characterized as being in a state of *nonuse* of technology (Satchell & Dourish, 2009) in relation to the production practice.

The introduction of social distancing can be seen as a methodological device (i.e., a large-scale "breaching experiment") (Garfinkel, 1967). Governmental laws or guidelines for such a relevant aspect of social life as the physical distance between people is a disruption of the social order. This disruption rendered visible the organization of work in that the established work practice had to be changed and norms guiding these work practices came under scrutiny.

To study these changes, we interviewed twelve TV and video professionals primarily from TV companies from Europe, Asia, the United States, and the Middle East. Based on these solicited accounts, we provide rich descriptions of how broadcast media adapt to the imposed regulations.

We identify two main themes that conceptualize the current socio-technical changes in live broadcast TV production: redistribution of work and automation. First, we discuss changes to the distribution of work and the new workflows and tools that support these adaptations. Second, we discuss how these changes also drive the move toward further automation of workflows. Third, we discuss how these changes should be seen in light of an underlying current of change that was already present before the new sanctions and norms for social distancing were introduced. In this way, the paper contributes to an understanding of the socio-technical changes that are occurring in the wake of the pandemic. In particular, the concept of redistribution is, as such, a way to see these changes as imposed and partially temporary, but still aligned with changes that were already ongoing in current practices.

2. Background

Traditional TV production usually involves a crew of people who create a broadcast for a mass audience (Perry et al., 2009), which stands in contrast to Internet services, which depend on to an increased extent, user-generated content of people working by themselves (Juhlin, Engström et al., 2014). Most of the interest in HCI and CSCW has been devoted to how the latter intersects with TV, such as the emergence of interactive TV (Obrist et al., 2008), second screen use (Courtois & D'heer, 2012), and peer productions (Yarosh et al., 2016). However, some studies unpack the characteristics of traditional TV production and how it relates to new Internet technologies and practices. Here, the area of journalism studies provides salient contributions. In the following, we first turn to the topic of remote work, and then discuss aspects such as current live TV practices, challenges for non-professionals and how they produce video, emergent changes in professional production, technical aspects of TV production, and production automation. These conditions are all relevant to understanding how the social-distancing requirements during the recent pandemic impact live TV production.

2.1. *Remote work*

Remote work has been a long-standing topic in HCI and CSCW. There have been studies of the use of videoconferencing (Mlynář et al., 2018), distributed work (Koehne et al., 2012; Olson & Olson, 2000), and working over time zones (Tang et al., 2011). More recent studies explore how on-demand workers and freelancers are integrating their work into the work practices of organizations (Blaising et al., 2019; Hui et al., 2019). In another study, Karis et al. (2016) showcase how the Google organization has made remote work options an integrated part of its workflows, and how the company relies on videoconferencing as the key tool for mediating these remote work processes. TV production companies, to the contrary, have been slow to adopt such tools in their work (Coche & Lynn, 2020).

2.2. *Collaboration in traditional live TV broadcasting*

A number of studies presented in CSCW and HCI examined the practice of live TV broadcasting. One study unpacked how vision mixers and camera operators coordinate through video-mediated indexical gestures to support mutual orientation and negotiate shot transitions between remote participants (Perry et al., 2009). Another study explored how live broadcasting often includes concomitant postproduction, such as when providing replays in live sport shows. This highly

time-critical activity is conducted in an uncertain setting and is achieved by using some regularities in the content itself (e.g., the live game) and by splitting up video material with different temporal trajectories on different media (audio and video) or on different screens (Engström et al., 2010).

2.3. *Collaborative nonprofessional TV broadcast*

The availability of low-cost mobile Internet technology has been used to invent and investigate whether current professional practices could be handled by nonprofessionals. However, some challenges have been identified. First, the possibilities to include amateur camera operators in multi-camera live broadcasts revealed how professional operators view things on behalf of the mixer's demands, whereas amateurs add their own interest in viewing for themselves. This double attention leads to a lack of interesting shot proposals for the mixer (Engström et al., 2012). Second, studies of so-called professional amateurs (pro-ams) reveal how learning to become a skilled broadcaster is both an individual task and an institutional endeavor. Organizers of recurrent events need to learn what to do for and during broadcasts, as much as individual producers (Juhlin, Engström et al., 2014). Recurrent production, in which producers cover a series of events over time, seems to be more than a quantitative sign indicating mastery of this medium. Rather, their recurrent production seems to be an important condition for both individual and institutional learning, making the broadcasts possible. Third, the new use of live technology for broadcasting events where the boundaries between producers and viewers were highly blurred (Webb et al., 2016) shows the need to develop new technologies for social communication.

In sum, what has been termed "mobile video literacy" requires competencies in handling the technology and only then in creating narratives and a comprehensible story line for an assumed audience (i.e., what camera angles to use, how to cut, and other aspects of video production) (Weilenmann et al., 2014). The specific professional competencies involved in producing video broadcasts are not easily disseminated to other types of users.

2.4. *Video production in social media*

Nonetheless, there has been tremendous growth in video production by Internet users, and such use of video in social media has been topicalized in CSCW and HCI. The research includes the study of variations in practice among different user groups (Yarosh et al., 2016) or the way in which the live aspect of mobile platforms has been appropriated (Juhlin et al., 2010).

Traditional broadcasters have acknowledged such social media content and have developed ways to incorporate such content within their own production. Their challenge has been to find content for their mass broadcasts, as well as add professional graphic content (Arndt et al., 2016, 2017).

Social media production differs from legacy TV production in many ways. Social media producers often do not have teams of professional workers supporting them, thus sometimes they struggle with managing media assets, such as overlay graphics, especially in live streaming (Tharatipyakul et al., 2020). At the same time, live streams, unlike legacy TV broadcasts, serve as "meeting grounds" for viewers (Hamilton et al., 2014) in which they can interact with each other reacting together to the broadcast (Musabirov et al., 2018). The viewers also interact with broadcasters who have to dedicate some of their resources to community management during live broadcast (Wohn & Freeman, 2020).

2.5. *New professional practices*

The availability of low-cost mobile Internet technology has made it possible to use professionals in new ways as well as provide content for new platforms. In an ongoing trend in news production, journalistic work has become embedded in a process "where the content is developed in front of and in some kind of dialog with the audience" (Nygren, 2014, p. 76). News workers are required to be skilled in more tasks and to be more flexible (Nygren, 2014). Such multitasking also includes video

production (Kumar & Haneef, 2018). Interview studies with journalists in TV production show that the increased requirement to provide fast turnaround for multiple platforms and multitasking has led to an increased amount of content but a decrease in investigative reporting (Higgins-Dobney & Sussman, 2013).

A salient question in the area of journalism studies has been to understand this shift's impact on the skills of journalists. Kumar and Haneef (2018) argue that it could be seen as de-skilling because journalists lack time to become proficient with the new tasks and technology, as compared to traditional tools. Based on a large interview study, Nygren (2014) argues that it instead leads to re-skilling and an increased emphasis "on production and adapting content for different channels" (p. 75). Although this has been described as a ubiquitous trend across the industry, a broad interview study in 2015 argued that it did not occur over the entire industry (Phillips et al., 2009). At that time, only a minority of journalists worked across the media. Phillips et al. (2009) argue that this is most effective in small organizations; in larger organizations, such as the BBC, multitasking seems to "to slow things down and to impact quality" (p. 78).

2.6. *TV production infrastructure*

Video is a demanding medium to work with. Transferring video signals from one device to another, processing those signals, adding graphics, and assembling video and audio signals from different sources into a TV signal that is then broadcasted to viewers requires a complex infrastructure based on contemporary technologies, such as video-encoding algorithms (Ksentini et al., 2006) and networking standards (e.g., Ethernet) (Buysschaert et al., 2020). However, the rising demand for higher quality TV (e.g., ultra-high definition TV) has been posing new challenges to broadcasters (Kostiukevych et al., 2019), as the production infrastructure has to provide video signals with higher definition and frame rates (Buysschaert et al., 2020).

In the wake of these challenges, the TV industry is shifting toward a software-based infrastructure, which implies higher flexibility and scalability, and promises to overcome geographical restrictions to allow TV production to be distributed across locations (Lapierre & Al-Habbal, 2018).

One distributed workflow that has been emerging is remote integration (Coche & Lynn, 2020), which shifts much of the signal processing to a centralized network facility where the production crew is located. In live productions using remote integration, only a technical crew is present at the event location to control cameras and make sure the signals are sent to the network facilities. This centralized production hub is a key part of the infrastructure to accommodate such a workflow. In the case of live sports broadcasts in the United States, most big networks have invested in such hubs. In addition to these large investments, some technological challenges with latency and delays in communication associated with reduced production quality have been seen as major obstacles and risks for adopting this new workflow. After the restrictions imposed due to the COVID-19 pandemic, most networks have adopted remote integration. (Coche & Lynn, 2020)

2.7. *Production automation*

To mitigate the complexity of production, TV companies employ innovations such as production automation (Hussein, 2015), that is, primarily but not exclusively, the automation of switchers – big control panels that producers and directors use to switch between different cameras, cue pre-filmed footage, and add graphics. The idea of broadcast automation is rather old (Freilich & Meyer, 1963), and production automation has become essential to news and news-like TV productions (Richards, 2010).

The effects of automation on the production process are complex as it changes the workflow by delegating what was previously human labor to machines. This change is sometimes met with resistance from workers as they perceive automation as a threat to their employment (Linden, 2017). However, studies show that in general (Arntz et al., 2016) and in the media industry

specifically (Linden, 2017), automation results in the redistribution of labor and complements human labor rather than just replacing it. By "codifying" tasks and substituting human workers in performing routine tasks, automation allows people to focus on the creative parts of work. Thus, automation can raise the value of each worker who performs unique tasks. However, many tend to ignore these positive effects and overestimate the automation's negative effects (Autor, 2015).

The automation of tasks may result in the de-skilling of the operators, that is, workers who observe automated processes and intervene if an automated system breaks down. As Bainbridge (1983) notes, the lack of regular practice of the task might result in the loss of skills essential for the operator to identify the breakdown and to correct the error.

Though the effects of automating workflows have been theorized and studied, research is still needed on the intraorganizational frictions that often accompany automation, how these are being resolved, how well-established practices and organizational norms adapt when automation is introduced, and how automation affects the resulting product (Napoli, 2014).

Automation has been a key topic in CSCW since the 1980s. CSCW partly appeared as a critique of the somewhat narrow perspective and conceptualization of work in office automation (Gerson & Star, 1986; Grudin, 1994). Further, Suchman (1993) criticized the concept of workflows in that they presuppose that formal descriptions of work are comprehensive representations of the work done. These fundamental debates from the dawn of CSCW are still relevant (Retelny et al., 2017) and have bearing on how we understand automation. With automation comes the need to formalize work practices and procedures, and due process (Gerson & Star, 1986) and articulation work (Schmidt & Bannon, 1992) will always accompany these formal procedures.

In a recent paper on everyday automation experiences, Fröhlich et al. (2020) discuss some of the implications automation may have on how we interact with technology in everyday life. They emphasize the way the interaction with automated systems is a matter of formalizing procedures and is more similar to end user programming than what is common in interaction with other technologies. The interaction between user and automated system becomes non-continuous and implicit. Thus, automation opens for substitutive experiences, as it frees the user to do other tasks or activities instead of continuously operating the system. While the authors say that these points are mostly relevant for everyday use of technologies, we argue that they are relevant also for more specialized and professional workflows involving trained personnel (Fröhlich et al., 2020).

3. Methods

We have conducted and analyzed twelve remote semi-structured video interviews with video professionals, including personnel working for high-profile TV companies in Europe, Asia, the Middle East, and the United States. In general, in the interviews, we focused on the changes to the working environments and routines of video workers due to social distancing. We discussed how these changes affected their work, and the work of TV companies and organizations in general, from a socio-technological perspective, that is, focusing on the amendments to workflows and tools that had to be applied. We also discussed the future of TV and video work with the participants. We transcribed the interview data and conducted thematic analysis in three iterations.

3.1. *Sample*

A combination of convenience and heterogeneity was applied as a sampling strategy. Heterogeneity was specifically applied to include participants from different TV stations and roles. Our strategy was to find people who worked at legacy TV production organizations and had relevant knowledge and experience to our research topic. We conducted 13 interviews in total. However, one interview did not give any relevant information as the interviewee did not have any relevant experiences with automation or redistribution as we were focusing on in our analysis. Thus, we excluded this interview from the analysis.

Table 1. Interviewees details.

	Country/Location	Current occupation/Background	Organization Type
I01	Northern Europe	Video editor, camera operator/TV editor, operator, producer	Higher Education
I02	Northern Europe	TV web editor, producer	National Broadcaster
I03	USA	TV studio technical director	National Broadcaster
I04	Western Europe	Specialist in automation, producer/TV director	International Media Technology Provider
I05	Western Europe	TV software development manager/TV and Video producer	International Media Technology Provider
I06	Europe	Product manager, field producer	International Media Technology Provider
I07	Southern Europe	TV journalist	National Broadcaster
I08	Middle East	TV journalist coordinator, manager, field producer/ Journalist	National Broadcaster
I09	Southern Europe	Coordinator, Manager	National Broadcaster
I10	Northern Europe	Head of Content, anchor, producer, manager/TV director, Art Director	International Media Technology Provider
I11	Asia	Head of the department for studios	National Broadcaster
I12	Latin America/Northern Europe	Film Director, producer, editor/TV editor	Higher Education/ Independent Filmmaker

The resulting sample (n = 12) included two journalists and ten managers and directors with different levels of expertise. Seven of the interviewees currently work for TV companies. Two worked for TV companies and currently work in the video production industry. Five interviewees worked for TV companies previously but currently work in adjacent industries, such as software and hardware development for TV studios. All interviewees, except I02 and I07, have very diverse responsibilities which are not limited to their formal roles and job positions. For example, I10, who is formally a "Head of Content," also participated in different productions as anchor, producer and manager. In Table 1, we list relevant roles and functions which interviewees performed during the pandemic, as well as their background and types of organizations they are working for. All participants, except I01 and I12, participated in TV content production. I01 and I12 worked with video production during the pandemic in another setting (higher education), but still provided us with valuable variation and supplementary views on the trends and future of production workflows. Participants I04, I05, I06 and I10 were not employed by TV companies during the pandemic; however, they worked closely with TV companies and participated in TV production in different capacities during 2020. Each interviewee had at least several years of experience of work in the TV industry.

3.2. *Data collection and analysis*

Data collection was carried out through individual semi-structured interviews that focused on participants' opinions and experiences in the TV industry. Two participants were interviewed in Italian, whereas other interviews were conducted in English. The interviews lasted from 30 to 60 minutes each. In addition, we had informal conversations with representatives of a TV technology provider to identify new ways of organizing live video broadcasts in the current situation. In the beginning of this study, these informal conversations helped us to develop the interview guide and identify interesting topics which we focused on during the interviews. We also used informal conversations to clarify professional terms and learn about specifics of certain technologies, such as TV networking standards.

In the semi-structured interviews, we mainly focused on the following questions: What is your daily routine and how did it change during the pandemic? How did the workplace change during the pandemic? What tools do you use to communicate and collaborate remotely and how do you use

them? What challenges and opportunities did pandemic introduce to your work? Do you think there will be any long-term consequences of the pandemic?

We must note here that the topics we discussed were considered sensitive as we were touching on subjects such as employment/unemployment and replacement of human workers due to workflow automation. Some interviewees expressed concerns about being identified even by a brief description. Thus, we do not specify their nationalities, nor do we name the companies they work for.

When the participants gave their consent, the interviews were conducted and video recorded using videoconferencing tools such as Microsoft Teams and Zoom. We transcribed, anonymized, and translated the available interview recordings into English. The study was ethically approved by the local institutional review board.

We analyzed the interview data by adopting a thematic analysis methodology in collaborative online sessions, and followed the six steps as identified by Braun and Clarke (2006): (1) reading and searching for meaning and patterns to familiarize yourself with the data; (2) generating initial codes where themes are data driven; (3) searching for possible themes; (4) reviewing themes by refining the data extracts; (5) defining and naming themes; and (6) producing this paper. We analyzed the data in three iterations. In the first, we searched for meaningful patterns and generated initial 72 codes. In the second iteration, we extracted eight general themes, which we defined and named. In the last iteration, we reduced the number of themes by filtering out less relevant and extracting one additional meta-theme (the future of work). The analysis was performed mainly by three researchers in collaborative remote sessions, the fourth researcher participated also in the second iteration.

4. Results

Certain adaptations of the work practice can be supported by adopting existing tools and applications, such as using Skype for live interviews. Other adaptations require deep infrastructural change, for example, the transition to Internet Protocol (IP)-based technologies. At the organizational level, these adaptations can involve changes in scheduling and role assignments or the knowledge and skills required to run operations.

The social-distancing rules disrupted the ordinary workflows in TV companies. Some productions, such as sports events broadcasts, had to be canceled because events were canceled and the content disappeared per se. Other productions, such as TV series, had to be postponed as it became impossible to organize a shooting with so many people involved.

In our data, we identify five major topics. In this section, we present the results in the following manner: First, we describe how TV companies adapted their workflows, specifically, how they employed remote work. Second, we address the topic of transformations in TV literacy. Third, we describe how TV companies adopted production automation and the implications therein. Fourth, we illustrate how TV companies perceive broadcast quality and how this perception changed during the pandemic. Fifth, we present interviewees' speculations about the TV industry's future and the long-lasting effects of the COVID-19 pandemic.

4.1. *Redistribution of work*

TV companies had to rapidly adapt remote production workflows to keep the production running. In this section, we describe how the nature of the work for different TV workers changed with the adoption of remote technologies.

To comply with social-distancing regulations, TV companies had to restrict access to buildings, allowing only "essential" personnel who are required to keep the TV broadcast running:

> I work in the control room. As a technical director, I have to because this is where the big switcher is. In theory, it can be done remotely, but, in practice, we have to be by the console. (I03)

The number of people in a control room had to be significantly reduced, from ten to eleven people to five people. Some workers, who were still required to work on the live production but could not be in the control room, moved to other rooms and communicated with their colleagues via intercoms and Voice over Internet Protocol (VoIP) services. Technical staff, who still worked on-site, had reduced mobility in that they could not change shifts or switch teams. Teams became segregated and isolated from each other.

In addition, the reporting workflow changed, as it became almost impossible to send a filming crew together with a journalist to a site. One of the interviewees, I08, who is responsible for managing field reporters, now tasks them with filming the footage by themselves with mobile phones.

TV anchors started hosting their shows from home, relying on the Internet connections that are shared inside the household. I03 described one such case:

> We have a show and the anchor has a home studio. Right when we went on the air one night, his connection suddenly dropped. His picture would freeze and pixelate. It turns out, it was because his teenage son was downloading software on the Xbox.

This quote illustrates how the socio-technical space created for doing the work remotely – from home – is a shared space. Juxtaposed with the social space of the home, the socio-technical space is dependent on what others in the household are doing. This setting is in stark contrast to how professional TV studios are organized as fully controlled settings for the purpose of recording and transmitting, including controlling noise, light, camera positions and angles, and network connections.

The redistribution of work in TV production involves adopting new tools, for example, video-conferencing services such as Skype, Microsoft Teams, and Zoom, in the production process. Where TV companies previously relied on sending a crew and using specialized equipment, they are now turning to such easily available software services.

The redistribution involves not only using these new tools but also accommodating and supporting remote work, as the social-distancing requirement mandates fewer people crowded together.

> The situation forces us to explore the potential of tools such as Microsoft Teams and Skype—this potential was not fully realized previously. Before that, to interview a person, they would go to this person with equipment and a camera operator—now they just call. (I07)

Further, an interviewee pointed out that this way of producing content actually makes the production process leaner (more lightweight):

> Doing Skype interviews from homes reduces production costs significantly because they need less work with light, camera, sound, and studio decorations. Also, it allows having more people than the studio could accommodate even without social distancing, as there is no need to prepare microphones and places for all guests. (I07)

This redistribution of work came with challenges connected to adopting the production process to incorporate these new workflows:

> I would say that in March, it was hard; today it is not, because the technology companies like Teams and Zoom have made this a lot easier. I just did an interview with five people at [TV Company] at the same time, all from their own homes. So, [I have] separate [video and audio] sources for every single person and myself, and, in March, that would have been very hard to do; now it's just a matter of getting everybody on Microsoft Teams. I'm able to take the Teams, turn that into an NDI source, and then bring that as individual sources onto my switcher. It becomes very easy. A quick level of innovation has happened in the tech industry to enable these things that were very hard and very expensive to do just a few months ago. (I10)

In this quote, several issues are addressed. The interviewee described how they are incorporating the new tools to allow for remote participation. In addition, technical changes have been made to the videoconferencing software to enable signals to be transferred directly into the switcher (a hardware/software piece that selects video and audio sources to be broadcasted). These signals can be

transferred directly to the controller for manipulation and editing. Another point made in this extract is how a transition to an IP-based tool network device interface (NDI) for production is key in this process.

In the process of accepting new tools, the challenge is to maintain the quality of signals and image resolutions that are now transmitted over the Internet while relying on the Internet provider's connections, method of encoding pictures, and so on:

> It can easily be shared outside the production building, even on a pure Internet connection. It technically is possible to have an NDI signal from Milan to New York without a major problem if the bandwidth is good enough. (I06)

The transition to an infrastructure that more easily supports a distributed and remote workflow of TV personnel, such as anchors, is further described by I06:

> Your workflow is restricted inside your walls because of your infrastructure and, because of COVID-19, because people are working remotely, they need to have a signal outside of your building, and it becomes a priority.

While some of the work was done remotely, the redistribution was also done inside the buildings of TV broadcasters, as I03 described: "Some of the producers work from home, we have some of the producers who work from the isolated rooms, but I work in the control room." With this redistribution, they need to keep a back channel open for coordination and communication between the team members:

> In the local studio, we have an RTS intercom panel, but for our connections to the outside world, we can have Internet tools as well as a cell phone or even landline; it depends. (I03)

A transition to incorporating remote work into the workflow of the production also required training and supporting the remote participants. This support included giving instructions on how to set up lighting, place microphones, and use cameras at home, as well as how to best operate the equipment at hand. These instructions varied according to the setup and expected level of quality of the production. In one case for a big sports event production, they put together standardized kits of equipment, including a two-camera setup and a light setup, to be assembled at the remote location:

> For the NFL draft earlier this year, they did the entire thing as a remote production. In fact, it was run out of the director's garage. But they knew who they were going to be talking to way ahead of time. They knew they were going to have an audience of 30 or 40 million people, so they built standardized kits that they could send to every single person they would interview, including a light setup and a two-camera setup, that they could just put down and turn on, and then the people at the NFL could remotely [log] in to it and do all the settings changes just to make sure that it's okay. That way, they can control the quality of the interview. That is what we're doing now too. We give them [interviewees] a small PC and a PPPoE [Point-to-Point Protocol over Ethernet] camera and a light ring, so [they] have the same setup. When I do an interview with them, I can log in to that computer remotely, adjust the camera, adjust lighting and all of that stuff, and set everything up so they don't have to do anything except sit there and talk. (I11)

In this quote, we see how different competencies and responsibilities are redistributed with the transfer of work from studio workers to guests and interviewees. This redistribution also involves adopting new tools into the distributed workflow, in this case, mobile devices and Internet services.

4.2. *Transformations of TV literacy*

The current situation makes explicit alterations of the distribution of skills among those who take part in a TV broadcast, as well as where these competencies are provided. I06 described a situation before COVID-19 where there was an elaborate separation of individuals with specific roles, related to certain skills:

> The main problem in the TV market is that the knowledge of other employees is so vertical. A cameraman is just able to work as a cameraman, and the video mixer as a video mixer.

I07 stated that to interview a person, they would go to this person with just equipment and a camera operator. This method of distributing skills among all involved (i.e., interviewees, journalists, and technicians) is in flux. Our interviewees gave accounts of TV productions where skills are redistributed. When it comes to the journalists, I04 stated that they have obtained new skills in handling technology, such as using digital applications:

> When I started working with journalists, they didn't have any clue about anything technical, but today anyone knows how to run an application and how to make a clip.

The redistribution of skills is most evident in how the interviewees are delegated the task to be both cameraperson and technician:

> In some sense, the work became easier because before we would be required to send a filming crew if we wanted to do an interview. Now we just call the person and ask them to record the answers with their phones and send the file. (I08)

When it comes to the technical staff with specialized skills, the situation is also changing:

> I think at some point the big change will be really in the culture behind the employees in TV. I think the new generation that TV channels will look for, especially because of what happened […] you will look for a guy that will [have skills to] manage a video gallery with audio mixers, vision mixers, and will have a clue about how to use a camera. (I06)

Thus, the technical staff is required to extend their skills beyond those of previous functional separations in the organization. The skills have been required in two different ways, that is, either through specific training in broadcasting or by engaging with consumer technology and services online.

The journalists that work with remote productions "need to have the skill, people that are able to use a camera, to edit, to send you the content" (I06). They are occasionally provided with formal training. I08 stated that they "did a two-day course for journalists teaching them how to use smartphones as cameras." As discussed previously, the redistribution of work depends on new skills among the interviewees. Journalists have become instructors and producers, and the interviewees have become production assistants and camera operators:

> We had to teach reporters how to film with phones so the reporters can teach guests how to film themselves with phones when recording answers. We teach them how to place the camera, what angle there should be and distance, and where to sit so the lighting would be good enough. (I08)

The skills that are used by interviewees, journalists, and technical engineers in current TV production draw in various ways on the participants' previous engagements with consumer technology and online services. Skills required online also change the competence of the technical staff, from legacy TV specialists to social media content creators, such as YouTubers. I11 described it further:

> I'm more interested in finding new, creative ideas on how we can implement things, for example, if they do YouTube on the side, they know smarter and cheaper ways to actually produce a video compared to someone who went through a traditional [production] route. They'll be like "oh I need to have very good quality, and I need to use this equipment specifically to produce this because that was what I was taught" […] I think people who went through that [nontraditional] route will have a different mindset. I think it's a good thing for people in TV to have an understanding that maybe we can tap on the skills from people who are from that route, not the traditional route.

Finally, the increased distribution of work means that those who are in charge of specific tasks might lack specific competencies, that is, skills that had been available in previous productions but are no longer at hand within the new type of broadcasts. I11 pointed specifically to how graphics require special skills that are not provided for in a distributed setting:

> We have to think how we can be better than the online platforms. We are paying more money, so how can we improve our content? Something that the [social media creators] can't provide. We have a lot more graphics

> presentation. We use a lot of graphics to present our stories. It's not something you can easily do in a home setup. It's something you can do in a big studio.

Furthermore, it is argued whether YouTube production experience is enough to produce a professional broadcast. I06 referred to TV as "a language" and stated the following:

> At some point, you need to do the proper thing, even if you use mobile phones. You need to know how to use the light and how to use the microphone because you need the proper quality.

In this quote, I06 emphasizes that there is a certain threshold in skills between professional video workers and other content creators.

4.3. *Automating production*

During the pandemic, some TV companies turned to automation as one of the instruments that allowed the production of content with fewer people in the studio:

> BBC is entirely automation, CNN is entirely automation, and Fox News is entirely automation, so it's become the norm now to go with this. (I10)

To understand the impact of COVID-19 and the role of automation, it is necessary to recognize live TV production as a labor-intensive process that involves many roles and tools. I03 described the control room in their TV company as the following:

> You have the person that is controlling the switcher, that's one or two people. You have the person that is controlling the audio, that's another person. You have a person that is controlling the graphics, that's another person. You have a person in the studio that is running the camera, that's another person. You have an engineer that is watching the video signals, that's another person.

The social-distancing restrictions have had direct consequences for this workflow organization. The inability to have workers in their usual places combined with the rising demand for new content created a situation in which automation becomes a very welcome solution:

> When you suddenly cannot have a lot of people in the control room, then what's your option? You still need to produce content because actually, at this moment, you have a captive audience that is sitting at home eager to have something to consume. You have to go to automation. It's the only way to survive and continue being able to produce content. (I10)

Automation in this context is about programming software to handle some tasks that normally are carried out by human workers:

> These roles can be combined into a single action by automation. So, when I say "take on camera 1" in the automation system, the microphone for camera 1 automatically pops on, and the camera moves over by robotics to the place that it needs to be. A graphic could come up saying that person's name on camera 1 at the same time, and the switcher is going to change a source. (I10)

The possibility of doing this has increased since the cost has decreased. Automation draws on a continuous digitization of production functions, which has reduced the need for specialized and expensive hardware:

> The smaller broadcasters are the last to really get into this because automation used to have a very steep price point in the past. It was very complex. […] The idea of software-based automation is lowering the price point and lowering the entry point. […] and it's not just the automation being software, it's the fact that all of the components are becoming software. (I10)

Furthermore, automation draws on an understanding and a practice where the work described previously is seen as repetitive and even boring:

> The show is always the same. [It is] the content that is different. You have the intro in the same way. You have the first shot in the same way. The camera is moving in the same way every day, four times per day. (I06)

When I06 referred to the show always being "the same," this is from a production perspective, which differs from the "content" that varies over time. I06 stated that even new shows repeat the same production principles: "It's just the same new show; it is so boring. It is always the same concept."

The interviews also discussed the consequences of this socio-technical shift. First, it will lead to broadcasts that in their structure are decided before they go live. This in turn requires that the current informal rules are made explicit and are decided upon in advance. I03 stated that "producers will be responsible for creating more of their content, not only for video clips but also for graphics. They would have to have a more structured rundown." I11 stated that their company tried to employ automation for years, and one of the major restricting factors was the negative reaction of the studio workers who "felt very handicapped with automation, so they couldn't produce the show that they wanted."

Second, the technology can lower costs for TV production, which would lead to redundancies in that specific case, as argued by I03:

> I have a feeling that this might cause lasting changes to the way we work because we are now finding out that we can automate some of our shows and eliminate staff.

However, I04 argues that there is a latent demand on the industry to produce more content, which will compensate for the leaner production:

> People most of the time are afraid about automation because they think that it's going to do everything, and it will take their jobs, which is not really correct. I know that the financial part they see, "oh we're gonna take out some," which is not wrong, but each TV station wants to produce more, and if you want to produce more, you cannot […] think about doing production without automation in the future.

Third, the programming of the machines could have been done in the wrong way, or the broadcast could need adjustments that are out of the scope of the current automation. However, I10 argued that they are less prone to errors than the current setup:

> [Company] has done a study on the effects of automation for their on-air errors and they have found that year to year going from nonautomated to completely automated, their errors have been reduced.

At the same time, handling errors and breakdown in the automated TV production requires a setup of numerous backup systems, procedures and countermeasures:

> We have multiple layers of backups. Number one [is], if, for example, the [automation system] server itself is having issues, we can swing it to our backup server immediately: it will take five seconds and then there would be minimal to no impact on air. So that is one layer of backup. If the automation system goes down, we have a procedure on how we swing to manual operations as well. The tricky part about that is getting the manpower in. If it is during the day, it's not too bad: I can just pull people in and run the show. In a really worst-case scenario when the entire control room is down - that's a whole different backup system. (I10)

4.4. *Broadcast quality*

Another topic addressed by several of our informants was how the new ways of working had implications for the quality of the product they were making. They reported a changed attitude toward the need for high quality in the content. Many said they had to accept a lower quality due to the new tools. "I do the best with the equipment I have," as I01 put it.

There is a distinction, however, between production quality and journalistic/content quality:

> What broadcasters are finding out is that in the past, they were very concerned about the quality of the video and the quality of the interview. They wanted to have this nice 4K camera shooting their interview subject because it's going to look better for the presentation; the audience is going to like it better. But really, the audience is just interested in what the person has to say. So even though we're not getting that same 4K quality of having the production crew on-site at this person's house, uploading to the satellite and coming back, and all these things, we're just bringing them in via Zoom or Skype or something. They're still getting the same story

> and that has become acceptable for the audience. And that has also changed, how much production value broadcasters actually have to put toward something in order to bring the story to their audience. (I10)

As we can see in this excerpt, there is a changing attitude toward what is important for making quality TV, which is not only tied to notions of using the best, top-rated, industry-grade equipment. Further, interviewees pointed out that there is also a shift in the expectations of the audience. They are getting used to seeing videos and clips that are not made with the best resolution or perfect lighting. This trend has been following the social media distribution of content:

> The quality is lower than it used to be, but it was getting lower anyway because more and more reports and interviews were done with these technologies, mobile, for example, breaking news. So, people were getting used to seeing reports and interviews in lower quality even before Skype became a mandatory tool. (I08)

Still, some also pointed out that there is a need for a certain threshold in quality. There are still some standards that they try to meet to stand out among other video producers, such as pro-am content creators. For example, I11 explained how they, the national broadcaster, use graphics as a tool to stand out from YouTubers and social media creators (see earlier comment from I11).

The new experiences with lower resolution, using videoconferencing tools and mobile phones, put the demand for expensive equipment and setups under scrutiny:

> Because of the union of the TV editors, they had to create the editing room with the proper audio noise leveling, so they spent something like 100,000 pounds for every editing room because of the noise, because of the light. Now, during COVID-19, all the editors are working from home with the noise, with the light, with the kids, and they are doing the exact same quality of production, so why should we be back? (I06)

We mentioned previously how TV companies try to preserve quality by teaching remote participants how to use their mobile phones as cameras. With the varying quality of signals coming to the editor, there is a renewed need to manipulate and process the signals:

> You have to deal with every single connection now. I have some color correction tools here, and I could do a live color correction to make it a little bit better, but it's still … you have what you have. (I10)

The acceptance of a lower quality when moving to remote production was also a recurring theme, as indicated by the "you have what you have" statement.

4.5. *Future of work*

When asked about possible implications for the future of the TV industry, each interviewee said that their workflows will definitely change in some way when social-distancing restrictions are lifted. In this section, we provide an outline of the possible changes that will occur in TV and video production.

I01 said that they would consider working from home when it comes to small projects, that is, short videos and stories, because small projects can be done completely remotely, without physical meetings with other participants. I06 speculated about the need for TV companies to have all the workers they used to have in the office:

> Right now, there is less than one-third of the people inside a building, and they do exactly the same level of production, so why should all those people be back in the building? Is there a reason for that? Is there a reason for all these people to travel again? To be in the car in a traffic jam?

Others, though, said they would prefer to return to the work in the studio "because we need to work as one team" (I08). However, that does not imply that their workflows will remain intact. I05 agreed with that, saying that it is easier to have work discussions and meetings face-to-face.

I08 suggested that their production company will rely more on mobile-produced content, that is, footage filmed with mobile phones by journalists, pro-ams, and amateurs. This more generalistic approach to TV production, which demonstrated its viability during the pandemic, is something that TV companies will look for in their personnel. For example, I11, whom we quoted previously, is

more interested in finding new, creative ideas to implement and is more interested in people with YouTube production experience than those with a traditional production background. I06 expanded on this topic, saying that the hyper-specialization of TV professionals ("vertical knowledge" [see earlier quote]) is a major problem for TV companies, including how it will be addressed in the future by hiring more multiskilled workers.

I11 also speculated that TV companies would be looking for video workers who possess more diverse skill sets and are familiar with nonprofessional and pro-am production tools and workflows, such as YouTubers and livestreamers, because TV companies are starting to accept the quality that amateur/nonspecialized tools provide:

> This is what the broadcaster will look for in the future. In terms of the broadcast quality, in the last three months, everybody was using Skype on a mobile phone because the content wins over what you see. The content is much more important than the quality you show.

Over the course of the pandemic, automation has become more and more popular and prominent. For example, several big TV companies with which I10 works have gone completely automatic over the course of just several months.

I06 said that there is an ongoing discussion among professionals: "We talk much more about remote production, especially for sports." I11 speculated about the future of remote production, remarking that not needing to travel to film an event would possibly become "a new norm" if the online event format maintains its popularity in the future:

> I feel like people are fond of these online events. We no longer have to travel to attend conferences. It's much cheaper and much easier to organize and that's why we are looking at a more long-term plan for how we can move forward without all these restrictions and without tying down our studio resources for this.

Aside from events, conducting interviews will also happen online more frequently

> Our reporters no longer conduct interviews out in the field. As much as possible, they try to do it via Zoom, so they're able to record it on their PC and send it to be edited. I think that those are some changes that would remain, but of course when it comes to our VIPs, we will still send a crew down to do the interview, but I don't really see us reversing any new changes that we implemented during this period. (I11)

The problem of limited resources, such as fewer incoming video sources (e.g., Skype calls) is addressed by the shift in technologies that continues to happen during the pandemic. In particular, there has been a shift from hardware-restricted to software-based protocols, as I04 elaborated:

> This is where we are going, and then we can talk about the NDI and the future. I don't know if it [the future] is the NDI, but it's not SDI [serial digital interface] anymore. SDI, I think, will just disappear very slowly. We still have some years with this SDI because the whole world is using SDI. It's not [going to change] in one day.

Another possible shift that is being adopted in some companies but not yet embraced by big TV studios is the employment of cloud-based production tools, which allow TV workers to produce content from various locations via the Internet. It is suggested that cloud-based production technologies can dramatically change the work for TV professionals:

> If you're in the US and you have the best person to do the task, but he lives in LA and you have to fly him over to New York four times a week to do that task because he likes living in LA, suddenly [by using the cloud] he doesn't have to fly anymore. (I05)

With these newly established remote workflows, participation that does not require traveling becomes more readily available. The benefits from cloud-based technologies are apparent, as they allow production to be scaled according to requirements:

> The power of software is that it allows you to have either one copy running or fifty copies running on [the cloud platform]. […] You can scale up and down based on the needs of one specific day. (I06)

However, not all TV companies are ready to adopt cloud-based production due to cybersecurity concerns:

> It's very interesting software, but our concern is always cybersecurity. We are a national broadcaster; our cybersecurity measures are very strict. Even internally in the building, we have a lot of restrictions, let alone producing a show from an outside location and tapping into our network. Those are some of the challenges that we will experience if we explore this route, but definitely, it is something that we want to look into. (I11)

In sum, our interviewees agree that after COVID-19 pandemic is over, TV production companies will change their working practices and accept more general public technologies in their production workflows, as well as new technologies, such as cloud-based, which allow remote workflows.

5. Discussion and implications

The statistics of live news broadcast during the pandemic shows (Castriota et al., 2020) an increased demand for live news TV. At the same time, the restrictions on social distancing meant that some of the previously available production resources are no longer available, which represents a challenge for the industry i.e., the COVID-19 challenge. In the following, we discuss how the TV industry address it through redistribution and automation, as well as the long-term implications of these strategies and how they interact.

5.1. *Addressing the COVID-19 challenge*

As argued by Guribye and Nyre (2017), live TV productions depend on an ecology of advanced and specialized equipment and tools, which are used in various workflows and organizations. The strategies selected add to the ecology of tools for TV production (Guribye & Nyre, 2017) in different ways, drawing on alterations of social practice as well as changes in the use of technology. In the following we summarize the identified general strategies, i.e., automation and redistribution.

5.1.1. *Redistribution of work, control, and skills*

In CSCW and HCI the concept of remote work is important to denote how mediated interaction across distance is happening (Karis et al., 2016; Koehne et al., 2012). Instead, we use the term "redistribution" to characterize what in many cases are adoption of remote work in live TV production. The concept is more encompassing than remote work and is meant to emphasize how practices change and how these changes involve more than adopting remote work per se and specific video conferencing tools. The concept of redistribution of work characterizes how practices are changing over many dimensions, and in particular it denotes the imposed and temporary adaptations done as a response to the social distancing measurements.

In this case, redistribution introduces new tools and practices into the ecology, i.e. ubiquitous Internet services. The use of Internet services such as social media for user-generated content, which lately has come to include video content, is abundant. When camerawork is redistributed to journalists, guests and other nonprofessional operators, it draws on this widespread literacy. The TV industry hooks into a different ecology of tools which includes general-purpose devices such as mobile phones.

In the studio, the remaining staff stuck to their traditional tasks but spread out in the office building, which was made possible by the use of voice communication technologies. The story is different when it comes to professional work outside of the studios. Here, the journalists had to learn new skills when camerapersons were no longer available. This shift from specialized tasks to multiskilling follows a trend that has been long discussed in research (Kumar & Haneef, 2018; Nygren, 2014; Phillips et al., 2009) and where there has been a discussion on the barriers to such a change.

Another, more remarkable, redistribution of work occurred outside the studios, when guests in TV shows were asked to film themselves. They instruct their remote guests on how to sit in front of the camera and where to sit in the room in relation to light and sound sources. In some cases, they send special hardware kits with light sources and cameras and provide detailed instructions

regarding how to use those kits to ensure the good quality of the produced content. In these cases, the professional work of an entire team was redistributed to a nonprofessional person behind the camera.

The redistributions occurring outside the studios also came with new demands, that is, providing journalists and participants with new skills. Journalists were provided with short courses where they were taught how to use their mobile phone to film the footage by themselves. Nonetheless, it was noted that the redistribution depended on learned skills that had been acquired in situations other than during formal education. Interviewees noted that during the pandemic, the redistribution could draw upon the fact that skills in using Internet services and mobile devices were ubiquitous. This was evident in I04 stating that all journalists these days know how to use an "application."

Still, we argue that the extent to which it depends on an increased familiarity of users, outside of the studios, needs to be recognized. Internet users have become increasingly proficient in photography through using services such as Instagram and Facebook (Juhlin, Zoric et al., 2014). They are even becoming skilled in producing moving image productions through videoconferencing services, Twitch, YouTube, and so on (Hamilton et al., 2014; Yarosh et al., 2016). Thus, the redistribution depended on the skills acquired for reasons unrelated to TV production.

The redistribution of specialized technical tasks, such as from camerapersons to journalists and nonprofessionals, implies that professionals decrease their control over broadcasts and must tolerate other standards that are more aligned with Internet practices, such as a TV host needing to share the local network with a gamer. A similar shift of control also occurred with the increased use of Internet services in broadcasts, such as videoconferencing technologies (e.g., Skype and Zoom). It is notable how these Internet service providers swiftly adapted their protocols to make them accessible for TV technologies such as NDI when the latter needed it during the first phases of the pandemic.

We also note that graphics production might have been distributed in the office facilities of the broadcast studio to meet social-distancing requirements, but the technology remains under control of the traditional broadcasting ecosystems. When camerawork, lighting, audio recording, and so on are shifted to ubiquitously available technologies and production is redistributed from specialists to multiskilled professionals and even to amateurs, it is argued by our interviewees that graphic production will continue to distinguish broadcast TV from online video production.

5.1.2. *Automation, planning and error reduction*

Automation is a prevailing trend that becomes increasingly available in studio work, when the industry implements standard digital technologies of various kinds, i.e. technologies which can be used for TV production through software applications. It draws on a shift from specialized digital hardware to general-purpose hardware, as well as on new specialized communication protocols that can be used on the Internet.

Previously, such automation has met resistance in the TV industry since, as noted by our interviewees, TV workers were "afraid" and "felt very handicapped." This experience of lack of capability and control over the production process might stem from the transition from continuous operation to implicit interaction with an automated system (Fröhlich et al., 2020). However, the imposed social distancing restrictions seem to have decreased institutional hesitations. We've been told about major TV companies automating much of their production during the pandemic.

Automation changes the "normal" working process, as TV workers have to adopt new workflows that require much more work in the preproduction phase. For example, workers are required to plan the work and develop scripts for automation systems, that is end-user programming (Fröhlich et al., 2020), and establish backup systems and breakdown-handling procedures. Thus, the production process becomes more formalized. This formalization should also be approached with caution, as it might overlook how the emergent work requires due process (Gerson & Star, 1986) and articulation work (Schmidt & Bannon, 1992), notwithstanding how detailed the formal representations are (Suchman, 1993). This formalization also reduces the room for human errors during live broadcast as production automation implies TV workers giving control over broadcast to a machine that

executes a set of predetermined commands. The participants in our study claimed that automation actually reduced the number of errors. We do not know, however, if the errors are getting more or less severe. Further, as I10 mentioned, there is a need for establishing complex backup systems and finding substitute workers rapidly in the case of automation system breakdown. At the same time, our informants say that automation allows TV companies to redistribute workers, who are now freed from continuous operating by automation, between productions and can produce more content as a substitutive activity (Fröhlich et al., 2020).

In sum, taking automation and redistribution into account, we see that the handling of the COVID-19 challenge has a broader impact on the work outside of the studio than inside. The studio workspace seems to be a resource that is used to resist multiskilling, or a barrier, if the trend is seen as positive.

5.2. *Long term implications for TV workflows*

The question is then what will stay of these changes and what will go back to how it was before the pandemic. In other terms, the COVID-19 restrictions on social distancing can be seen as a push (Nimako & Ntim, 2013) that fuels those trends. But what happens when this push is no longer available? We have identified a number of topics that can either pull, or more vaguely "moore," (Stimson & McCrea, 2004) the industry back toward its traditional workflow. The factors that pull the industry toward both automation and redistribution includes those of availability of general-purpose technology; multi-tasking, wide skilling and changed viewing experiences. The factors that "moore" the work flows back into traditional studio work include lack in skills, need of control and data security.

5.2.1. *Factors that push the workflows toward redistribution and automation*

We argue that the TV industry will continue to be pushed in the current direction because of changes in technology, production, and viewing practices. As mentioned above, both automation and redistribution, from a technical perspective, draw on a shift from specialized to general-purpose technologies.

This shift *lowers the costs* of TV production but also makes it possible to *do new things*, e.g., scripting studio work in software or a more mobile journalistic practice. From a production practice perspective, there is a push by an *increasing number of producers* who are getting skilled in live video broadcasting online (Hamilton et al., 2014). From a viewer's perspective, there is another push that supports the redistribution trend. As I10 noted, the TV industry has always striven for higher broadcast quality. However, the Internet ecology has created *viewing experiences with lower image qualities*, made available on consumer Internet services, including mobile devices and Wi-Fi connected laptops. We argue that the "tolerance" of the quality of pandemic TV productions depends on this new ecology of tools and Internet practices.

5.2.2. *Conservative factors*

There are obviously factors that can make the industry pull back, or at least moore the changes back, to the state the industry was in before the pandemic. Participants' current production *skills might not always be sufficient*. As we mentioned earlier, during the pandemic, TV workers had to instruct their remote guests, as well as provide them with better equipment. It might be argued that this has not been sufficient. Having people back to the studio will make it *easier to control* all these features. The resistance to automation, might re-occur when the physical locations are open as before, as might the resistance to the lower broadcast quality imposed by the use of general-purpose Internet services and mobile devices. This resistance can become stronger as the physical restrictions are ended. *Lack of data security* is another important concern when considering cloud-based production technologies from becoming widely accepted in the TV industry. Cloud-based production implies the possibility of redistribution of traditionally in-house specialists, such as graphics operators and producers, to

remote locations. This redistribution introduces new uncertainty as the production will rely heavily on public infrastructure, the Internet.

5.2.3. *Interdependecies*

We have previously discussed automation and redistribution either as parallel trends, or in some sense even dependent on the same underlying Internet technology. In the following we will discuss if there are characteristics that make it difficult for both to sit together at the same time. We learned that both emerged as strategies to address the Covid-challenge. But will they continue to emerge side by side or are they affecting each other? This is perhaps the most critical topic to address for the industry. Automation is very much an insiders' trend where the professional TV producers programme the machines to do their routine work. Such routines emerge out of their standardized practices when making a recurrent broadcast. The redistribution trend is rather decreasing the professionals' control of the broadcast and increasing the uncertainty as to what will happen in the broadcast. Hence, automation is less likely to be of use in these situations.

Thus, on a general level, these trends are of a different kind. Automation and redistribution do not really go together hand in hand because automation is giving control to machines and formalizing the workflows, whereas redistribution implies that TV professionals lose control to other people, who are less trained and operate in unpredictable settings. It can create an opposition: redistribution pushes things out of the studio, allowing workers to work from home or remote locations, but automation pulls stuff to be going on in the studio, in a controlled environment. If we assume that the redistribution trend is the strongest, then the automation of studio work will be of less interest in the future.

5.3. *Implications for TV work*

Our findings confirm previous research on automation. For example, Autor (2015) suggests that people in general tend to overestimate the negative effects and neglect the positive effects of automation. Linden (2017) suggests that the introduction of automation to journalistic work does not result in job losses, as there is more to journalism than just writing stories, and only entry-level jobs might be in danger.

With the broader acceptance of TV production automation, we speculate that some jobs might become rudimentary, such as the job of a graphics producer during a live TV broadcast. However, the work will still exist in the form of preproduction graphics handling. In line with Bainbridge (1983), we see that automation has certain effects on skill requirements for jobs, but we do not find evidence that TV studio jobs are being de-skilled. Instead, they are rather being "re-skilled" (Nygren, 2014), as the producer in the automated studio needs fewer special skills to run the production but needs to be more polymathic.

Further, we speculate that the trends toward more versatile skills and employment of generally available Internet technologies will lower the barrier to new people and organizations which want to enter the TV industry. These trends are not new (Phillips et al., 2009). However, they have significantly accelerated during the pandemic. Along with I11, we speculate that legacy TV companies will start looking more for talented people outside of the industry who are familiar with Internet-based technologies and prove their creative potential via Youtube and other video-based social media sites. At the same time, the entry requirement for new employees will decrease as the required skills and expertise will be more related to Internet-based technologies that are already available to the general public.

The imposed adoption of general-purpose tools forced TV companies to reevaluate different aspects of TV work, such as broadcast quality and skill requirements. With such changes in place, TV companies could start taking advantage of services, tools and practices for video production and delivery commonly associated with video-based social media, such as Twitch and YouTube. For example, live streamers employ chat applications to connect with the audience and allow the viewers

to participate in the live video content production. These interactions between video producers and consumers and practices and social norms around them have been addressed in HCI and CSCW research (Abokhodair et al., 2015; Hamilton et al., 2014). Interaction with the live audience entails development and maintenance of audience management mechanisms which help broadcasters keeping track of the audience history, aggregating and visualizing various data, and categorizing the audience (Wohn & Freeman, 2020).

From the production perspective, live streamers are much leaner, as they seldom employ production teams and any specialized software and hardware (Wohn & Freeman, 2020). This practice of single-person inexpensive production might also be adopted by some TV companies as they start accepting lower quality and adopting automation.

5.4. *Limitations and further research*

This work aims to shed light on processes which happen inside the TV industry during the COVID-19 pandemic. While we were able to identify important and interesting trends from our rich material, there are some inherent limitations following our choice of method and the sample. We conducted an exploratory study based on interviews with representatives from different media organizations including world-renowned TV production companies. However, we do not know to what extent our findings and analytic generalizations are relevant to other legacy media organizations, including local TV channels. In the data gathering, we mostly focused on workflows inside TV studios and we show that some of them are being redistributed to the outside via, for example, interview kits that are being sent to show guests. How such changes were experienced by guests on TV shows and how the audiences have reacted to these changes in production quality is outside the scope of this paper. Further research is needed on *how* legacy TV companies adopt internet tools and practices into their workflows. Such studies should include observation of actual production practices. Furthermore it would be of interest to CSCW and HCI to study more closely the *interaction, coordination and planning practices* involved in automating live TV productions, including breakdown handling and establishment of backup systems and procedures.

6. Conclusion

In this paper, we looked at how regulations in the wake of the COVID-19 pandemic have influenced practices in live TV production. The transformations can be seen as temporary and imposed, and we focused on how they align with and accelerate existing changes. In particular, we discussed how the ecology of tools and practices in live TV production is being supplemented by Internet technologies and services. Some adaptations of the work practice require deep infrastructural change, such as adopting new protocols and hardware. Further, the integration of Internet technologies into live TV production will make available a number of innovations from this area. We discussed how a redistribution of work not only concerns the geographical location of people but also the adoption of new skills, new tools, and an adjusted conception of quality in TV production. Further, we discussed how the adoption of automated workflows can allow for streamlined production of content, but that it requires the work to be formalized. The notion of redistribution can contribute to conceptualizing the response to imposed and temporary conditions and how it aligns with ongoing, slower socio-technical changes.

Acknowledgments

The study is a part of the project "Better Video workflows via Real-Time Collaboration and AI-Techniques in TV and New Media," funded by the Research Council of Norway under Grant No.: 269790. This work was also partly supported by industry partners and the Research Council of Norway with funding to "MediaFutures: Research Centre for Responsible Media Technology and Innovation," through The Centres for Research-based Innovation

scheme, project No.: 309339. We would like to express our gratitude to Vizrt's Even Normann and Roger Saetereng. We would also like to thank our interviewees for their input.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Pavel Okopnyi (Pavel.Okopnyi@uib.no) is a PhD Candidate focusing on social aspects of TV and Video production; he has a background in software engineering, Sociology, and HCI.

Frode Guribye (Frode.Guribye@uib.no) is a Professor of Information Science focusing on HCI and the social implications of ICT; his research spans different application areas such as technology enhanced learning, computing and mental health, video production and mobile journalism.

Valentina Caruso (valentinacaruso83@gmail.com) is a Senior UX Researcher with experience in software and web-based solutions; She holds a Ph.D. in Educational Sciences from the University of Fribourg (CH) and an MA in Technology and Language Teaching; Her interests include HCI and learning technologies.

Oskar Juhlin (oskarj@dsv.su.se) is a Professor at the Department of Computing and Systems Sciences at Stockholm University, and a guest professor at the MediaFutures research center at the University of Bergen.

Funding

This work was supported by the Norges Forskningsråd [269790].

ORCID

Pavel Okopnyi http://orcid.org/0000-0001-7034-2733
Frode Guribye http://orcid.org/0000-0002-3055-6515

References

Abokhodair, N., Yoo, D., & McDonald, D. W. (2015). Dissecting a social botnet: Growth, content and influence in Twitter. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* Vancouver, Canada (ACM), 839–851.

Andueza-López, B., & López-Plaza, M. (2020). The TV-production shift during the COVID-19 health crisis: How TV language changed as a state of alarm was enforced in Spain. *Tripodos*, 2(47), 161–172.

Arndt, S., Perkis, A., & Räty, V.-P. (2016). Opportunities of social media in TV broadcasting. *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* Gothenburg, Sweden (ACM), 1–6. https://doi.org/10.1145/2971485.2995346

Arndt, S., Räty, V.-P., Nieuwenhuis, T., Keimel, C., Ibáñez, F., & Perkis, A. (2017). Enhancing use of social media in TV broadcasting. *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video* Hilversum, Netherlands (ACM), 51–56. https://doi.org/10.1145/3084289.3089923

Arntz, M., Gregory, T., & Zierahn, U. (2016). *The risk of automation for jobs in OECD countries: A comparative analysis* OECD Social, Employment and Migration Working Papers . https://doi.org/10.1787/5jlz9h56dvq7-en

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30. https://doi.org/10.1257/jep.29.3.3

Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Bijker, W., & Law, J. (1992). Do technologies have trajectories. In W. E. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 1–16). MIT Press.

Blaising, A., Kotturi, Y., & Kulkarni, C. (2019). Navigating uncertainty in the future of work: Information-seeking and critical events among online freelancers. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* Glasgow, UK (ACM), 1–6. https://doi.org/10.1145/3290607.3312922

Bødker, S., & Klokmose, C. N. (2012). Dynamics in artifact ecologies. *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* Copenhagen, Denmark (ACM), 448–457.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
Buysschaert, C., Descampe, A., & Lorent, J.-B. (2020). Creating bandwidth-efficient workflows with JPEG XS and SMPTE ST 2110. *SMPTE Motion Imaging Journal*, 129(7), 33–40. https://doi.org/10.5594/JMI.2020.3001465
Casero-Ripolles, A. (2020). *Impact of COVID-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak* (SSRN scholarly paper ID 3594133). Social Science Research Network. https://papers.ssrn.com/abstract=3594133
Castriota, S., Delmastro, M., & Tonin, M. (2020). *National or local? The demand for news in Italy during COVID-19* (SSRN scholarly paper ID 3733273). Social Science Research Network. https://papers.ssrn.com/abstract=3733273
Coche, R., & Lynn, B. J. (2020). Behind the scenes: COVID-19 consequences on broadcast sports production. *International Journal of Sport Communication*, 13(3), 484–493. https://doi.org/10.1123/ijsc.2020-0231
Courtois, C., & D'heer, E. (2012). Second screen applications and tablet users: Constellation, awareness, experience, and interest. *Proceedings of the 10th European Conference on Interactive TV and Video* Berlin, Germany (ACM), 153–156. https://doi.org/10.1145/2325616.2325646
Engström, A., Juhlin, O., Perry, M., & Broth, M. (2010). Temporal hybridity: Footage with instant replay in real time. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Atlanta, GA, USA (ACM), 1495–1504.
Engström, A., Perry, M., & Juhlin, O. (2012). Amateur vision and recreational orientation: Creating live video together. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* Seattle, Washington, USA (ACM), 651–660.
Freilich, A., & Meyer, S. (1963). The "STEP" system a unique, low-cost TV automation system. *IEEE Transactions on Broadcasting*, BC-9(1), 16–25. https://doi.org/10.1109/TBC.1963.265906
Fröhlich, P., Baldauf, M., Meneweger, T., Tscheligi, M., De Ruyter, B., & Paternó, F. (2020). Everyday automation experience: A research agenda. *Personal and Ubiquitous Computing*, 24(6), 725–734. https://doi.org/10.1007/s00779-020-01450-y
Garfinkel, H. (1967). *Studies in ethnomethodology* (Prentice Hall).
Gerson, E. M., & Star, S. L. (1986). Analyzing due process in the workplace. *ACM Transactions on Information Systems*, 4(3), 257–270. https://doi.org/10.1145/214427.214431
Grudin, J. (1994). Computer-supported cooperative work: History and focus. *Computer*, 27(5), 19–26. https://doi.org/10.1109/2.291294
Guribye, F., & Nyre, L. (2017). The changing ecology of tools for live news reporting. *Journalism Practice*, 11(10), 1216–1230. https://doi.org/10.1080/17512786.2016.1259011
Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on twitch: Fostering participatory communities of play within live mixed media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Toronto, Ontario, Canada (ACM), 1315–1324. https://doi.org/10.1145/2556288.2557048
Heinonen, K., & Strandvik, T. (2020). Reframing service innovation: COVID-19 as a catalyst for imposed service innovation. *Journal of Service Management*, 32(1), 101–112. https://doi.org/10.1108/JOSM-05-2020-0161
Higgins-Dobney, C. L., & Sussman, G. (2013). The growth of TV news, the demise of the journalism profession. *Media, Culture & Society*, 35(7), 847–863. https://doi.org/10.1177/0163443713495078
Hui, J., Cranshaw, J., Kotturi, Y., & Kulkarni, C. (2019). The future of work(places): Creating a sense of place for on-demand work. *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* Austin, TX, USA (ACM), 487–491. https://doi.org/10.1145/3311957.3359432
Hussein, N. (2015). *Broadcast Automation System: Newsroom Production* [Fi=AMK-opinnäytetyö|sv=YH-examensarbete|en=Bachelor's thesis|]. Metropolia Ammattikorkeakoulu. http://www.theseus.fi/handle/10024/92773
Juhlin, O., Engström, A., & Önnevall, E. (2014). Long tail TV revisited: From ordinary camera phone use to pro-am video production. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Toronto, Ontario, Canada (ACM), 1325–1334. https://doi.org/10.1145/2556288.2557315
Juhlin, O., Engström, A., & Reponen, E. (2010). Mobile broadcasting: The whats and hows of live video as a social medium. *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* Lisbon, Portugal (ACM), 35–44.
Juhlin, O., Zoric, G., Engström, A., & Reponen, E. (2014). Video interaction: A research agenda. *Personal and Ubiquitous Computing*, 18(3), 685–692. http://dx.doi.org/10.1007/s00779-013-0705-8
Karis, D., Wildman, D., & Mané, A. (2016). Improving remote collaboration with video conferencing and video portals. *Human–Computer Interaction*, 31(1), 1–58. https://doi.org/10.1080/07370024.2014.921506
Koehne, B., Shih, P. C., & Olson, J. S. (2012). Remote and alone: Coping with being the remote member on the team. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* Seattle, Washington, USA (ACM), 1257–1266. https://doi.org/10.1145/2145204.2145393
Kostiukevych, I., Vermost, W., & Ferreira, P. (2019). Analyzing SMPTE ST 2110 streams using EBU's open-source software. *SMPTE Motion Imaging Journal*, 128(4), 1–6. https://doi.org/10.5594/JMI.2019.2899712

Ksentini, A., Naimi, M., & Gueroui, A. (2006). Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture. *IEEE Communications Magazine*, 44(1), 107–114. https://doi.org/10.1109/MCOM.2006.1580940

Kumar, A., & Haneef, M. S. M. (2018). Is Mojo (En)De-skilling? *Journalism Practice*, 12(10), 1292–1310. https://doi.org/10.1080/17512786.2017.1389291

Lapierre, J., & Al-Habbal, M. (2018). Bridging the gap between software and SMPTE ST 2110 SMPTE 2018 Annual Technical Conference & Exhibition Los Angeles, CA, USA. , 2018 (SMPTE), 1–7. https://doi.org/10.5594/M001841

Linden, C.-G. (2017). Decades of automation in the newsroom. *Digital Journalism*, 5(2), 123–140. https://doi.org/10.1080/21670811.2016.1160791

Mlynář, J., González-Martínez, E., & Lalanne, D. (2018). Situated organization of video-mediated interaction: A review of ethnomethodological and conversation analytic studies. *Interacting with Computers*, 30(2), 73–84. https://doi.org/10.1093/iwc/iwx019

Musabirov, I., Bulygin, D., Okopny, P., & Konstantinova, K. (2018). Event-driven spectators' communication in massive esports online chats. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* Montreal, QC, Canada (ACM), LBW564.

Napoli, P. M. (2014). On automation in media industries: Integrating algorithmic media production into media industries scholarship. *Media Industries Journal*, 1(1 33–38). https://doi.org/10.3998/mij.15031809.0001.107

Nimako, S. G., & Ntim, B. A. (2013). Construct specification and misspecification within the application of push-pull-mooring theory of switching behaviour. *Journal of Business and Management Sciences*, 1(5), 83–95 http://pubs.sciepub.com/jbms/1/5/2/.

Nygren, G. (2014). Multiskilling in the newsroom–de-skilling or re-skilling of journalistic work? *The Journal of Media Innovations*, 1(2), 75–96. https://doi.org/10.5617/jmi.v1i2.876

Obrist, M., Bernhaupt, R., & Tscheligi, M. (2008). Interactive TV for the home: An ethnographic study on users' requirements and experiences. *International Journal of Human–Computer Interaction*, 24(2), 174–196. https://doi.org/10.1080/10447310701821541

Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human–Computer Interaction*, 15(2–3), 139–178. https://doi.org/10.1207/S15327051HCI1523_4

Perry, M., Juhlin, O., Esbjörnsson, M., & Engström, A. (2009). Lean collaboration through video gestures: Co-ordinating the production of live televised sport. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Boston, MA, USA (ACM), 2279–2288.

Phillips, A., Singer, J. B., Vlad, T., & Becker, L. B. (2009). Implications of technological change for journalists' tasks and skills. *Journal of Media Business Studies*, 6(1), 61–85. https://doi.org/10.1080/16522354.2009.11073479

Retelny, D., Bernstein, M. S., & Valentine, M. A. (2017). No workflow can ever be enough: How crowdsourcing workflows constrain complex work Proceedings of the ACM on Human-Computer Interaction 1 CSCW . , 1–23. https://doi.org/10.1145/3134724

Richards, G. (2010). Here is the news [automation broadcasting]. *Engineering Technology*, 5(9), 36–38 1750-9637 .

Satchell, C., & Dourish, P. (2009). Beyond the user: Use and non-use in HCI. *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7* Melbourne, Australia (ACM), 9–16.

Schmidt, K., & Bannon, L. (1992). Taking CSCW seriously. *Computer Supported Cooperative Work (CSCW)*, 1(1), 7–40. https://doi.org/10.1007/BF00752449

Stimson, R. J., & McCrea, R. (2004). A push – pull framework for modelling the relocation of retirees to a retirement village: The Australian experience. *Environment and Planning A*, 36(8), 1451–1470. https://doi.org/10.1068/a36206

Suchman, L. (1993). Do categories have politics? *Computer Supported Cooperative Work (CSCW)*, 2(3), 177–190. https://doi.org/10.1007/BF00749015

Tang, J. C., Zhao, C., Cao, X., & Inkpen, K. (2011). Your time zone or mine? A study of globally time zone-shifted collaboration. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* Hangzhou, China (ACM), 235–244. https://doi.org/10.1145/1958824.1958860

Tharatipyakul, A., Li, J., & Cesar, P. (2020). Designing user interface for facilitating live editing in streaming. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* Honolulu, HI, USA (ACM), 1–8. https://doi.org/10.1145/3334480.3383037

Túñez-López, M., Vaz-Álvarez, M., & Fieiras-Ceide, C. (2020). Covid-19 and public service media: Impact of the pandemic on public television in Europe. *El Profesional de La Información* 29 5 , e290518. https://doi.org/10.3145/epi.2020.sep.18

Webb, A. M., Wang, C., Kerne, A., & Cesar, P. (2016). Distributed liveness: Understanding how new technologies transform performance experiences. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* San Francisco, California, USA (ACM), 432–437.

Weilenmann, A., Säljö, R., & Engström, A. (2014). Mobile video literacy: Negotiating the use of a new visual technology. *Personal and Ubiquitous Computing*, 18(3), 737–752. https://doi.org/10.1007/s00779-013-0703-x

Wohn, D. Y., & Freeman, G. (2020). Audience management practices of live streamers on twitch. *ACM International Conference on Interactive Media Experiences* Barcelona, Spain (ACM), 106–116. https://doi.org/10.1145/3391614.3393653

Yarosh, S., Bonsignore, E., McRoberts, S., & Peyton, T. (2016). YouthTube: Youth video authorship on YouTube and Vine. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* San Francisco, California, USA (ACM), 1423–1437. https://doi.org/10.1145/2818048.2819961

Chapter 8

Article III: Designing for Collaborative Video Editing

Okopnyi, Pavel, Oskar Juhlin, Frode Guribye

Nordic Human-Computer Interaction Conference (NordiCHI'22), 1-11 (2022)

Designing for Collaborative Video Editing

Pavel Okopnyi
Pavel.Okopnyi@uib.no
Department of information science and media studies, University of Bergen
Bergen, Norway

Oskar Juhlin
oskarj@dsv.su.se
Department of Computing and Systems Sciences, Stockholm University
Stockholm, Sweden
MediaFutures, Department of information science and media studies, University of Bergen
Bergen, Norway

Frode Guribye
Frode.Guribye@uib.no
MediaFutures, Department of information science and media studies, University of Bergen
Bergen, Norway

ABSTRACT

This paper explores the design space of collaborative video editing through a series of design workshops with video editors. Collaborative video editing can be supported by adding awareness features or other well-known collaborative features found in existing software and introducing new features designed specifically for video editing software. The paper identifies different design concepts that illustrate how such collaborative features can be included in non-linear video editing software and discusses the challenges of introducing such features. Some design concepts are explicitly inspired by existing collaborative tools. However, we suggest that introducing such features might not be straightforward. In other cases, alternative abstract representations of time-based media might be necessary to support collaborative video editing.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative interaction**; **Collaborative content creation**; **Graphical user interfaces**.

KEYWORDS

video editing, collaborative video work, video production, user interface design

ACM Reference Format:
Pavel Okopnyi, Oskar Juhlin, and Frode Guribye. 2022. Designing for Collaborative Video Editing. In *Nordic Human-Computer Interaction Conference (NordiCHI '22), October 8–12, 2022, Aarhus, Denmark*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3546155.3546664

1 INTRODUCTION

Video editing is a complex activity embedded in a social setting, including collaborators with different roles and skillsets, such as montage, colour correction, sound levelling and graphics production [3]. At the same time, the editing per se is most often done in solitude by a single person, and the supporting software is primarily designed for an individual user. Through a number of participatory design sessions, we investigate challenges and opportunities in enabling online collaboration in video editing.

Video editing software has a specific challenge that it tries to address -— how to represent the temporal dimension of media and allow its manipulation [28, 37]. Existing video editing software, such as Adobe Premiere and Apple Final Cut, utilises abstract representations and metaphors, such as a timeline and thumbnails [16, 17]. The timeline is commonly a horizontal representation of the temporal dimension, going from left to right. In the timeline, media elements, such as clips and images, are often represented as rectangles, with the length of a rectangle representing the duration of the media. The media elements are often supplemented by thumbnails demonstrating a frame from a clip [16]. This abstract representation presents limited information about the underlying video and mediates the editing of vital aspects of the video such as narrative, movement and sound. Thus video editors are required to repeatedly recall and review the video during the editing process to see and evaluate the results of their editing [16, 32, 34].

Given the emergence of new ways of working, such as remote work in TV and video production automation [15, 24, 31], as well as the availability of new technologies, there is a growing interest in the industry in the development of video workflows that would support online collaboration and concurrent video editing. Video production used to be exclusive to large organisations due to its high cost. New ways of producing video content have emerged as both high-speed communications networks and production tools have become cheaper and more readily available for distribution. It has become more affordable to conduct editing at multiple places due to the emergence of inexpensive digital video recorders and non-linear video editing software [22]. Furthermore, the introduction of social distancing rules during the COVID-19 pandemic forced TV and video professionals to adapt to the new ways of working and to tackle the idea of cloud-based video work [31]. In sum, there is a growing mismatch between the collaborative production setting and the individualised functionality of the editing software [34].

The topic of supporting collaborative video editing is further inspired by research on collaborative tools for other types of media production. Examples include research to make writing [1, 13], music creation [4, 5] and drawing [21, 40] collaborative. Further, research in the area of Computer-Supported Collaborative Work (CSCW) [38] has shown the importance of "workspace awareness" [8, 11, 20] in real-time distributed collaborative work. Gutwin and

NordiCHI '22, October 8–12, 2022, Aarhus, Denmark

ACM ISBN 978-1-4503-9699-8/22/10.
https://doi.org/10.1145/3546155.3546664

Greenberg [20] suggest that user interface (UI) design should provide users with information about the state of the shared workspace by answering questions in several categories, such as "Who is present?", "What are they doing?" and "Where are they working?" [20].

When collaborators are continuously cognisant of each other's ongoing work, it helps them coordinate and adapt to each other. The focus on concrete and ongoing practical details in work settings has led to the generation of new applications and services [23, 31, 36]. Additionally, CSCW has over the years also identified a number of challenges in designing for collaboration [7, 18, 19, 27].

The topic of collaborative video editing has already attracted interest in the area of human-computer interaction (HCI), and then in the part of video editing that is currently most concretely about collaboration, that is, the reviewing [32, 34]. We argue that there is a potential to expand existing research on collaborative video reviewing by providing a broader perspective on the possibilities for collaborative editing, including features that are anchored in the concept of workspace awareness. In this paper, we focus on how we can redesign existing non-linear video editing software to include collaborative features and support new collaborative workflows.

To address the question, we carried out this study in two steps. First, we invited twenty persons to take part in five design sessions to generate a set of design ideas on the topic of awareness support in video editing. The participants ranged from media students to professional editors. The sessions resulted in design concepts such as "Time Slot Separation" and "Scaffolding." Second, we analysed the discussions around the concepts and identified high-level themes such as "focus" and "resemblance" that put said concepts in the context of their potential uses.

In our analysis, we identify challenges of enabling collaboration in video editing, such as the need to account for potential distractions introduced by new elements of the user interface. The analysis shows how some challenges go beyond those found in other shared workspace applications (such as collaborative writing tools, e.g., Google Docs) because the video editors work with abstract representations of media (the timeline) and the complexity this adds to the workspace.

2 COLLABORATIVE VIDEO WORK: PREVIOUS RESEARCH AND EXISTING TOOLS

There is a strong interest in enabling collaborative video work in the industry illustrated by the emergence of multiple tools such as frame.io[1], ftrack[2] and ReviewStudio[3]. These services allow uploading or casting a video or other media to remote participants and gathering feedback through annotations and online video sessions. They also all aim to enable collaborative reviewing. These tools, however, do not support multiple video editors to work collaboratively in a shared workspace.

There are attempts to create collaborative video editing software. Service motionbox.io[4] provides a simplified web-based video editing tool that allows multiple participants to work on the same video project simultaneously. However, the service is still under development and does not provide any specific collaborative features that would help co-editors to communicate and coordinate their activities.

Some professional video editing suites, such as Avid Media Composer and DaVinci Resolve, allow multiple participants to work on the same project by locking certain parts of the project. This feature is known as "bin locking." It implies that only one user can manipulate elements of video, that is, clips, that are inside the bin – a virtual folder, while other users can access the bin in a read-only mode. DaVinci Resolve also includes a built-in chat to enable communication between video production team members[5].

In research, there have been multiple attempts to design, prototype and develop collaborative systems for video work when it comes to technological viability, means of interaction and reviewing practices.

As early as 2006, Fonseca and Carrapatoso [14] provided a paper that argued for the potential of collaborative video editing given new Internet technologies. The topic was motivated by achievements in the research area of CSCW. They argued that collaborative activities, such as communication, sharing and joint visualisations, could be of benefit also in video editing. Their study showed that the idea was technologically viable, but at the same time stated that: "The implementation of a complete cooperative video editing tool is a quite complex task and would require the involvement of a large multidisciplinary team."

The technical viability of collaborative video editing was further demonstrated by Klokmose et al. [25], who presented a web-based toolkit "Videostrates." The toolkit provided the basis for the development of collaborative video tools with personalised interfaces that are tailored towards specific tasks, and emphasised the need for further investigation of real-time collaboration around video.

The possibility of digital video editing sparked further design research on the means of interaction with video. Although the media is visual and aural, it was argued that editing could be collaborative and allow mutual awareness if making cuts and realigning the time sequences became more tactile. Taylor et al. [41] presented a tabletop editing system named VideoPlay, which presented video clips in the form of plastic tiles and allowed their combination into video sequences. The prototype aimed to make the video editing experience playful and more engaging and allowed several people to work on the video. Zigelbaum et al. [43] presented a prototype that aimed to combine the benefits of traditional linear and more modern non-linear video editing. The prototype presented video clips as plastic tokens, allowing participants to rearrange them in order to create a coherent video. Terrenghi et al. [42] designed an interactive tabletop application for collaborative video editing that aimed to enable video work as a spontaneous activity, allowing all interested parties to be involved in the editing process. Another tangible video editor prototype was presented by Merz et al. [29], who aimed to create a more engaging video editing experience for school children and promote digital storytelling literacy. Similarly to previous works, the prototype required participants to be physically present in the same space and was limited in its video editing functionality.

[1] https://frame.io/
[2] https://www.ftrack.com/
[3] https://www.reviewstudio.com/
[4] https://motionbox.io

[5] https://www.blackmagicdesign.com/no/products/davinciresolve/collaboration

The call for research on collaborative video applications [14, 30, 39] has generated a specific interest in reviewing practices. Pavel et al. [34] developed a video reviewing system that allowed reviewers to leave feedback for video editors. Their system allowed reviewers to record video feedback that would include non-verbal cues to the editors. Okopnyi et al. [32], in their study of communication between video editors and other stakeholders, noted that video work is often open-ended and requires video editors to devise various strategies to avoid or resolve conflicts and misunderstandings. They suggested three design directions for video reviewing and editing software: scaffolding, iconic referencing and suggestive editing.

Bartindale et al. [2, 3] developed and evaluated experimental tools such as TryFilm [2] and StoryCrate [3] to support collaboration during and immediately after the filming process. These tools allowed filming crews to review the filmed footage, reflect upon its quality and perform basic video editing such as cutting and rearranging clips. Similarly to other tangible and tabletop systems, these tools required the physical presence of participants.

In sum, the technical opportunities in supporting collaborative video editing have been mostly topicalised in research in two ways: by enabling playful, spontaneous and more engaging interaction in a co-present setting and by introducing sharing mechanisms in the reviewing process. The former type provides new forms of co-located interaction in a studio following traditional linear editing principles where cuts were made, and sections were rearranged and glued back. Reviewing has received some attention in both academia and industry. The focus and research on this practice is a starting point for design to support remote collaborative video editing. Still, there are other collaborative practices that can also be supported by video editing tools. In the following, we explore some of these.

3 METHODS

Our approach to expanding the design space for collaborative video editing includes a set of design workshops and inductive analysis of the generated material, that is, the design concepts and the transcripts of participants' discussions. We conducted two rounds of design workshops on the topic of supporting collaboration in a video production team. The first round focused on the topic of workspace awareness [20]. The second round had a broader scope. In all workshops, the participants were instructed that the ideation should focus on remote video editing, including both synchronous and asynchronous collaboration between video editors who use current non-linear editing software tools in a professional capacity and work in production teams.

3.1 The design workshops (WS1-2)

The first round was dedicated to ideation in two separate workshops (WS1-2), with eight participants in total. The initial workshop was conducted with three master's students who studied media and human-computer interaction. With them, we verified the workshop protocol. The evaluation included topics such as whether the task descriptions and other materials were clear enough, whether the time for completing the tasks was enough and whether the overall pacing of the workshop was good.

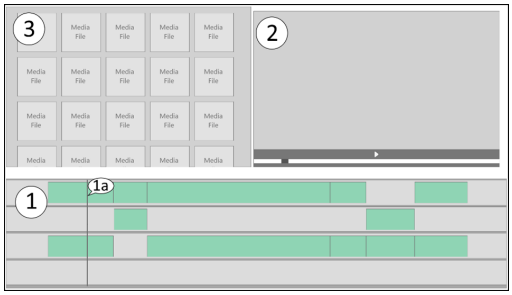

Figure 1: Schematic Video Editor UI

In the second workshop, we recruited five people via social media. All participants had experience in video editing ranging from participating in high school video projects and hobby projects to studying TV and video production at university and doing professional video work.

Both workshops in the first round (WS1-2) followed the same protocol. All participants were introduced to the workshop's topic and provided with general CSCW theory as inspiration, that is, the definition of workspace awareness as presented in [20], and examples of workspace awareness categories and features in existing online collaborative services, such as Google Docs, Overleaf and Balsamiq. We also presented a schematic illustration of the video editing software UI (Fig. 1). The schematic constitutes several elements that are commonly found in existing video editing software. First, there is a timeline (Fig. 1-1) that contains graphics, video and sound tracks. Inside the tracks, there are clips that represent pieces of media. On top of the timeline, there is a playhead (Fig. 1-1a) that represents the current position of the user in the timeline. Second, there is a preview window (Fig. 1-2) that allows watching the current version of the video. Last, there is a media catalogue (Fig. 1-3), a set of media files that can be ingested into the timeline, such as raw footage files, pre-recorded audio files, and various graphical elements such as images.

Thus, we suggested that participants should think of ways to augment or change the existing software user interface along with new workflows that they would like to introduce. Therefore, we tried to ground design ideas in existing software and refrain from inventing completely new user interfaces that would be unfamiliar to video workers.

After the presentations, we had a brainstorming session in which participants worked individually to suggest as many design ideas as possible. The ideas were written down on sticky notes. They were then developed into prototyping sketches on paper. Then, each participant presented their ideas and discussed them. The most promising ideas were selected and re-developed by the participants. Finally, they discussed the relevance and feasibility of the produced ideas.

The second workshop in the first round was video-recorded and transcribed.

3.2 The second round of workshops (WS3-5)

In the second round (WS3-5), we used design ideas generated during the first round (WS1-2) as an inspirational starting point. Due to the COVID-19 pandemic and social distancing restrictions, we conducted three online workshops attended by four participants each. The participants included:

- three current TV editors, directors, and producers;
- one TV automation specialist;
- one multimedia production specialist with a background in motion design and TV show production;
- four current freelance video editors and filmmakers with experience ranging from YouTube video production and documentaries to professional TV work; and
- three UX designers with experience in design for professional visual media production tools and video work ranging from professional video editing to running a YouTube channel.

The participants have relevant work experience in various European countries, including Bulgaria, Denmark, Norway, Serbia, Ukraine and the United Kingdom.

Each workshop comprised three stages. First, participants were introduced to the workshop's topic – what features an online collaborative editing tool is required or expected to have. Participants also introduced themselves and described their video work experience. We demonstrated the design ideas developed in the first round of workshops (WS1-2) and discussed the relevance of the presented ideas to the field and participants' work experience. Participants were encouraged to express their opinions and use these ideas to inspire the design. Second, we encouraged participants to develop the presented ideas further or propose new design concepts. We used the Balsamiq.cloud web application to work collaboratively on prototypes. We reflected on the ongoing online collaboration experience and discussed possible design interventions to enable concurrent work. Third, we concluded the workshop with a short discussion of the feasibility and relevance of the developed prototypes.

All workshops in the second round were video-recorded and transcribed.

3.3 Data Analysis

The workshops yielded two types of data that were further analysed: design concepts and the transcripts of participants' discussions. We searched the transcripts for the reasoning behind the design concepts. Specifically, for each design idea, we tried to answer three questions: How should it work? Why is it beneficial? What are the drawbacks or disadvantages? While we did not explicitly specify the setting of video production for the ideation process, the participants, with their broad and varied experience in editing, brought up many examples of relevant contexts and settings in their discussions.

We employed content analysis, a research technique that aims to make valid inferences from texts and other meaningful materials in the context of their use. With this approach, we were able to reduce collected data, that is, transcripts, by identifying core patterns and ideas [33]. Then, from the volume at which the subject matter was discussed, we inferred the importance that workshop participants assigned to identified patterns and ideas [26].

The data were coded inductively in three iterations. We started with the most apparent aspect of each design idea as workshop participants voiced it. Then we looked for various motivations and possible challenges for design ideas. In the third iteration, we identified overarching themes and patterns that describe such motivations and challenges. The analytical work laid the foundation for the coding scheme, which was refined through revisiting the material between coding sessions. In the end, we arrived at high-level categories that address various aspects of collaborative video editing and put design concepts in the context of their potential use.

4 RESULTS AND ANALYSIS

During the workshops, the participants produced multiple design ideas. Here, we present design concepts, which were developed into either pen-and-paper or digital sketches and then discussed by the participants. In order to convey the meaning of the sketches better to non-participants, the authors of this paper have developed the sketches further. We also present our categories (in italics) that express participants' descriptions of the advantages and potential challenges of the concepts.

Here, we first mention several design concepts that might be considered "generic" to collaborative tools in general. Then we describe in more detail design concepts which, in our opinion, require more attention from research and design communities.

4.1 Enabling Workspace Awareness

Some design concepts were to a large extent inspired by other collaborative tools, such as Google Docs, and the workspace awareness categories as described in [20]. Those ideas include using colour-coding on avatars (Fig. 2-1) and other interface elements, such as playheads (Fig. 2-2), emphasising salient actions performed by other users with visual effects, integrating chat, and adding an editing log. The participants suggested that the implementation of such concepts "[would be] on par with most real-time collaborative systems, like Google Docs." Thus, due to the *resemblance* to existing tools, these features should be recognisable and not require additional effort from users to learn the functionality.

Even though these features exist in various collaborative software, the participants still expressed multiple concerns regarding their presence in collaborative video editing software. Specifically, additional dynamic UI elements, such as multiple coloured playheads (see Fig. 2-2) or new message notifications, might become a source of *distraction*:

> "[I]f I'm editing, and I'm doing my work, and I suddenly see a lot of highlighted [elements] around me, it's really distracting."

At the same time, the participants recognised the potential of such additions to help work collaboratively by providing awareness of "who" is present, "where" they are and "what" they are doing in the shared workspace: "you would have a focus on what he's doing, so you can follow what's happening." For example, an integrated chat could be augmented with support for indexical refencing –

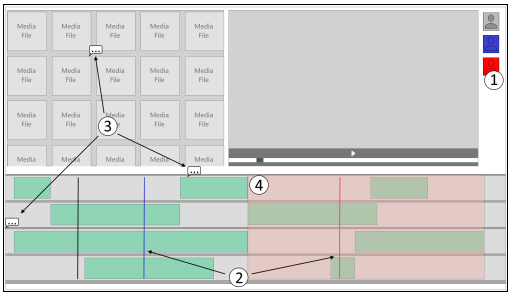

Figure 2: Color-coded avatars and 2. playheads. 3. Annotations, and 4. Time Slot Separation

timecodes entered in the chat should be transformed into references that take users to the referred-to time frame when clicked on.

Previous research on reviewing [32, 34] shows that indexical referencing is essential to communication around video. In a co-located setting, pointing is one of the most common activities in which co-editors engage [42]. Thus, the ability to refer to certain points of video via timecodes and address certain people in discussions is crucial for video workers and should be supported by online collaborative video editing software.

4.2 Annotations

The idea suggests adding textual notes to various elements of the video project, such as media clips, images, and different elements of the timeline (Fig. 2-3). Notes can be addressed towards a person or a group of people and are intended to draw collaborators' attention to areas of demand for further work:

> "I'm a video editor, and I'm making notes, for example, to a sound designer: Here, [the] noise is a little too harsh, [can] you clean this a little bit?"

Annotations in time-based media have been topicalised in academic research [32, 34] and industrial applications such as frame.io. Our participants, who were inspired by tools such as Google Docs, wanted to expand these ideas and employ annotations as a task-tracking tool, as well as means of communication between collaborators. For example, in contrast to frame.io, our participants suggested that annotations should not be attached to static points in the timeline (timecodes) but to clips and other elements that editors can move.

There could also be an overview of all notes addressed to the editor in order to help track the progress and to navigate problematic areas of the project, without the need to go through the timeline:

> "You also need one place where all notes [are listed] so you don't really miss any of those. You give the notes [to] a sound designer, he can open all notes in one list, so he doesn't need to go through [the] timeline, so he's sure he does not miss any."

Again, the participants argued that the idea was motivated by its *resemblance* to existing software, specifically the commentary feature in Google Docs, which allows using notes to track progress, communicate vital information to collaborators and assign tasks to each other:

> "Like [in] Google Docs, [...] you make notes for someone, and when this problem is solved and [marked as] solved, [the note] disappears."

Further, it allows switching to a new *digital workflow* instead of "taking notes manually," as a participant described their collaboration:

> "I used to take notes manually, so [the producer] would watch [the video], stop it, I will write timecode or sometimes mark it on the timeline, but still I will have to write the comment in my notebook."

However, another participant expressed concern that the implementation of this design concept might be obstructive to the editing, as it will take up *limited space of the user interface* and present too much information to the editor:

> "[There is] a thing with taking up the space of the interface [...] you would always have to be aware of what's the most important to look at."

4.3 Spectating Mode

A spectating mode is a special mode added to the collaborative editing software that allows other editors to look at others' work while working on other parts. Video work in a co-located setting has many advantages, as was demonstrated by previous research [41–43]. The participants drew on their experiences of working in a face-to-face setting and suggested this design concept that allows remote editors to look at what their collaborators are currently doing. The mode works as follows. The user clicks on the collaborator's avatar in the UI (Fig. 3-left), and the software goes into spectating mode. It then presents the shared workspace from the point of view of the selected collaborator, allowing real-time observation but not interaction with it (Fig. 3-right).

As a participant suggested, this concept would allow a new *digital workflow* for participants who otherwise had to be physically near to the editor in order to collaborate:

> "You can discuss what you're doing. If the director or someone is not in the same place, instead of coming into the editing room, you can just do it like this."

It would also allow *saving time* in situations when the editor needs to confirm the edits with other project members:

> "I've lost [a lot of] time because when I had to confirm with my clients or my team project, I [had] to render [the video] and send [it to] them."

However, one participant expressed concern regarding *privacy of work* in this kind of setup:

> "What I would prefer [...] if I'm editing and someone wants to look at what I'm doing, they [would] have to be granted access to my editing for viewing, so they will not be able to monitor what I'm doing without me knowing."

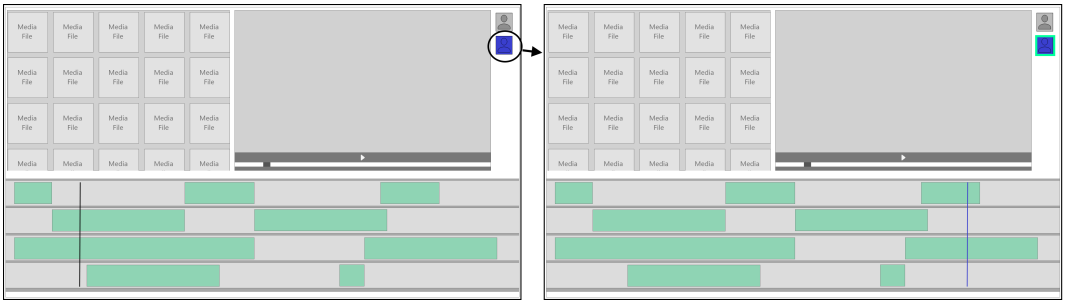

Figure 3: Spectating mode: video editor UI (left) and spectating another editor (right)

4.4 Integrated Reviewing Tool

This idea is inspired by existing software, such as frame.io, which allows third parties to leave comments on the video attached to certain timestamps and have a discussion around the video as a whole.

The integrated reviewing tool allows immediate communication between the editor and a collaborator who is not directly involved in the editing. For example, an editor can demonstrate specific fragments of the video to a producer without needing them to watch the whole video. The reviewing tool also allows reviewers to leave notes with their feedback attached to the timeline via a web-based interface. The notes are immediately available for the editor in the video editing software (Fig. 4).

Existing software, such as frame.io, lacks integration with editing software, and it requires video editors to perform *additional work* to get feedback:

> "Frame.io is very convenient. I just would like to skip the step of rendering and sending [the video]. [It would be] a huge help [if] they could see the [video in] real time in good quality because either you send it to them in bad quality or they have to wait [for] a few hours."

The concept was positively regarded by the participants. They suggested that having the ability to initialise a feedback session instantly, without the need to render the video file and upload it to a third-party service, is beneficial to the workflow.

Reviewing and quality of communication between parties are essential in commercial video production. However, in previous research [34] and industrial applications such as frame.io, the reviewing process is treated as an asynchronous activity that happens in iterations. Our participants suggested that synchronous real-time collaboration between editors and reviewers would be a "huge help" to the video production process. Similar to the spectating mode (4.3), this design idea allows *saving time* by eliminating certain steps of the reviewing workflow:

> "It would be really interesting [if] I can just call to my client [and say]: Hey, can you look at the 1 minute 32 seconds? Do you like this? What do you think about it?' without rendering the video. This is probably the first problem which I have as an editor at the project."

However, again one participant suggested that such a tool might be *distracting* to the editors if the feedback is not limited. They suggested that the editor should be in control of the situation: "[Clients] have to be let in only when they can be let in, when you're ready." Another participant mentioned that they would prefer not to *introduce new tools that other parties are not familiar with* into the workflows and communicate with other parties via more commonly used software: "If a client contacted me on Skype, we are going to have the call on Skype."

4.5 Time Slot Separation

This design concept is based on the idea of dividing a project into several areas that can be distributed between co-editors to ensure concurrent work. An editor, who works on a segment of the video, defines a slot, that is, a section of the timeline, by selecting video clips in the timeline and marks this section as something they are working on. Other collaborators are not allowed to make alterations, but they can see what is being worked on by their colleagues (Fig.2–4). When the acting editor has completed the task, they mark the section as ready for the next step of post-production.

The concept suggests separating collaborators into various time segments of the shared workspace. Our participants discussed various ways of segmentation. They concluded that the separation should be based on media elements, that is, clips. However, this approach still needs to be evaluated and compared to other possible methods of separation, for example, based on concrete timecodes.

The participants argued that the benefits of such a concept were the way it allows editors to *focus* on their work and not worry about the integrity of their editing:

> "[There is] a section that I want to focus on because I need to do some modifications within it, and I don't want other people to do anything in this area right now."

Further, the possibility to restrict access provides the editor's *privacy of work*:

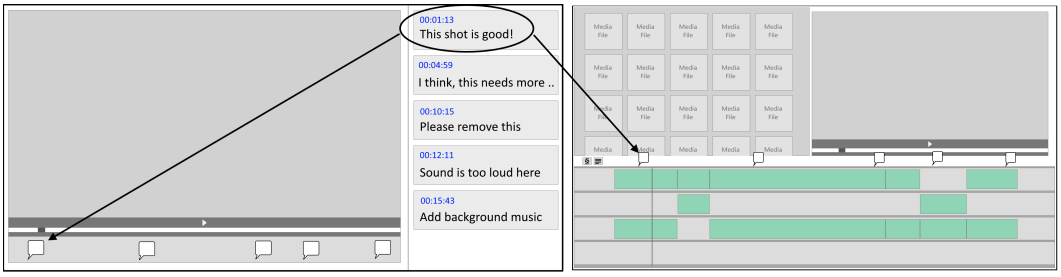

Figure 4: Reviewing tool: reviewer UI (left) and video editor UI (right)

> "Only I [should] see what I'm doing. [So] I could concentrate better. Because sometimes I try something [that] I don't like."

On the other hand, a participant suggested that this feature might be excessive and bring in more *complexity* to an already challenging process of video editing: "I think it adds a level of complexity [and video] editing is already complex."

4.6 Master Styles and Functional Separation

Most current collaboration occurs between people with different roles, such as the interaction between an editor and the graphics producer. In order to improve on such work, the participants suggested a feature that allows decoupling of the content of graphics, such as text and style. The editor who works on montage and is responsible for the content of the video can insert a graphics placeholder and put the necessary text into a free-form input field while the graphics style is incomplete (Fig.5—left). When the graphics producer updates the style, all graphical elements in the project are automatically updated (Fig.5—right). The concept extends to organised, functional separation. It enables concurrent work in between multiple collaborators instead of the way such interaction is now sequentially pursued.

Again, the concept draws on its *resemblance with existing software*, specifically Adobe Photoshop and Adobe XD:

> "The first [graphical element] is like a master component [in Adobe XD], and if you change the style of it all [graphical elements] are going to change as well."

It is suggested to be beneficial by *saving time* when editing long videos, such as feature films:

> "When you have a project [that is] one hour long, and then you decide [that] you want to tweak up this template, then it will change like all, like 50 or 100 elements in the whole timeline. [This] could take you two hours to change by hand."

4.7 Scaffolding

Workshop participants emphasised that in collaborative work, there is a need to have a general organisational structure for the video. This structure is often provided by a director, a writer, or a project manager, who does not participate in the actual editing. Similar to the design ideas presented in [2, 3, 32], they suggested attaching documentation describing such structures in the editing software. It would be represented as another layer of meta-information that is generated during the filming process:

> "When [...] working with the director [...] there should be some kind of a file attached to [the project] where they write guidelines for general editing."

In order to make such notes easier, the participants suggested a mobile app that would be synchronised with the editing software. Specifically, it would allow selecting a fragment of video, highlighting it with colour, and adding a voice note during production and recording of videos (Fig.6—left). When the footage is ingested into the video editing application, the editor can identify where on the timeline there are additional comments (Fig.6—right).

The idea resembles the current usage of clapper boards in filmmaking [9]: it is a device that helps designate and mark various video segments. It provides an opportunity to add metadata, such as scene titles, in the form of notes or plain text messages. The concept of scaffolding suggests enabling similar digital notes that would be associated with certain timeframes and can be presented in the timeline. This enables a *digital workflow* that otherwise unfolds with analogue artifacts, such as hand-written notes or via third-party software:

> "Some clients have a certain idea of what they want to use and what not to use before the editing starts. Instead of giving me just plain text messages, they could just access the software and do it in there."

The concept also suggests a *holistic approach* to UI design, keeping all necessary tools in one software suite and preventing users from having to "go between two screens all the time."

At the same time, a participant suggested that this solution might not be welcomed by external people, such as clients, as it suggests introducing *new tools that they are not familiar with*:

> "Some clients might be a little bit apprehensive about going into a software and doing anything in there; they would just prefer to [write a] message."

5 DISCUSSION

Presented design ideas help illustrate and make concrete the scope of supporting collaborative video editing. As mentioned in 4.1 and

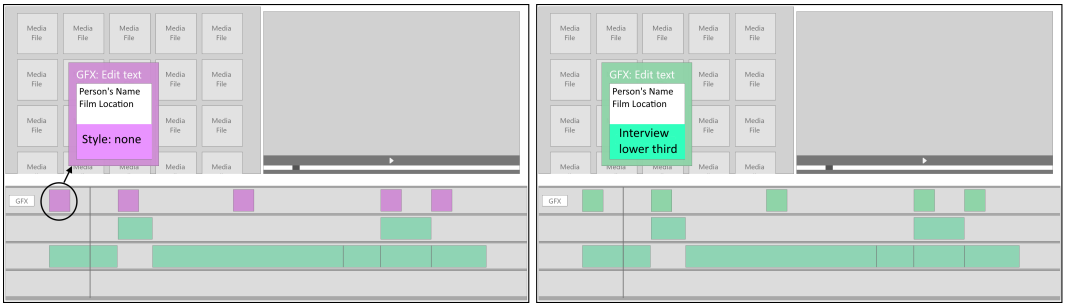

Figure 5: Master Styles: Editor enters text for graphics into a free form without template (left). Graphics template is finished and applied to graphical elements (right).

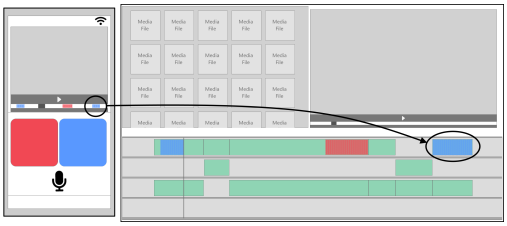

Figure 6: Mobile application (left) and areas of the footage highlighted in the video editing software UI (right)

4.2, some of the design concepts, such as colour-coded avatars and annotations, are explicitly inspired by existing collaborative tools (e.g., Google Docs and Frame.io) and the workspace awareness framework [20]. Still, these concepts have to be qualified for collaborative video editing software. As the participants suggested, it might be challenging to implement such concepts in a working environment successfully. Other concepts are not explicitly inspired by existing tools, though we can find similar design ideas in other software. For example, Time Slot Separation (4.5) can be considered similar to file-based separation in text-editing and software development tools [10, 27] as well as to bin-locking feature in DaVinci Resolve and Avid Media Composer. However, the lack of mass adoption of such concepts in the video industry suggests that they are still to be evaluated in a working environment. In our analysis, we identify three aspects that present significant challenges to designers and developers of collaborative video editing tools: focus and distraction, new workflows and tools, and working with the temporal dimension.

5.1 Focus and Distraction

Previous work on collaborative video editing showed that it is possible to enable collaboration when the editing happens in real time and in a face-to-face environment: the communication flows smoothly, and collaborators can see what others are doing at any moment [34]. In this study, the participants mention concerns about suggested features as being distractive (4.1), complex (4.5) or articulating the need to be able to focus (see 4.1). Since such topics are raised when discussing various applications, we suggest that it is a general and serious concern. Indeed, distributed collaboration demands additional effort compared to a face-to-face environment: communication between group members is limited, maintaining group focus is more difficult, and the work itself can be confusing and chaotic [7, 12].

One of the presented design concepts (4.5: Time Slot Separation) suggests that the introduction of additional structural mechanisms might be able to help deal with these issues. Specifically, this concept envisions a feature that prevents access by other editors and, essentially, limits possible collaboration: other editors cannot edit the same part of the video. Thus, the editor who is responsible for that part can work without distraction.

5.2 New Workflows and Tools

In a number of design concepts (see 4.2, 4.3 and 4.7), our participants suggested providing support for various activities that usually occur beyond the scope of video editing software. For example, the reviewing process typically involves rendering the video and sharing it with other parties via some file-sharing service. Previous studies suggest improving the reviewers' experience by providing them with additional instruments to support annotating and indexical referencing [6, 32, 34]. However, the introduction of these new instruments does not solve the underlying problem, that is, the time and effort required to share a video with another person.

One design concept (4.3: Spectating Mode) essentially proposes adding to video editing software functionality that is already implemented in remote desktop services and screen-sharing features of communications applications. Our participants suggested that a mere integration of such functionality would be beneficial for the overall video editing process. Specifically, video editors would be able to get feedback from their colleagues promptly and discuss the edits while in the process of editing. This change in the mode of reviewing from iterative and asynchronous to immediate and synchronous might have a significant impact on the whole editing process. However, it is hard to tell what this impact would be, as previous research has indicated that video editors sometimes

limit communication with other parties to ensure the quality of the feedback [32].

Besides video itself, video production often involves various media, such as textual notes and documentation. Two design concepts (4.2 and 4.7) specifically suggest integrating these types of media into video editing software to save time and reduce the effort that is spent on switching [7] between different applications and contexts when editors work with documentation, messages, and notes. However, integrating new tools and inviting non-editors to participate in the editing process might not be welcome, as our participants suggested that the introduction of unfamiliar tools can confuse less experienced users. Thus, in line with [25], we suggest that a more personalised approach to user interface design could be preferable.

5.3 Working with the Temporal Dimension

The interface of a video editing software application visually displays the time-based media in a time-independent form, that is, a line. In our design exercises, this time conversion is also a given when suggesting new collaborative features. All features suggested by the participants are ideas on how to support collaboration in a setting where the time feature in a video is described as a line, that is, a so-called timeline [16].

This preconception might have influenced the participants to draw inspiration from existing collaborative production tools of media, which are less time-based than video. Concrete examples refer to collaborative features in text writing. Such associations and claims on resemblance are mentioned in several of the suggested applications (see 4.1, 4.2 and 4.6).

However, it might be that the "resemblance" between line-based time representation in video software and text-based collaborative editors would not work as straightforwardly as first imagined. For example, indexical referencing (see 4.1 and 4.2) supports coordination and collaboration in a shared workspace [8, 20] in settings such as collaborative text writing [27]. However, we suggest that, in the case of video, there is a potential mismatch between a reference that exists in the user interface and a referred-to object that can exist in the medium itself. In terms of semiotic theory [35], the sign and the object are on different levels of abstraction. This example highlights the mismatch between what is happening with the medium during rearrangements and what is visible in the interface. Pavel et al. [34] address this mismatch in their design, allowing users to video-record their feedback and gesture "over the source video with the mouse" during the recording, essentially creating a new "feedback session" timeline.

In the editing process, the editors need to comprehend the connection between the abstract representation of video and the video itself [16, 17]. We suggest that this need might restrict the introduction of new collaborative elements similar to those that exist in other collaborative systems. If that is the case, the concerns raised might be fundamental and challenging. This might be one of the reasons why video editing software is so far designed mainly for individual users.

Previous research demonstrated alternative approaches to present time-based media, such as blocks [29], tokens [43], and storyboards [16]. These approaches include putting a structure on top of continuous media, breaking it into pieces. Specifically, each block, token, or storyboard frame represents a clip – a discrete segment of video; editors rearrange these clips to create a new video. One design concept (4.5: Time Slot Separation) suggests a similar approach – to impose a structure on top of the existing timeline by splitting the video into a series of time slots. Thus, the whole video can be represented as a collection of time slots, similar to a storyboard, which requires less effort to observe and comprehend [16]. We suggest that storyboard-based representation may be a viable alternative or a supplement to a timeline-based design which is prevalent in current video editing software. The idea of using storyboards for video editing in itself is not particularly new [28], however, it is still to be explored in the context of online collaboration.

The fact that collaborative skills and experiences, along with new types of software, are introduced to edit other types of media is inspiring but not decisive for the future of video editing. The intent of this study was to explore the design space of collaborative video editing. However, an important next step is to evaluate such concepts in practice and possibly shrink the space.

6 LIMITATIONS

This study focuses on a very specific domain, remote collaborative video editing. Thus, the suggested design concepts might be not applicable outside this domain.

In the workshops, we focused and limited the ideation process by introducing the topic of workspace awareness, existing collaborative tools, and the timeline-based video editing software interface. Thus, design ideas are anchored in existing video editing software, which is designed with a single user in mind.

The results of this study should be considered preliminary as we did not evaluate the suggested design concepts.

7 CONCLUSION

In this paper, we have examined how we can support collaborative video editing. In a series of design workshops, we identified, illustrated and discussed design ideas to support collaboration. We have identified key topics that put design concepts in the context of their potential use and discussed accompanying challenges.

Our results suggest that adding collaborative features might introduce unwanted distractions into video work, and we conceptualised the ways of mitigating such effects through providing additional structuring. We have looked beyond the scope of current video editing software and suggest providing support for collaboration and interactions, which often happen outside of such video editing tools. Further research on collaborative video editing should investigate alternative ways of video representation in editing software and focus on the evaluation of collaborative tools in a practical context.

ACKNOWLEDGMENTS

The study is a part of the project "Better Video workflows via Real-Time Collaboration and AI-Techniques in TV and New Media," funded by the Research Council of Norway under Grant No.: 269790. This work was also partly supported by industry partners and the Research Council of Norway with funding to "MediaFutures: Research Centre for Responsible Media Technology and Innovation," through The Centres for Research-based Innovation scheme, project

No.: 309339. We would like to express our gratitude to Taume Dery, Kari Raudstein, Sara Stegane and Oda Nordberg for their help with the workshops, and Vizrt and Vizrt's Even Normann and Roger Saetereng for their participation in the project. We would also like to thank our participants for their input.

REFERENCES

[1] Ronald M. Baecker, Dimitrios Nastos, Ilona R. Posner, and Kelly L. Mawby. 1995. The user-centred iterative design of collaborative writing software. In *Readings in Human–Computer Interaction*. Elsevier, 775–782.

[2] Tom Bartindale, Guy Schofield, Clara Crivellaro, and Peter Wright. 2016. TryFilm: Situated Support for Interactive Media Productions. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1412–1422. https://doi.org/10.1145/2818048.2819929 event-place: San Francisco, California, USA.

[3] Tom Bartindale, Alia Sheikh, Nick Taylor, Peter Wright, and Patrick Olivier. 2012. StoryCrate: Tabletop Storyboarding for Live Film Production. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 169–178. https://doi.org/10.1145/2207676.2207700 event-place: Austin, Texas, USA.

[4] Michele Biasutti. 2015. Assessing a collaborative online environment for music composition. *Journal of Educational Technology & Society* 18, 3 (2015), 49.

[5] J. Bligh, K. Jennings, and B. Tangney. 2005. *Designing interfaces for collaborative music composition*. Technical Report. Centre for.

[6] Soon Hau Chua, Toni-Jan Keith Palma Monserrat, Dongwook Yoon, Juho Kim, and Shengdong Zhao. 2017. Korero: Facilitating Complex Referencing of Visual Materials in Asynchronous Discussion Interface. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW (Dec. 2017), 34:1–34:19. https://doi.org/10.1145/3134669

[7] Andy Cockburn and Steve Jones. 1995. Four principles of groupware design. *Interacting with Computers* 7, 2 (1995), 195–210.

[8] Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. ACM, 107–114.

[9] Emilie Dumont and Bernard Mérialdo. 2008. Sequence alignment for redundancy removal in video rushes summarization. In *Proceedings of the 2nd ACM TRECVid Video Summarization Workshop (TVS '08)*. Association for Computing Machinery, New York, NY, USA, 55–59. https://doi.org/10.1145/1463563.1463572

[10] Lisa S. Ede and Andrea A. Lunsford. 1990. *Singular texts/plural authors: Perspectives on collaborative writing*. SIU Press.

[11] Sandy El Helou, Manolis Tzagarakis, Denis Gillet, Nikos Karacapilidis, and C. Yu Man. 2008. Participatory design for awareness features: Enhancing interaction in communities of practice. In *Proceedings of the 6th International Conference on Networked Learning*. Issue: CONF.

[12] Clarence A. Ellis, Simon J. Gibbs, and Gail Rein. 1991. Groupware: some issues and experiences. *Commun. ACM* 34, 1 (Jan. 1991), 39–58. https://doi.org/10.1145/99977.99987

[13] Robert S. Fish, Robert E. Kraut, and Mary DP Leland. 1988. Quilt: a collaborative tool for cooperative writing. In *ACM SIGOIS Bulletin*, Vol. 9. ACM, 30–37.

[14] Benjamim Fonseca and Eurico Carrapatoso. 2006. Coview: A Cooperative Architecture for Digital Video Editing. *SMPTE Motion Imaging Journal* 115, 11-12 (Nov. 2006), 482–491. https://doi.org/10.5594/J16132 Conference Name: SMPTE Motion Imaging Journal.

[15] Andreas Girgensohn, John Boreczky, Patrick Chiu, John Doherty, Jonathan Foote, Gene Golovchinsky, Shingo Uchihashi, and Lynn Wilcox. 2000. A Semi-automatic Approach to Home Video Editing. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology (UIST '00)*. ACM, New York, NY, USA, 81–89. https://doi.org/10.1145/354401.354415

[16] Dan B Goldman, Brian Curless, David Salesin, and Steven M. Seitz. 2006. Schematic Storyboarding for Video Visualization and Editing. In *ACM SIGGRAPH 2006 Papers (SIGGRAPH '06)*. ACM, New York, NY, USA, 862–871. https://doi.org/10.1145/1179352.1141967

[17] Dan B. Goldman, Chris Gonterman, Brian Curless, David Salesin, and Steven M. Seitz. 2008. Video object annotation, navigation, and composition. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*. 3–12.

[18] Jonathan Grudin. 1988. Why CSCW applications fail: problems in the design and evaluationof organizational interfaces. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*. 85–93.

[19] Jonathan Grudin. 1994. Groupware and social dynamics: Eight challenges for developers. *Commun. ACM* 37, 1 (1994), 92–105. Publisher: ACM New York, NY, USA.

[20] Carl Gutwin and Saul Greenberg. 2002. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)* 11, 3-4 (Sept. 2002), 411–446. https://doi.org/10.1023/A:1021271517844

[21] Claudia-Lavinia Ignat and Moira C. Norrie. 2006. Draw-together: Graphical Editor for Collaborative Drawing. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 269–278. https://doi.org/10.1145/1180875.1180917

[22] Oskar Juhlin, Goranka Zoric, Arvid Engström, and Erika Reponen. 2014. Video interaction: a research agenda. *Personal and Ubiquitous Computing* 18, 3 (March 2014), 685–692. https://doi.org/10.1007/s00779-013-0705-8

[23] Demetrios Karis, Daniel Wildman, and Amir Mané. 2016. Improving Remote Collaboration With Video Conferencing and Video Portals. *Human–Computer Interaction* 31, 1 (Jan. 2016), 1–58. https://doi.org/10.1080/07370024.2014.921506 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370024.2014.921506.

[24] Seungwon Kim, Gun A. Lee, Sangtae Ha, Nobuchika Sakata, and Mark Billinghurst. 2015. Automatically Freezing Live Video for Annotation During Remote Collaboration. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1669–1674. https://doi.org/10.1145/2702613.2732838 event-place: Seoul, Republic of Korea.

[25] Clemens N. Klokmose, Christian Remy, Janus Bager Kristensen, Rolf Bagge, Michel Beaudouin-Lafon, and Wendy Mackay. 2019. Videostrates: Collaborative, Distributed and Programmable Video Manipulation. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 233–247. https://doi.org/10.1145/3332165.3347912

[26] Klaus Krippendorff. 2018. *Content analysis: An introduction to its methodology*. Sage publications.

[27] Ida Larsen-Ledet and Henrik Korsgaard. 2019. Territorial Functioning in Collaborative Writing. *Computer Supported Cooperative Work (CSCW)* 28, 3 (June 2019), 391–433. https://doi.org/10.1007/s10606-019-09359-8

[28] Wendy Mackay and Daniele Pagani. 1994. Video Mosaic: Laying out time in a physical space. In *Proceedings of the second ACM international conference on Multimedia*. 165–172.

[29] Allison Merz, Annie Hu, and Tracey Lin. 2018. ClipWorks: A Tangible Interface for Collaborative Video Editing. In *Proceedings of the 17th ACM Conference on Interaction Design and Children (IDC '18)*. ACM, New York, NY, USA, 497–500. https://doi.org/10.1145/3202185.3210758

[30] Boris Novikov and Oleg Proskurnin. 2003. Towards Collaborative Video Authoring. In *Advances in Databases and Information Systems (Lecture Notes in Computer Science)*, Leonid Kalinichenko, Rainer Manthey, Bernhard Thalheim, and Uwe Wloka (Eds.). Springer Berlin Heidelberg, 370–384.

[31] Pavel Okopnyi, Frode Guribye, Valentina Caruso, and Oskar Juhlin. 2021. Automation and redistribution of work: the impact of social distancing on live TV production. *Human–Computer Interaction* 0, 0 (Nov. 2021), 1–23. https://doi.org/10.1080/07370024.2021.1984917 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370024.2021.1984917.

[32] Pavel Okopnyi, Oskar Juhlin, and Frode Guribye. 2020. Unpacking Editorial Agreements in Collaborative Video Production. In *ACM International Conference on Interactive Media Experiences (IMX '20)*. Association for Computing Machinery, New York, NY, USA, 117–126. https://doi.org/10.1145/3391614.3393652

[33] Michael Quinn Patton. 2014. *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.

[34] Amy Pavel, Dan B. Goldman, Björn Hartmann, and Maneesh Agrawala. 2016. VidCrit: Video-based Asynchronous Video Review. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 517–528. https://doi.org/10.1145/2984511.2984552 event-place: Tokyo, Japan.

[35] Charles Sanders Peirce and Justus Buchler. 1902. Logic as semiotic: The theory of signs.

[36] Mark Roseman and Saul Greenberg. 1996. Building real-time groupware with GroupKit, a groupware toolkit. *ACM Transactions on Computer-Human Interaction* 3, 1 (March 1996), 66–106. https://doi.org/10.1145/226159.226162

[37] Stephanie Santosa, Fanny Chevalier, Ravin Balakrishnan, and Karan Singh. 2013. Direct space-time trajectory control for visual media editing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1149–1158.

[38] Kjeld Schmidt and Liam Bannon. 1992. Taking CSCW seriously. *Computer Supported Cooperative Work (CSCW)* 1, 1 (March 1992), 7–40. https://doi.org/10.1007/BF00752449

[39] Naimdjon Takhirov and Fabien Duchateau. 2011. A cloud-based and social authoring tool for video. In *Proceedings of the 11th ACM symposium on Document engineering (DocEng '11)*. Association for Computing Machinery, New York, NY, USA, 65–68. https://doi.org/10.1145/2034691.2034705

[40] John C. Tang and Scott L. Minneman. 1991. VideoDraw: a video interface for collaborative drawing. *ACM Transactions on Information Systems (TOIS)* 9, 2 (1991), 170–184.

[41] Stuart Taylor, Shahram Izadi, Kursat Ozenc, and Richard Harper. 2007. VideoPlay: Playful and Social Editing of Video using Tangible Objects and Multi-touch Interaction. https://www.microsoft.com/en-us/research/publication/videoplay-playful-and-social-editing-of-video-using-tangible-objects-and-multi-touch-interaction/

[42] Lucia Terrenghi, Torsten Fritsche, and Andreas Butz. 2008. Designing environments for collaborative video editing. (2008).
[43] Jamie Zigelbaum, Michael S. Horn, Orit Shaer, and Robert JK Jacob. 2007. The tangible video editor: collaborative video editing with active tokens. In *Proceedings of the 1st international conference on Tangible and embedded interaction*. ACM, 43–46.

Errata for
"Collaborative Video Editing"

Pavel Okopnyi

Thesis for the degree philosophiae doctor (PhD)

at the University of Bergen

04.01.2024 _____ _____

(date and sign. of candidate) (date and sign. of faculty)

Errata

Page 45 Misspelling: “Redistribution or Work and Responsibilities” – corrected to “Redistribution of Work and Responsibilities“

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230864432 (print)
9788230844083 (PDF)