The Never-Ending Digital Dilemma

A Thesis

Presented to

The Faculty of the Moving Image Archiving and Preservation Program

of the Department of Cinema Studies

New York University

In Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

By

Soojin Park

May, 2020

Table of Contents

Abstract

It has been 13 years since the Academy of Motion Picture Arts and Sciences released its first version of Digital Dilemma in 2007, and eight years since the second edition, which was geared toward independent filmmakers, documentarians, and non-profit audiovisual archives, was released in 2012. Among the conclusions of the Digital Dilemma series was that digital data is much harder to preserve than celluloid film – despite the advantages that digital technologies bring to creating and sharing movies. While it is true that digital files and storages have significantly crept into how we create and handle movies since then, especially since 2012 which was the year when at least half of top hundred US-grossing movies were shot digitally for the first time, it is equally true that technology has been growing at a pace one could hardly keep up with. This interview-based paper does not serve as a quantitative analysis with figures and statistics. Rather, it is a reflection on the circumstances of larger institutions such as studios and smaller archives/independent filmmakers that make their archival practices different, or on the contrary, similar. The thesis first broadly discusses the shift in our moviegoing habit, which will be followed by a closer look at the status of archiving and preservation of feature-length movies in the movie industry, paying particular attentions to the Interoperable Master Format (IMF) and cloud technology.

Acknowledgements

Thank you to my thesis advisor, Nicole Martin, Associate Director of the Moving Image Archiving & Preservation program at New York University. She has been an invaluable source of knowledge and encouragement throughout the research process. I am also greatly indebted to the faculty and students of the Moving Image Archiving and Preservation program at New York University for their support.

I would like to thank James Snyder, Senior Systems Administrator at the National Audio-Visual Conservation Center, at the Library of Congress. His extensive knowledge across the history and technology of audiovisual preservation was integral to establishing the questions explored here.

The completion of this thesis could not have been possible without the expertise of Buzz Hays, Global Lead of Entertainment Industry Solutions for Google Cloud. He helped inspire many of the ideas expressed in this thesis. His unwavering enthusiasm for movies and the technology behind it kept me engaged and motivated with my research.

I would also like to express my gratitude to the members of the Inventory and Archive Department at Paramount Pictures, particularly Senior Manager Kathryn Claypool. Thanks to her generousness in inviting me to the Paramount Studio Lot in Los Angeles, I was able to broaden my horizons in the understanding of the archiving strategies implemented in a studio as prestigious as Paramount.

Finally, I would like to extend my appreciation to my family who has been there to support me throughout my studies.

Introduction

In the digital age, the sheer volume of movies being created in higher and higher resolutions and the never-ending problem of obsolescence across hardware and software are posing a significant set of problems in terms of storage and searchability. Institutions across the field including archives, standardization bodies, and major storage service providers are finding ways to deal with these problems. Despite the effort, there seems to be no concrete answers that will magically solve the digital dilemma. It has been 13 years since the Academy of Motion Picture Arts and Sciences released its first version of Digital Dilemma in 2007, and eight years since the release of the second edition, which was geared toward independent filmmakers, documentarians, and non-profit audiovisual archives. Among the conclusions of the Digital Dilemma series was that digital data is much harder to preserve than celluloid film – despite the advantages that digital technologies bring to creating and sharing movies. While it is true that digital files and storages have significantly crept into how we create and handle movies since then, especially since 2012 which was the year when at least half of top hundred US-grossing movies were shot digitally for the first time, it is equally true that technology has been growing at a pace one could hardly keep up with.[1] The volume of data and the intricacies of how we need to handle them is neither getting any smaller nor simpler. The level of technical knowledge that is expected from audiovisual archivists has, if anything, soared. Movies directed or produced by acclaimed filmmakers, especially when produced or distributed by big studios,

[1] Stephens Follows, "Film vs Digital – What Is Hollywood Shooting On?," 2016, https://stephenfollows.com/film-vs-digital/.

will be preserved at any cost because of the status of the directors and actors involved as well as the commercial values associated with it. There are non-profit audiovisual archives that usually preserve the works of renowned ‘auteurs’ and maybe some filmmakers who produce arthouse or experimental works, but where does this leave the works of independent filmmakers whose works are not picked up by film festivals and are not stored in a managed archive? While they wait for their third film to finally receive any recognition, no one tells them that with digital contents, media fails, bit loss creeps in, corrupted files overwrite good backups, and proprietary file formats become inaccessible. Most importantly, with born-digital contents, there are no physical copies to return to when lost. But the truth is, no matter how successful or famous a movie is, movies shot digitally are inherently fragile by nature; the only medium that has been proven to last for decades, if not a century, is film. Neglecting born-digital files does not lead to decay, but to a complete loss.

Much to everyone’s surprise, digital files decay as well. And as obvious as this may sound, unlike silver particles found in celluloid films, pixels are directly proportional to the overall image quality and therefore is more sensitive by nature. Digital video information can be thought of as thousands of images recorded and presented one after the other in quick succession, where the resolution of the images is determined by the number of pixels used to represent it. Binary encoding can represent more than just the written word; they can represent images, sound, video, databases and websites, each with their own general types of encoding - and a change in a single bit from zero to one can cause degradation.

As Heather Ryan prefaces in her book *The No-Nonsense Guide to Born Digital Content* (2018), in this era where we are "forever catching up to the present", it seems unlikely that we will "land in a new normal or if there is a new normal, it's to expect a constantly changing digital knowledge ecosystem". To say that digital media does not provide a stable means to provide long term preservation is equal to saying that they are in a state almost as precarious as many silent films from one hundred years ago, which are only available as digitized copies now.[2] To that end, this paper aims to explore the current status quo of the manner through which movies are watched, stored, and distributed holistically, considering the different situations and needs of studios, archives, and filmmakers. The paper will first broadly discuss the shift in our moviegoing habit, and will take a look at how digital techniques, resulting in the highest resolutions, unconventional frame rates, and visual effects, are driving the movie industry, and, in turn, archiving and preservation. This will be followed by a closer look at the status of archiving and preservation of mainly feature-length movies in the movie industry, paying particular attentions to the Interoperable Master Format (IMF) and cloud technology. The paper is an interview-based reflection rather than a survey or a quantitative report. Further, considering the timeline given for this project, the scope of the research was confined to the US.

[2] Cynthia Lucia and Rahul Hamid, *Cineaste on Film Criticism, Programming, and Preservation in the New Millennium* (Austin: University of Texas Press, 2017), 325.

I. From Films to Files: The New Moviegoing Habit

1. Movies On-Demand

It is no longer easy to clearly define what television is, or what cinema is. Whatever the definitions may be, the advance in technology, internet connectivity, mobility, device fragmentation, and platform diversity are impacting how and what we watch. According to *The Nielsen Total Audience Report Q3 2018* published in 2019 by Nielsen, the leading global measurement and data analytics company, nearly seven in ten TV households have a device that can stream content and/or a streaming Subscription Video On-Demand (SVOD) service, resulting in what has been coined as out-of-home (OOH) TV viewing. It also found that, on average, US adults spend 10 hours and 30 minutes per day connected to media. Additionally, adults aged between 18-34 spend over one-third of their daily media usage on smartphones. Internet Connected Device/Smart TV app usage experienced the largest year-over-year growth in Q3 2018, and enabled smart TV ownership had the largest year-over-year growth across all races and ethnicities.[3]

For many parts of the world, it is no longer a luxury for families to own personal computers, let alone each member owning their own personal computers, and of course, smartphones. According to a research conducted by the Pew Research Center in 2017,

[3] "THE NIELSEN TOTAL REPORT 2018," 2019, https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/q3-2018-total-audience-report.pdf.

"more than half of all households in the US contained a cellphone but not a landline telephone". It goes on to note that "some 84 percent of American households contain at least one smartphone, according to a Pew Research Center survey conducted in fall 2016". Further, "desktop and laptop computers are nearly as common – 80 percent of households contain at least one of these devices," and while "tablet computer ownership is somewhat less widespread with 68 percent of households containing at least one tablet," "39 percent of households contain at least one streaming media device, such as an Apple TV, Roku or Google Chromecast".[4] As demonstrated by these facts and figures, the overall landscape of how we watch moving image contents are shaped by the growth of technology.

The growth of the Internet has generated new channels for motion picture distribution. Theaters spent huge amounts of money on digital projection systems, which have been introduced to replace the traditional film projection systems in movie theatres; according to the 2018 THEME Report by the Motion Picture Association of America (MPAA), "in 2018, the number of digital screens in the United States increased by one percent from 2017, accounting for 99 percent of all US screens."[5] This meant that it is not only individual consumers but also movie theatres that became dependent on the internet to 'download' digital movie files. The dominance of digital equipment and the internet means reduction in distribution costs. In theory, digital movies have the power to reduce the marginal costs of

[4] "A Third of Americans Live in a Household with Three or More Smartphones," May 25, 2017, https://www.pewresearch.org/fact-tank/2017/05/25/a-third-of-americans-live-in-a-household-with-three-or-more-smartphones/.
[5] "2018 THEME Report," MPAA, 2018, https://www.motionpictures.org/wp-content/uploads/2019/03/MPAA-THEME-Report-2018.pdf.

duplication to almost zero, thus overcoming the constraints that used to define theatrical exhibition.[6]

One of the effects of the rise of online streaming platforms is that the distinction between film and television has significantly blurred. Once upon a time, to watch a movie meant going out and dressing up, while television was the stay-at-home type. However, somewhere in the past few years or so, the lines started to blur. There are two main reasons behind this phenomenon. Firstly, we are witnessing the golden age of television. Today's shows have high production value and rich story lines that you would be forgiven for thinking you saw it in the theater. Starting from American Movie Classics channel's *Breaking Bad* (2008-2013) to HBO's *Game of Thrones* (2011 to current), *Silicon Valley* (2014 to 2019), Netflix's *Stranger Things* (2016 to current), and *House of Cards* (2013-2018), the quality of television shows has risen to a point where it would seem outdated for anyone to boast their preference of cinema over television. In the book *Visible Fictions,* first published in 1982, Producer and Professor John Ellis's argues that "a film on TV yields a very different experience to its viewer, unless that viewer is able to suspend the sense of watching TV and imagine instead the sense of being in a cinema".[7] While it still does come across as a reasonable argument, the sentiment is questionable in today's media viewing landscape. And as evident in the title of a lecture that the author gave in 2016 at the University of Groningen, '*Is 'Visible Fictions' Still Relevant'?',* the author also seems

[6] Virginia Crisp and Menotti Gabriel Gonring, *Besides the Screen* (New York: Palgrave MacMillan, 2015), 57.
[7] John Ellis, *Visible Fictions: Cinema, Television, Video (2nd Ed.)* (London: Routledge, 1992), 24.

to have noticed that television has become a very different medium since the book was first published.[8]

The second reason would obviously lie in the fact that the majority of us watch movies the exact same way we watch television, on television, computer, and smartphone screens. Inevitably, it is much more convenient, cheaper, and immediate. Even the so-called cinephiles rely on online streaming services to not only watch the latest movies but also old black-and-white or cult movies. For these reasons, today's definition of what it is like to love cinema (hence the term cinephilia) cannot be described the same as when there was a clearer distinction between film and television. This is not to say that film and television are now a completely unified entity. However, it is indeed a phenomenon which cannot be ignored. Historically, media scholars have contrasted film and television spectatorship, noting that moviegoing has been associated with an immobilized spectator seated in a theater directing a fixed gaze toward a giant screen. Television, by contrast, has long been associated with a smaller screen and a viewer capable of moving around while watching. However, once viewers began consuming movies at home, whether on television, videocassette, or DVD, new modes of spectatorship that place new emphasis on the role of mobility has made it possible for viewers to choose from vast programming menus wherever an internet connection is available.

[8] "Is 'Visible Fictions' Still Relevant? By John Ellis," October 19, 2016, https://www.rug.nl/research/icog/research/conferences/research-colloquium-film-studies-fall-2016/visible-fictions?lang=en.

On-demand culture raises new questions about how, when, and where we access movies. It is bringing both cultural and financial changes to the movie industry. For example, the persistent availability of movies through different VOD services has altered their value, often with the result that consumers feel less urgency to own copies of individual films. This essentially takes away one of the primary sources of income for moviemakers, which is a situation that has been especially damaging to independent producers. The blurred line between cinema and television, driven by on-demand culture, also means that it is not just television but also movies that are expected to be readily available in different "versions" at your fingertips. This resulted in both movies and television shows to use the Interoperable Master Format (IMF) as the "de-facto standard" for both movies and television shows, for submissions to streaming services like Netflix. According to James Snyder, Senior Systems Administrator at the National Audio-Visual Conservation Center (NAVCC) at the Library of Congress (LOC), the Library is "already receiving (their) first IMFs through Copyright submission". He added that "the format is now the de-facto international standard for production and distribution and most of the world's content creators are moving toward fully implementing it in their workflows," and hence "now is the perfect time to bring the subject of IMF up in the preservation community since we will start seeing IMFs in collections that New York University students, for example, will be seeing starting in two to five years."[9] As a matter of fact, in LOC's *2019-2020 Recommended Formats Statement*,

[9] James Snyder, email message to author, "Thank you (Soojin from NYU Tisch)" November 6, 2019.

IMF is listed as the recommended format for preservation for file-based video works.[10] The technical details and its implementations will be further discussed in Section Three.

II. Dealing with Data

1. Digital Techniques as a Storytelling Tool

Technology trends are driving an evolution of modern media workflows, including higher resolution imagery, higher frame rates, higher dynamic range, wider color gamut, and simply more cameras per shoot. Their impact on the digital storage required to capture, process, distribute, and archive this richer digital video content is significant. Today, a single large Hollywood production may shoot 100 hours or more of footage. Productions embracing these advanced technologies generate a large volume of materials. The David Fincher film, *Gone Girl* (2014), which generated about 500 hours of raw footage during its multi-camera 6K (4:1) Red Dragon production, is equivalent to about 261 TB of storage. Back in 2009, James Cameron's *Avatar* was hailed as the first movie to break one PB of total video production content. Five years later in 2014, The *Amazing Spider Man 2* (2014) required 2.4 PB of storage to scan from film to 4K.[11] The volume of data in the movie industry is growing substantially, especially as our expectations in regards to internet bandwidth and digital storage spaces increase while the costs associated with them

[10] "Library of Congress Recommended Formats Statement 2019-2020." (2019). https://www.loc.gov/preservation/resources/rfs/.
[11] "4K, HDR, HFR: Calculating the Storage Impact in Media Workflows," 2016, http://landing.quantum.com/rs/561-AAR-658/images/Tom_Coughlin_Whitepaper_2016.pdf.

decrease. However, there are many valid reasons as to why we shifted to digital methods for filmmaking. Low-budget productions with short uninterrupted shooting period and limited crew members sometimes may not even have the time to view dailies. Instead of one-light work prints as dailies, digital production provides low-resolution digital proxy files for editing.[12] Even if you can afford film, this means that you still have to worry about the small probability of lab errors resulting in the loss of irreplaceable footage, mistaken exposure, and film stocks being scratched.[13] Most importantly, as a result of the increase in the computer's ability to manipulate data, 'processing power', it became possible to edit digital movies on a personal computer. This is indeed beneficial for many independent filmmakers who could even use their phones to shoot a feature-length movie, much like what Sean Baker had done with *Tangerine* (2015).

Some argue that the cost that goes into shooting a movie on film is exaggerated when considering the high-end camera and crew with up-to-date technical expertise that are required for digital production, as well as for archiving and preservation. However, it is true that with digital, the filmmaking business becomes far more efficient and affordable in the immediate term. In fact, it is not just the cost and time efficiency that had resulted in switch from film to digital. For example, Director Ang Lee, who once went so far as to even avoid the digital intermediate (DI) process that had become a standard post-production practice in Hollywood, said during the promotion of *Gemini Man* (2019) – in which Lee explored the

12 Sonja Schenk and Ben Long, *Digital Filmmaking Handbook* (Boston: CENGAGE Learning, 2014), 4.
13 Drew Campbell, *Technical Film and TV for Nontechnical People* (United States: Allworth Press, 2002), 171–72.

uncanny realism of combining high framerates (HFR) of 120 FPS, together with 4K and 3D – that "it (digital) is a different media with different perception, different requirements. Digital doesn't want to be film, it wants to be something else. I think we need to get past that and discover what it is."[14]

2. Visual Effects

The days when visual effects (VFX) belonged to the exclusive domain of big budget sci-fi and superhero movies are long gone. VFX are now used extensively in almost any type of movie and television in a wide variety of genres—not only to create spectacular visuals but also to support storytelling, overcome production limitations and reduce costs. VFX as we know today could not have been possible without the ability to convert images into a series of numbers, and the development of computers that could handle and manipulate these numbers and spit out virtual imagery. Before digital technologies took center stage, visual effects were achieved through a combination of on-set and in-camera practical work, animatronics, stop-motion, painted glass panels, optical tricks, and chemical film manipulations (which would now technically be classified as 'special effects' (SFX) to differentiate the practical, real-life effects performed on the set and captured by the camera from digital manipulations which happen primarily during post-production).[15] Some of the most beautiful and creative visual effects in the history of film have been done with such

[14] Chris O'Falt, "Ang Lee Deserves More Respect for the Ambitious Digital Experiment of 'Gemini Man,'" October 10, 2019, https://www.indiewire.com/2019/10/gemini-man-ang-lee-interview-promise-digital-hfr-3d-cinema-1202180316/.
[15] Eran Dinur, The Filmmaker's Guide to Visual Effects: The Art and Techniques of VFX for Directors, Producers, Editors and Cinematographers (Focal Press, 2017), 7.

non-digital means, from Georges Méliès' *Le Voyage Dans la Lune* (1902) to George Lucas's *Star Wars: Episode IV - A New Hope* (1977).

Yet the dinosaurs that roamed *Jurassic Park* (1993) could not have been created, animated, and rendered without the emerging digital technologies and 3D applications, which owe its birth to Industrial Light & Magic. These technologies did not only open up almost endless new possibilities for visual effects, they revolutionized their use in the industry—from unique events in a handful of movies to a widespread filmmaking tool that's used extensively in every genre and on any budget.

Volumetric capture is a good example that shows how production can affect the works of archivists. Intel's massive volumetric capture stage, built to "help Hollywood and others record actors, musicians and athletes for augmented reality (AR), virtual reality (VR), holograms and anything in-between" by recording them in virtually any angle, measures "10,000 square feet and uses more than 100 individual 8K cameras, mounted all the way from the floor to the dome's concentric ceiling". Here, "each camera is connected to a battery of Intel servers housed on-site through a network of five miles of fiber-optic cable, capable of transporting data at a speed of terabytes a minute".[16] The level of volume of data that the servers need to deal with is so heavy that Intel's team had to build an insulated room to make sure that the noise was not interfering with the audio recording on stage. The

[16] Janko Roettgers, "Inside Intel Studios, the World's Largest Volumetric Capture Stage," October 7, 2019, https://variety.com/2019/digital/features/intel-studios-volumetric-capture-holograms-ar-vr-1203358126/.

final product that is produced from using this technology will be larger than ever, as it is capturing the subject from every angle possible, at 8K resolution. This begs the question: what to keep and what not to keep?

Figure 1. Intel's volumetric capture stage.[17]

During a discussion about the volumetric capture with Buzz Hays, the Global Lead of Entertainment Industry Solutions for Google Cloud, he said that "there are many very large data sets that create new challenges for archivists. The use of game engines now adds more complex time-based data to the pile, not to mention all of the volumetric file sets you have mentioned. If you look at this new information as more source material, it becomes a

[17] (Roettgers, 2019)

massive storage issue." More importantly, it calls attention to issues beyond storage. He added:

> I would posit that the workflows which create media from this data are less understood from an archiving standpoint. We need to embrace a production model which captures not only the source material, but all of the relevant metadata as well in a meaningful workflow archiving structure so that any one final asset can be recreated from the source material and its workflow metadata at a later date. This will require new workflow metadata standards in order to be effective. Workflows need to be treated as archive assets. I really think this poses a much bigger challenge than the volume of data at the moment.

Storage issues, while extremely important, is only one aspect of concerns that arise with using more complicated or new VFX technologies. As filmmaking procedures evolve, new workflows enter the picture. Without adequate measures to enable the recording of relevant metadata and workflows, recreation of final works would not be made possible.

VFX may sound like a domain that only belongs to big budget studio movies, but it is important to remember that it is also a tool that helps moviemakers save budget. The ability to change, modify, or enhance locations during the post-production process helps reduce costs, time, and bureaucratic complications. The ability to extend practical sets with VFX means that less time and money are spent on constructing large sets. Crowd-tiling techniques make it possible to populate big scenes with just a limited number of extras. Sure, visual effects cost money too; sometimes a lot. But careful planning and a good understanding of the medium can keep VFX costs considerably lower compared to real-world practical solutions. Removing an unwanted street sign on location, for instance, can

easily cost several thousand dollars for labor and municipal permits, while the VFX removal may be much cheaper, all without the legal and bureaucratic hassle.

3. 4K Ultra High Definition (UHD) and Beyond

The benefits of 4K UHD video go beyond pixel dimensions. Color and dynamic range are greatly enhanced, and these factors mean the saturation and contrast differences in images for 4K are well beyond HDTV. ITU-R's UHDTV standard Rec. 2020 currently defines color in terms of 10 or 12-bit data, which translates into the ability to render more than one billion colors.[18] Ever since the first version of Rec. 2020 was posted on the International Telecommunication Union (ITU) website on August 23, 2012 , 4K UHD has become commonplace.

UHD also allows for high dynamic range imaging, which is a photographic technique used to reproduce a greater range of luminosity or brightness in a scene. The result mimics the human visual system's ability to see detail in shadows or bright sunshine. UHD currently provides for frame rates up to 60 frames per second.[19] However, this comes at a cost. For example, to deliver 1920 ×1080 HD video at 60 frames per second (i.e. progressive) requires a data rate of 4.46 Gigabits per second. UHD video with 4:2:0 color subsampling

[18] "Recommendation ITU-R BT.2077-2," June 2017 [illegible] E.pdf.
[19] "Understanding High," November 2015, https://www.smpte.org/sites/default/files/users/user29811/wp-uhd-update-v3.pdf.

requires 13.36 Gigabits per second.[20] This requires a robust data networking infrastructure capable of meeting sustained bandwidth demands. As with all new technologies, UHD also seems to be seen less as something extravagant. Tim Schwab, Editor at the Culture House Productions, a full-service media production company, said that "even for smaller budget projects, we have generally been shooting UHD (or 3200x1600 in the case of the Alexa Mini) so we have some room to stabilize and do punch ins".

The race to control the 8K market among television makers is competitive. Samsung Electronics, the world's largest television vendor, was the first to commercialize 8K resolution televisions in late 2018. LG Electronics, the runner-up, has launched its own televisions in the top-notch resolution during the second half of 2019. NHK presented their challenges for preserving UHD content. Last year, NHK began 4K/8K test broadcasting on satellite channels and was planning on starting regular broadcasting in 2018 with a view to broadcast the Tokyo Olympics and Paralympics in UHD in 2020. The demand for archiving UHD materials and retrieval is increasing as 4K/8K program production is rapidly growing. NHK Archives are building a trial online archival system for UHD to assess many challenges. The storage capacity of 8K video is phenomenal so it is compressed to a 2K copy for search and preview. Although LTO-6 tapes are used as back up, they found that operationally this is a very slow process NHK is currently developing a business continuity

[20] Tom Coughlin, "4K, HDR, HFR: Calculating the Storage Impact inMedia Workflows," 2016, http://landing.quantum.com/rs/561-AAR-658/images/Tom_Coughlin_Whitepaper_2016.pdf.

plan to duplicate the data on cloud storage, and is aiming to move to a new broadcasting center by 2025. As a result, they are trialing a testbed of “everything in the cloud” to test scalability, reliability, migration and operational costs.

4. Structured vs. Unstructured

Structured data refers to data that uses a predefined and expected format. This can come from many different sources, but the common factor is that the fields are fixed, as is the way that it is stored (hence, structured). This predetermined data model enables easy entry, querying, and analysis. On the contrary, rather than predefined fields in a purposeful format, unstructured data can come in all shapes and sizes, and can come in many forms to be stored as objects: images, audio, video, document files, and other file formats.

5. The Originals

The management of the original materials are crucial in archiving and preservation because without them, works cannot be recreated. However, they are huge in size, and while major studios have the resources to make it a mandate to preserve the source files, the same does not apply for independent producers. Sandra Schulberg, who started the Independent Feature Project (IFP) and IndieCollect, has been on a quest to save American independent

cinema from technological obsolescence.[21] Schulberg noticed that there are filmmakers who were lamenting the fact that they do not know where their source materials are anymore heightened by the fact that the smaller independent film distribution companies are gone out of business. Conversely, she has been working with the WRS Motion Picture and Video Laboratory which used to handle a huge volume of overflow prints or foreign release prints – those that fell outside of the category of materials handled by Technicolor and Deluxe, which the big Hollywood studios used to process all their prints and archiving for years.

The WRS Lab was where independent filmmakers would send original negatives and dupe negatives to. When the company went out of business in 2001, all their resources were locked in a warehouse. 55,000 film elements were stored, many of which are original negatives from independent films. It was impossible to exactly identify the damaged proportion due to the warehouse's hostile environment involving a leaky roof. Schulberg has been making efforts to raise funds to move the items to a dryer location until a more articulate measure can take place. In addition, Schulberg launched an initiative to build the most comprehensive database of independent film and filmmakers possible, which involves tracking down all the elements that can be located around the world. Schulberg travels to Los Angeles often for fundraising, and many of her presentations are funded by the Hollywood Foreign Press Association. She has started to build a coalition of people who

[21] IFP is essentially the first independent cinema-focused production entity in the US and IndieCollect is a film preservation organization.

are genuinely interested in rescuing independent film, but it is, at the end of the day, a needle-in-haystack problem, because elements are scattered through the winds and when something is found, they are not necessarily the originals. However, this effort is still meaningful, as whatever that can be located can contribute to find valuable information or metadata, to see the evolution of a filmmaker. The elements are scattered everywhere, and the money to improve this situation is always lacking.

Filmmakers are eager to find any platform to show their works. When they finally find a distributor who is interested, and strike a deal, the original elements are passed on and the filmmaker often lose all track of where those assets are. As per usual, archiving is an afterthought.

No part of the film budget, up until very recently, included any line of items for archiving. It simply did not exist; it was just a part of studio overhead. Hence, while archiving remains as an afterthought across the field, it is true that the works produced by big studios are organized in a consistent and maintained archive infrastructure. This also means that the studios have the capacity to treat original files as utmost important source materials. According to Kathryn Claypool, Senior Manager at the Inventory and Archive Department at Paramount Pictures, “if it is a Paramount production we have what is called the entire deliverable schedule, which is a list of things – much like a legal document – that contains information such as what kind of production it is and how it was shot. The production team needs to provide it to us for archiving and preservation”. She added:

> The original RAW files are the furthest upstream items and materials that can be used to then make other materials. Therefore, the closer you are to the actual RAW material, the more options you have in the future too. And different types of technologies will be continuously introduced, so those RAW files are going to be the neutral material whereby you can down and upgrade as necessary. They are the most important. And keeping the original element is important because there is still information you can get out of the original negative. If we had our first generation of scanning as a 2K file for example, and had gotten rid of the original negatives, there is no pushing further. With photochemical, you can record more information than you can see. And you can also push the colors in a lot of ways. Maybe HDR 4K is the best you can get now but who knows what the future holds? And to truly preserve something you want to actually keep the original.

To gain a consensus around how institutions outside of the big Hollywood studios manage their original files, Editor Tim Schwab from the Culture House Productions, where contents from big studios as well as smaller production houses or independent filmmakers are edited, was consulted. He said:

> For raw media, storage and preservation is always a concern, but the cost of storing them is a bigger one. Our plan is to store it on LTO tapes and then eventually find a more permanent home for them in cloud storage. On the client side, they may do more to preserve the footage than we do, but we can't store media like we were running an archive without running into cost concerns.

6. Filmmakers' Perspectives

Unfortunately, it is rare to find filmmakers who are concerned with what is going to happen with their materials afterwards. If filmmakers are left unaware of the specific elements that need to be considered when forming a digital material such as the container format or codec, we are complicating the problem as the ideal moment to begin archiving, especially for born-digital materials, is at the moment of its inception – which also calls into question

the responsibilities that filmmakers, further, film schools, have. As Buzz Hays, also a producer himself, noted: “if you are not worried about preservation and are just creating a bunch of files, you could be spending a fortune saving stuff you cannot read in the future”. So, what is the point in trying to fix a problem without finding the root cause? Hays added: “I do not think we will ever change people’s minds to say ‘Stop producing content! We must solve the archiving problem before we can go back to making more content.” However, it is detrimental that filmmakers ask themselves questions such as ‘What am I creating? Where is it going to live? And how am I going to access it in the future?’. It can be said that the ‘Studio in the Cloud’ system, in which everything that occurs in creating content stays permanently, could be helpful in terms of files being organized from the production stage. As outlined in *The Evolution of Media Creation*, a whitepaper released by MovieLabs, the system will provide a convenient environment for filmmakers to speed up rendering times, provide workstations on-demand for artists and editors around the world, and make it easier to securely manage petabytes of data for hybrid or full cloud pipelines.[22] Kathryn Claypool notes:

> It would force people to organize files in a unified way because one of the biggest problems is that people spend a lot of time trying to find a file through hard drives or whatever ends up being in front of them when they are editing or any part of post-production process. If it is in the cloud, and gets uploaded straight on to the cloud, you could build in validation that says you need specific metadata items in a certain format, etc. So, it would basically standardize things. However, it would still require a great deal of human effort. I think there is still room for human Quality Control (QC) on some of our files that are ready to be packaged and sent to platforms. Those would be the materials that come from vendors. With them, you have to double

[22] “THE EVOLUTION OF MEDIA CREATION,” MovieLabs, 2019, https://movielabs.com/production-technology/.

> check their details and conditions, so those are some things that might not work as well when it is gone straight into the cloud.

Accordingly, from what it looks like now, it is tricky to say that for archiving and preservation, the 'Studio in the Cloud' system is going to bring an unprecedented benefit. The same cannot be said for making and cutting a movie together. It means that you would not have to transmit files back and forth. The files are left there, users access what they need, put it back, and let the next person take it. Whatever benefit the 'Studio in the Cloud' system brings to archiving and preservation, it seems inevitable that a new set of strict policies will have to be in place so that the integrity of archival workflows can be maintained. It could also mean that archivists need to be more digitally literate than ever before as it will require more complex discussions and negotiations on archival procedures and policies with engineers. However, it is still uncertain whether the same system would be something that can be leveraged to equal degree by independent filmmakers with more limited budget. Buzz Hays noted that "in the near term, it will be difficult for non-technical independent filmmakers to fully utilize cloud for content creation unless they use some prepackaged solutions like Adobe Creative Suite (once it's ready in the cloud)". This is because "the complication lies in the initial setup of the compute, storage and networking required to make use of cloud resources, which is fairly technical". Apart from the initial setup, without the resources to manage the system, it would be difficult to take full advantage of a cloud-based system. He continued: "Managing the monitoring of cloud resources can be challenging for independents as the costs can spiral out of control if they're not paying close attention to the utilization of these resources in

order to stay on budget. That's tough to do when you're trying to get a film done with limited resources." He also shared that there are Cloud orchestration tools available that lets you remotely access powerful computer in the cloud like BeBop Technologies, Teradici, iTopia and Conductor – that make the deployment part easier for smaller scale productions. While not impossible, it will be some time before independent filmmakers without a technical team and on a budget can take full advantage of the 'Studio in the Cloud'. However, the studios are set up to implement a full virtual studio implementation over a much shorter period of time.

Whether or not the 'Studio in the Cloud' system takes place in the near future, the fact remains that filmmakers need to be aware of what they are shooting. The cost of storage came down to the point where filmmakers have the luxury to shoot as much as they want to. Consequently, there is not enough planning that goes on for it to inform the most expensive part of a moviemaking process; production. As opposed to having an intent going in, the editing process becomes a sifting-through-the-rubble situation. Does this necessarily mean better quality? A blockbuster movie, which will remain unnamed, apparently used a 300:1 shooting ratio. The crew numbered the cameras they were with different alphabets, and they ran out of alphabets – which meant that it surpassed 26. It looks like the work discipline we used to see with movies shot using celluloid film still needs to be adopted in digital shooting. Gradually, we are going to start to feel the effects of this in a few years when archiving will be taken more seriously with the ever-growing volume of digital files.

7. Case Study: How does a filmmaker's lack of awareness and knowledge in moving image preservation affect the works of archivists?

At the Association of Moving Image Archivists (AMIA) 2018, a panel discussion titled *Collecting Born-Digital Material at the Source: Acquisition Strategies & Lessons Learned* was held. The hour-long session featured moderator Erwin Verbruggen, Project Lead Research and Development at the Netherlands Institute for Sound and Vision, and panelists Ann Gant, Head of Film Conservation and Digital Access Department at the EYE Museum; Laura Davis, Digital Project Specialist at the Library of Congress; and Chew Tee Pao, Archivist at the Asian Film Archive (AFA), who discussed "challenges of and solutions for getting born-digital content from makers and how to treat born-digital materials once they arrive."[23] During the session, Archivist Chew Tee Pao from AFA shared some of the challenges that arose when working with a public agency, during the agency's collection of born-digital materials as deliverables. According to the agency, the list of format types AFA suggested for submission (such as DCDM, DPX, in addition to the mandatory formats such as DCPs, and ProRes 422/4444 HQ as well as H.264: the three formats that the most common formats that filmmakers have been working with), were "non-exhaustive", and stated that they would rather specify a single format that AFA thinks would suffice for preservation, which obviously does not exist in the case of digital files. This resulted in

[23] "Preliminary Program Topics," December 5, 2018, http://www.amiaconference.net/preliminary-program-topics/.

AFA narrowing down the list (to H.264, ProRes 422 or 4444, along with different versions if applicable, and DCP if available). These were the formats that were 'easy' enough for filmmakers to produce. This meant that AFA, it its own capacity, had to approach the filmmakers for the intermediate and archival formats as it deemed necessary. He added that it "at least guarantees the accessibility of the films in the short term, until AFA figures out a better and economical solution for the long term". While he believes "the agency could have been the body to dictate the archival formats it should receive", he also acknowledges that "they were bound by certain constraints as a non-archival agency".[24]

Chew emphasized that filmmakers must understand how born-digital materials 'work', as he went on to share that AFA often receives a chunk of unkempt files with names that do not follow the appropriate file naming conventions. On top of this, the fact that the files are not even thoroughly checked before being passed on to AFA slows down the process of selection, accessioning, and ingest – as they would often receive a huge list of files including dailies and outtakes. After the discussion, when asked about the types of topics that he thinks film schools should teach their students, he listed areas such as the fundamental difference between film and digital, overall history of preservation, the progress of different medium, and challenges that archivists face as the result of the absence of such knowledge on the filmmakers' part.

[24] Chew Tee Pao, email message to author, "Re: Born-digital files related question from NYU student (AMIA 2018)" December 13, 2018.

III. Introduction to the Interoperable Master Format (IMF)

1. The Abundance of File Formats: Why IMF?

The range of different file formats in an archive is simply astonishing. Tim Schwab, Editor at Culture House Productions, said that for lower budget projects, they have often been working with ARRI Alexa Mini camera recordings (MOV files with a ProRes 422 (HQ) or ProRes 4444 codec) or C300s (MXF with an XF-AVC codec). And for their larger projects, they are "usually recording RAW, so, ARRIRAW with an Alexa LF or R3D when shot with a RED". For Paramount Pictures, for example, there are TIFF and DPX sequences, a huge volume of ProRes files, the original camera files from different cameras like RED and ARRI, and there are also output files from different software such as Avid or Premiere, as well as different transcoder projects. In terms of submissions to Paramount Pictures' clients, Kathryn Claypool said that each client has different specs. She used an example of a work with only Standard Definition (SD) version available. If the client wants a High Definition (HD) version, Paramount Pictures' Media Services Team will go and check if they even have a HD version of that title. If they do not, they have to go back to the client and let them know that they simply do not have it. She said:

> We are not going to make a HD version for like, two clients. It is too expensive. So, whether or not we create more copies of something all boils down to the deals that are made for acquiring materials for different platforms. They are going to have to pay more for the HD version of that title.

In addition, according to Claypool, the deliverable schedule that the Production Team has to pass over to the Asset Management Team requires the Production Team to provide the Asset Management those materials in certain formats, but they are separate from the transcoding that the Asset Management Team has to carry out to deliver to streaming and all kinds of other actual consumers. She said: “Suppose Amazon wants a certain movie. They have very specific requirements as to how they will have their files. There are client profiles and that enables us to deliver those files that are usable to each client depending on what their requirements are.” Paul Charron, Executive Director of Asset Management at Paramount Pictures, said: “Suffice to say, there is just a wide variety - which is why the Interoperable Master Format (IMF) will be useful.” Before diving into the details of IMF, it needs to be emphasized that it is a componentized package that holds different audiovisual and metadata information – it is not a single file format developed solely for the purpose of preservation. There have been other efforts made by the Society of Motion Picture and Television Engineers (SMPTE) with their work on the Archive eXchange Format (AXF); another package type format, as well as the non-profit audiovisual archiving community with Matroska File Format with FFV1 video encoding in which its standardization is currently being undertaken by the Internet Engineering Task Force (IETF)’s Codec Encoding for LossLess Archiving and Realtime transmission (CELLAR) working group that have been built specifically to meet the demands of archiving and long-term preservation. Despite that, this chapter will explore IMF specifically because with OTT platforms like Netflix and broadcasters like Walt Disney/ABC now requiring that

HD/UHD/4K material must be delivered to them as an IMF package, many vendors such as Avid, BlackMagic, Design, Dalet, EditShare, Ooyala, Signiant, Sony, Tedial, Telestream, and Tektronix have released products that support IMF.[25] While IMF is not yet a popular topic of conversation among audiovisual archivists, especially in non-profit audiovisual archives, opening up a dialogue about a widely used format in the commercial film and television world that will soon become as common as DCP would be beneficial. As mentioned previously, according to James Snyder, Senior Systems Administrator at the National Audio-Visual Conservation Center (NAVCC) at the Library of Congress (LOC), the Library is already receiving IMF copyright submissions and in LOC's *2019-2020 Recommended Formats Statement*, IMF is listed as the recommended format for preservation for file-based video works.[26]

Before the world went file-based, exchanging media between or within institutions like studios or post-production sites was fairly simple. Typically, you would send your tangible videotape to the receiver along with the metadata written on a piece of paper. As long as you had the right decks, there would not be major issues for playback. Ever since media started coming in files from the 1980s, the number of file formats has been growing exponentially. Interoperability became more and more problematic for institutions where exchanging media was a significant part of their work. For example, a Windows media file

[25] Michael Grotticelli, "Distribution & Delivery Global Viewpoint," (November, 2018). https://www.thebroadcastbridge.com/home/category/distribution-and-delivery/entry/12383/netflix-mandate-prompts-increased-interest-in-the-interoperable-master-form.

[26] "Library of Congress Recommended Formats Statement 2019-2020." (2019). https://www.loc.gov/preservation/resources/rfs/.

might not play in a Mac operating system and vice-versa. A standard was needed to manage the production process from editing to distribution. Long story short, this explains how the Media eXchange Format (MXF), which provided the framework for Interoperable Master Format (IMF), came about in 2004.[27]

In an era where the internet has allowed viewers from all over the world to enjoy audiovisual contents dubbed or subbed in a particular language from the comfort of their own countries, homes, or specific locations, major network broadcasters and content distributors are increasingly expanding their worldwide reach and sending a single piece of content to multiple outlets simultaneously. IMF, a component-based interchange format published by SMPTE in 2013, is in essence an evolution of the Digital Cinema architecture that addresses the problem of creating and managing many different master versions of the same material.[28] These versions may differ in content, such as the theatrical release of a feature film, the airline edit, and the broadcast television edit. Localized versions will also have different audio and subtitles or captions for alternate languages and perhaps different video segments for the titles, end credits, and even localized portions of the program material itself. IMF also manages different versions based on the playout format, such as delivery by broadcast television or delivery by an OTT (Over-The-Top) streaming service.[29]

[27] Tom Butts, "IMF: Unleashing the Benefits of File-Based Content," (February 23, 2018). https://www.tvtechnology.com/news/imf-unleashing-the-benefits-of-filebased-content.

[28] "How IMF can Benefit a Facility Where Versions Matter." (2017). https://www.dalet.com/sites/default/files/2018-03/Dalet_IMF_WhitePaper_2017.pdf.

[29] "Quality Control for File-Based Video Workflows." (October 24, 2016). https://www.telestream.net/pdfs/general/Quality-Control-for-File-Based-Video-Workflows-25W608940.pdf.

IMF is still a fairly new format and was created with the intent of being a predominant format for business-to-business (B2B) transfer of media in the supply chain, not specifically for preservation purpose per-se[30]. However, the fact remains that it is a worldwide standard that is being increasingly implemented by the post-production world as a format for 'finished' audiovisual works, and is starting to be seen as a format that is suitable for preservation as well. As such, this chapter has been written to introduce a broad overview of the history of the development of IMF, its technical structure that makes it ideal for today's digital landscape, the benefits that it is bringing to the film and television industry, discussion on MXF and DCP (Digital Cinema Package), and its effectiveness as a preservation format.

2. The History and Current State of IMF

In 2007, talks around the need for a distribution file format arose among the studios. During the following year, in 2008, the University of Southern California's Entertainment Technology Center (USC – ETC) created a common "safe room" where leading Hollywood studios such as Warner Brothers, Sony Pictures, Universal, Paramount, and Walt Disney/20th Century Fox were encouraged to meet and discuss how this common format would manifest. In 2011, USC – ETC published the IMF spec 1.0.[31] Later in the same year, the 35PM50 Working Group at SMPTE started to work on IMF.[32] SMPTE released the first

30 Annie Chang, "SMPTE Standards PDA Webcast: IMF (Interoperable Master Format)" (January 16, 2013). https://www.youtube.com/watch?v=bmhv36hmSP4.

31 This particular specification published by the Entertainment Technology Center (ETC) is still available here: "Interoperable Master Format (IMF) Specification," (February 19, 2011). https://wikileaks.org/sony/docs/05/docs/IMF/IMF_Specification_V1.0.pdf.

32 Refer to this link for SMPTE standards committee outline: "SMPTE Standards Ecosystem." (2015).

IMF specification in 2013, and later that year, France's Commission Supérieure Technique (CST) released a Film Archive extension proposal (and eventually came up with Application #4: another form of JPEG2000 variant, in 2016).[33] IMF was strictly focused on 'Hollywood' until the Fall of 2014, when Netflix announced that all Ultra-High Definition (UHD) submissions must be in IMF. From January 2016, major studios began distributing sample IMF packages to the supply chain. This was followed by the National Association of Broadcasters (NAB) announcing IMF as their standard library format and a number of major US broadcasters adopting IMF.[34]

Since then, a lot of work went into creating different Applications Specifications for different purposes. Parties outside the US also became involved in this effort. For example, in February 2018, SMPTE, the UK-based Digital Production Partnership (DPP), the North American Broadcasters Association (NABA), and the European Broadcasting Union (EBU) announced that they were working together to develop a global IMF specification for broadcast and online media – the Technical Specification (TSP) 2121-1:2018 – Application DPP (ProRes). Based on the IMF standards specification (SMPTE ST 2067), the goal was to agree on and publish a global standard for the file-based interchange of multi-version, finished audiovisual works by using ProRes, one of the most commonly used codecs by

https://www.smpte.org/sites/default/files/images/1664_SMPTE_Wallchart_v5c.pdf.

[33] "Interoperable Master Format - Business Situation (by Marquise)." https://documents.pub/document/imf-interoperable-master-format-business-situation.html.; It is difficult to find information written in English for CST's work on Application #4 but free online translation will allow you to understand the general narrative of this article: Elsa Sepulveda, "La CST-RT021 Aboutit À La Publication Du Standard « Cinema Mezzanine » À La SMPTE," (December 4, 2016). http://www.mediakwest.com/cinema/item/la-cst-rt021-aboutit-a-la-publication-du-standard-cinema-mezzanine-a-la-smpte.html.

[34] (Chang 2013); Kennet Eriksson, Björn Isakson and Kojo Mihic, "IMF – Interoperable Mastering Format, for Local Markets," (August 22, 2018). https://www.linkedin.com/pulse/imf-interoperable-mastering-format-local-markets-bj%C3%B6rn-isakson/.

international distributors for program mastering and exchange.[35] Building on the work of SMPTE, DPP and NABA have worked together to achieve is the development of SMPTE Technical Specification for an IMF Application based on the requirements common to many in the broadcast and online sector.[36] DPP and NABA continued with the specification work with the publication of the second TSP for the IMF for broadcast and online use cases – SMPTE TSP 2121-4:2019 – Application Constraint DPP (JPEG2000) in August 2019.

The abundance of different platforms that we use to watch film and television today, as well as the different languages, subtitles, or cuts that are required, means that versioning has become a very important issue. A component-based workflow offered by IMF avoids the replication of assets and allows for creating deliverables on-the-fly. And more importantly, although IMF was first developed out of the needs specified by Hollywood studios, over time, it has become applicable to broadcast and online media as well.

3. The Technical Structure of IMF

IMF is a collection of files which together make up the content (be it a full-length movie, a documentary, or an episode of a series). Each version of a piece of content is called a "composition", which is the set of all essence and metadata that makes up a particular deliverable. An example of a composition might be the US television version of the

[35] "DPP and SMPTE have Got Together, but what does this Mean for You?" (October, 2018). https://www.tedial.com/blog/article/dpp-smpte-together/.; "Technology Fact Sheet: IMF." (September, 2018). https://tech.ebu.ch/docs/factsheets/IMF_v2.0.pdf.
[36] "SMPTE TSP-2121 Application DPP." https://www.smpte.org/technical-specifications/tsp2121-app-dpp?action (accessed December 2, 2019).

theatrical cut of a movie, or the French version of an episode within a series. When you are playing back an IMF Composition, a device first looks for a file called a Composition Playlist (CPL). IMF eliminates the need to create a dozen plus master copies by separating each market requirement into individual Composition Play List, or CPL, that references the available essence components (MXF media files) included in an IMF package (the IMF "bundle" that holds a uniquely arranged CPL for a specific version would be the Interoperable Master Package, or IMP).[37] The CPL is a bit like an edit decision list—it defines the playback timeline for the composition and includes metadata applicable to the composition as a whole. One very important point here is that the CPL is not designed to contain essence but rather references external (MXF) Track Files that contain the actual essence.[38] This allows multiple compositions to be managed and processed without duplicating common essence files. Due to the flexible construct of the CPL mechanism, the playback timeline can be decoupled from the underlying track files, allowing for economical and incremental updates to the timeline as needed.[39]

[37] (Grotticelli 2018)
[38] "Media Composer 2019." (May, 2019) [illegible] final.pdf.
[39] "IMF: A Prescription for Versionitis." (March 7, 2016). https://medium.com/netflix-techblog/imf-a-prescription-for-versionitis-e0b4c1865c20.

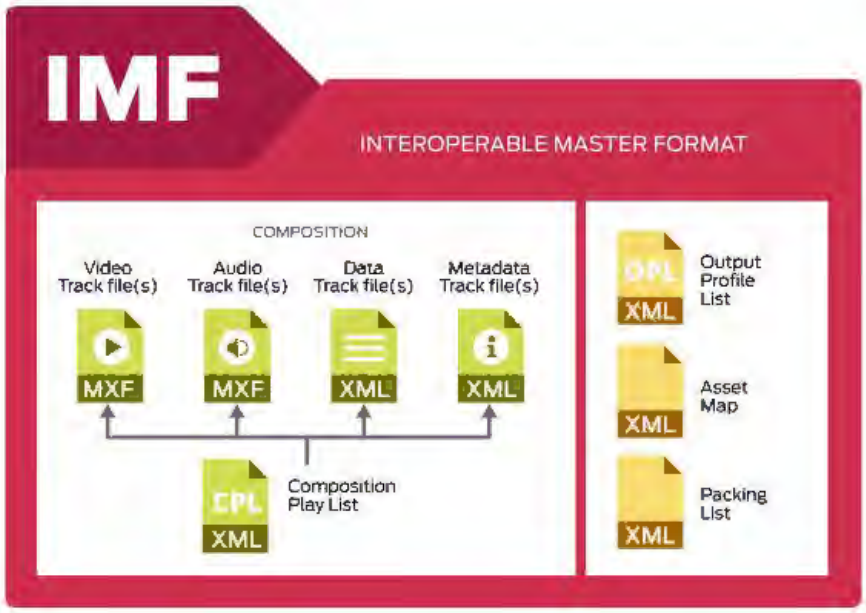

Figure 2. The structure of IMF.[40]

IMF also allows you to create different distribution formats from the same composition, if required. This is accomplished through another file, called the Output Profile List (OPL), which specifies the processing/transcoding instructions required for each distribution format. For example, you can change the display aspect ratio without generating any new media. The Asset Map is a mapping of the files inside a package to their location inside the little bundle. The Packing List tells you exactly what is inside the shipment and passes on the information to the receiving end. It is the checksums for all files in the package to verify package integrity.[41]

[40] Ulrich du Bosque, "Transforming IMF with OPL," August 16, 2019, https://www.thedpp.com/news/transforming-imf-with-opl.
[41] "IMF - SMPTE ST 2067."Tektronix (2019). https://www.abacanto.net/wp-content/uploads/2019/02/TEKTRONIX-TIPSTRICKS-158-IMF-ES.pdf

But it will take a while for us to get used to this complex format. The MXF and XML files are identified by Universally Unique Identifiers (UUID) instead of file names or paths, because IMF is meant for computers and machine systems. You need an IMF-aware system that can ingest these files and register them and manage them for you and know where they are at. This enables automation, which in turn enables fast workflow. One of the key things about IMF is that you have lots of identifiers inside it. You can identify a version of a composition, or a specific piece of media, or a specific segment on a timeline. You can even identify the two resources that go together and join up to make that segment. Devlin adds that "there are identifiers at every level of IMF, and once you have them you can then start to track things. And once you can track you can start to automate movement. You can look back in time to replicate a decision you made yesterday with a new piece of material automatically. We are not relying on file names; everything is done with identifiers."[42] Due to its flexibility, IMF requires significantly reduced storage space for an archive to keep multiple versions of a work. Technically, you would only need a single set of core assets with different supplemental packages to create different versions instead of solid, heavy master files for each version. It should also be noted, however, that using UUID in data storage may not sit well in the context of small or non-profit archives, and even individual production companies.

[42] George Jarrett, "Interoperable Master Format for TV: The Solution to Versioning Woes?" (May, 2017). https://www.ibc.org/manage/imf-for-tv-the-solution-to-versioning-woes/1975.article.

An example of the value of IMF was demonstrated with the *House of Cards* television series in 2015. Right before the release date for Season Three, Netflix modified its opening logo to add sound. This was after the variants such as languages had been mastered and delivered to the playout servers. Thaks to IMF, only a trivial adjustment mad to be made to the CPLs. As a result, the process did not make any effect on audio/video/subtitle synchronization, no additional Quality Assurance step was required, and the series was successfully released on its original schedule.[43]

4. The Technical Structure of IMF Part Two: Application Specifications and ACES

To improve interoperability across several workflows, IMF Applications exist which encase some mastering parameters to specific delivery constrains, hugely simplifying technical encoding efforts. As can be seen from the below table, IMF has modular Applications that are "plug-ins" to the core framework. They allow for specific functionality to determine areas such as codecs and specific resolutions or frame rates.[44] They set specific requirements that IMPs must obey regarding either naming conventions, compulsory metadata, package and virtual timeline structure, etc.

[43] "IMF: A Prescription for Versionitis." (March 7, 2016).
[44] Walter Arrighetti, "The Interoperable Master Format (IMF) in Film Preservation," (October, 2019). https://search.proquest.com/openview/5ed0858b9cf37009acd7dc43211c1198/1?pq-origsite=gscholar&cbl=29010.

Table 1. IMF Applications Specifications[45]

IMF Application	**Description**
App.1	Uncompressed image data
App.2	UltraHD 4K video content based on JPEG2000 compression. App.2 Extended (2E) also exists, targeting HDR formats (in the International Telecommunications Union (ITU) Recommendation Rec.2020, display referred colorimetry)
App.3	Video content based on Sony SStP studio profile codec (HDCAM SR)
App.4	Cinema Mezzanine which targets Digital Cinema content for theatrical screening. App.4 is mostly limited to theatrical distribution, sharing the scope and some limitations of the DCP format
App.5	Specifies additions to the IMF Master Package (IMP) structure when video encodes ACES colorimetry. Other workflows may be combined with App.5, including the one proposed here with respect to ACES-based content
RDD 45 (non-standard)	App.2E but using ProRes
TSP 2121-1	DPP ProRes
TSP 2121-4	DPP JPEG2000

While many Application Specifications can be considered for use in archives because MXF is used to wrap the essences, the Application Specification that may be of interest to audiovisual archivists is App.5 that came about in 2018 (although the most commonly used in the distribution sector would be App.2). It uses the Academy Color Encoding System (ACES) – a color management pipeline for ways we create content for both episodic and feature contents. This image format was already a standard but the Academy of Motion Pictures Arts and Sciences had no way to wrap audio and subtitles to it. As such, they

[45] Walter Arrighetti, "Review on ACES Colour Spaces," June 16, 2016, https://acescentral.com/t/review-on-aces-colour-spaces/171.

decided to incorporate the ACES image format into IMF (also known as SMPTE ST 2067-50), which is meant for archive master.[46] The growing interest by studios and OTT companies in encoding mezzanine/master formats to preserve the original ACES colorimetry is also evident, as this means that interchange and archival copies are stored in one shared colorimetry. The de-facto standard for this purpose has become IMF.[47] The ACES system, an open-source digital image file interchange and color management system, is also applicable for low-budget movies, and was designed to create a digital image file interchange and color management system that was not dependent on any specific camera, display, or production or postproduction tool.[48] Regardless of the budget, movies today use many different cameras or even visual effects facilities. The color system inoperability issues that arise from this can be removed using ACES as the unifying image interchange and color management system, which offers maximum image fidelity can be achieved throughout the production process.[49] As the ACES2065-1 color space (the standard ACES color space) provides a larger coverage than the human eye, detail that may not be able to be viewed or used today will be available as soon as future technology can reproduce it. Accordingly, integrating this information with IMP means that future ACES productions can be made without image quality loss. In addition, since ACES is backward-compatible to earlier versions, future accessibility is assured.[50]

46 (Chang 2018)
47 Edoardo Provenzi, *Color Image Processing*, First edition ed. (Basel; Beijing; Wuhan; Barcelona; Belgrade: MDPI, 2018), 211.
48 *International Cinematography Summit at the Academy (Tech Level: Low)* (Academy ACES, 2018), https://www.youtube.com/watch?v=PGXwnMLT9YU.
49 "ACES Primer, Glossary and Quick Start Guides - Get up and Running with ACES!," ACES Central, August 12, 2018, https://acescentral.com/t/aces-primer-glossary-and-quick-start-guides-get-up-and-running-with-aces/1433.
50 Walter Arrighetti, "Review on ACES Colour Spaces," June 16, 2016, https://acescentral.com/t/review-on-aces-colour-spaces/171.

5. What is MXF?

MXF and IMF are fundamentally different in that MXF is just a container format while IMF is a componentized format and therefore the two formats cannot be compared side-by-side. However, it might still be worthwhile to remember that MXF AS-02 (An MXF Application Specification for storage of MXF program components to enable versions & inventories, for use in a multi-version, multi-lingual, multi-delivery media environment) makes use of *MXF* to select the "playable range" of these single Essence Component Files while in IMF, the same fundamental concept used in Composition Play Lists (CPL) of Digital Cinema is used, which use human-readable eXtensible MarkUp Language, *XML*.[51] To borrow the words of Bruce Devlin, Standards Vice President at SMPTE or 'Mr. MXF', "there are maybe twenty or thirty really good MXF binary programmers in the world today; XML is much more generic, and there must be hundreds of thousands of top-quality XML programmers out there". He adds that "given the growing amount of localized versioning that we are now faced with, it makes sense to use a more generic technology like XML to represent the various content versions whilst maintaining the proven MXF AS02 media wrapping to store the essence components".[52]

Targeted at the interchange of audio-visual material with associated data and metadata, MXF is a SMPTE-approved open file format developed by organizations such as Pro-

[51] To explore different MXF Application Specifications, visit: "AMWA Application Specifications." (2019). https://www.amwa.tv/specifications.
[52] Bruce Devlin, "MXF AS02 and IMF: What's the Difference and Can They Work Together?" (July, 2015). https://www.dalet.com/news/mxf-as02-and-imf-whats-difference-and-can-they-work-together.

MPEG, the EBU, AAF Association, and more, to establish interoperability of content between various applications used in the production chain.[53] Ideally, this would lead to operational efficiency and creative freedom through a unified networked environment. Unlike formats defined by single companies, MXF was developed by SMPTE committees representing broadcasters and equipment vendors from across the industry.[54] It is compression-format independent and contains a standard data model for metadata; it allows essence and metadata transfer without the metadata elements having to be manually reentered, thus making it suitable for versioning and robust file exchange between institutions.[55] The MXF standard was defined as a wrapper which supports a number of different streams of coded 'essence', encoded in any of a variety of image and audio compression formats, along with a metadata wrapper which describes the material inside the package.

The MXF standard, with timecode and metadata support, was intended as a stable distribution method for professional video and audio applications, regardless of the platform. As a result, the MXF design was encompassing and managed to include support for diverse applications between servers, and carriage of a subset of the Advanced Authoring Format (AAF) data model, under a policy known as the Zero Divergence Directive (ZDD), which enables MXF/AAF workflows between non-linear editing (NLE)

53 Bruce Devlin, "MXF— the Material eXchange Format," (July, 2002). https://tech.ebu.ch/docs/techreview/trev_291-devlin.pdf.
54 "Quality Control for File-Based Video Workflows." (October 24, 2016). https://www.telestream.net/pdfs/general/Quality-Control-for-File-Based-Video-Workflows-25W608940.pdf.
55 Nick Wells et al., *The MXF Book* (Oxford: Routledge Ltd, 2006), 9.

systems using AAF and cameras, servers, and diverse devices employing the MXF wrappers.[56] However, although it was a widely adopted format with various shims that allow different Application Specifications like AS-02 for versioning, AS-07 for archiving and preservation, and more, it was clear that in the studio business there was a need for a format that ensures a higher level of asset compatibility from one institution to its suppliers, clients and outlets.[57]

The standard was effective at media interchange but by 2005, there were interoperability issues because the specifications were too broad a design. MXF compliant files were not necessarily interoperable with other MXF compliant systems that had adopted a different application of the details of the specifications, and different vendors adopted different sub-format versions of the specification in order to promote their products that were technically compliant but not necessarily interoperable with similar systems from a different vendor.[58] Hollywood learned from the MXF experience and understood the need to define a digital media specification for reliable cinema distribution, which resulted in the Digital Cinema standards.

[56] (Wells and others 2006) 283.
[57] Shim is “an application-specific constraint set that constrains an Application Specification in order to tailor the general specification to a specific purpose,” from: "MXF Archive and Preservation Format Registered Disclosure Document." (December, 2018). http://www.digitizationguidelines.gov/guidelines/rdd48-2018_published.pdf.
[58] “A notorious example of these vendor to vendor interoperability problems was the Sony XDCAM and the Panasonic DVCPRO P2, both of which were MXF compliant, yet their files were mutually unintelligible to the other system,” from: Jay Batista, "Software Infrastructure Global Viewpoint - IMF Supports Interoperability and Increases Performance," https://www.thebroadcastbridge.com/home/category/software-infrastructure/entry/11410/using-imf-supports-interoperability-while-increasing-performance

6. Why IMF When We Already Have Digital Cinema Package (DCP)?

While the DCP is about theatrical content distribution, IMF is about providing businesses with a master format for creating multiple tailored versions of the same piece of content for different audiences. In today's digital landscape for film and television culture, in which the distinction between film and television has noticeably blurred, a myriad of different screens and platforms are thrown into the equation. In this respect, it seems only logical that IMF had to be developed despite the existing robust componentized structure preceded by DCP.

In reality, DCP is not an ideal choice for preservation purposes because it uses intra-frame, JPEG2000-based, visually lossless (yet lossy) compression. DCP is also constrained to projection technology standards, limited as regards resolution (HD, 2K, 4K), aspect ratio (Scope, Flat, or "Scope-across-flat" Container), frame rates (24, 25, 30fps plus their HFR variants), soundfields (stereo, 5.1, 7.1, Dolby Atmos and other immersive audio). Also, encrypted DCPs, which is frequently used as archived masters, can be played back on pre-determined devices only and pending certificates' expiration, with the risk of preventing generation of any new decryption keys, unless a resilient preservation chain for the whole Public Key Infrastructure (PKI) is established and maintained. Once KDMs expire, or authorized playback devices become obsolete, the content will be forever locked inside the encrypted DCP.[59] IMF is free from this worry as it is a B2B tool assuming that agreed-upon content-security measures are already in place among the exchanged parties.[60]

[59] (Provenzi 2018) 220.
[60] (Provenzi 2018) 219.

7. IMF for Preservation

In addition to the advantages of IMF listed above, the most impressive trait of IMF is that it uses consistent references across its asset files via both hashes and UUIDs, thus making it possible to archive IMPs in a physical media that are spread across vaults, or sublimate them in several Cloud "data lakes". It will still be possible to address the individual content as if played back, at logical level, from one file system.[61] There are also several open-source tools available for processing IMPs, as can be found in the official IMF User Group website.[62]

However, since IMF was not developed specifically *for* preservation purpose, there needs to be more discussions around how it can be molded so that the advantages of IMF can be used efficiently and fully as a preservation format. An example of how this is manifesting was shared by David Deelo, Executive Director of Engineering & Technology at Sony Pictures Global Mastering & Servicing, at the "Workshop on Preserving Movies with IMF" event (organized by the IMF User Group and held in Amsterdam on September 13, 2019) – where he said that Sony has been "a big proponent of using IMF for all the distribution mastering, and has also been looking at the Archive eXchange Format (AXF) as a solution for archive mastering".[63] AXF is a standardized interoperable format for data and metadata preservation that was designed with long-term preservation as a key characteristic. It is

[61] (Arrighetti 2019)
[62] "Open Source."https://www.imfug.com/open-source/ (accessed December 2, 2019).
[63] "Workshop on Preserving Movies with IMF." (November 15, 2019). https://www.youtube.com/watch?v=7oyJFhgIrzw.

storage agnostic where objects can span across medium, can store and relate all kinds of data, of any type, and preserves full context of data within a standardized archive. It can also hold metadata unlimited in languages, formats used, and standards applied. As such, there is a growing interest in combining the traits of IMF and AXF together to create a synergized effect.[64] Technically, IMF can be stored within AXF where each AXF Object is a fully self-contained, encapsulated collection of files, metadata and any other ancillary information which adds relevancy or value to its contents. AXF is designed to handle a single file encapsulation as easily as it can for hundreds of millions of files. Regardless of whether they are created on data tape, spinning disk or flash and optical media with or without files systems, AXF Objects are equivalent. As such, creation and handling of AXF Objects on differing media become easier. Using preservation information such as fixity and provenance as specified by the OAIS model, AXF enables long-term archiving of content. In addition, data structure redundancy and cryptographic hash algorithms ensure resilience of data.[65]

[64] "Archive eXchange Format Community Home. "http://www.axf.io/history.html (accessed December 5, 2019).
[65] S. Merrill Weiss, "AXF-Archive EXchange Format: Interchange and Interoperability for Operational Storage and Long-Term Preservation," SMPTE Motion Imaging Journal 123 (SMPTE, January 2014), 42–52, https://doi.org/10.5594/j18371XY.

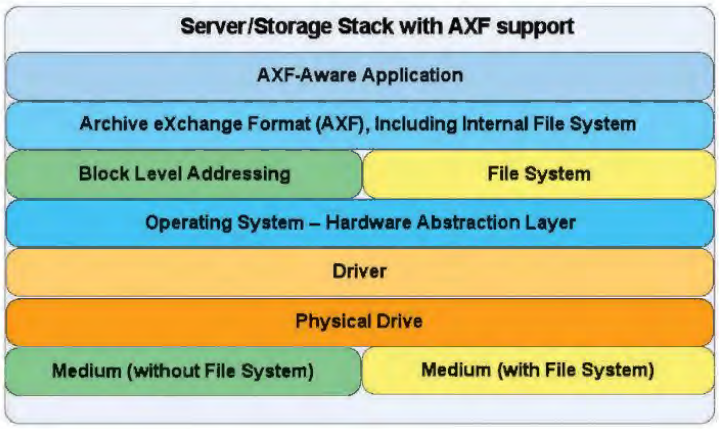

Figure 3. Hardware/software stack incorporating AXF writing to and reading from media.[66]

As can be seen from Figure three, AXF exists on its own level in the stack. This enables to "change the AXF-aware application, switch between use of a file system or block-addressing, change the operating system and hardware abstraction layer, change the medium, the physical drive, and the driver for the drive, without having to modify the standardized AXF protocol," and AFX can therefore "remain standardized according to the needs for AXF functionality without concern for the changes in technology surrounding it".[67] The audiovisual archiving community will need to figure out where the crossover is between AXF and IMF so that their advantages can be best utilized for preservation purpose.

Since IMF was developed as a business-to-business (B2B) solution, it can be assumed that it will not be a popular choice amongst non-profit or small archives, in which the lack of

[66] (Weiss, 2014)
[67] (Weiss, 2014)

employees with strong technical background could also play a part. However, the shift in our digital landscape is extremely rapid, and as far as digital preservation goes, archival trends and practices could possibly take unexpected turns. More importantly, it is crucial that communities share knowledge and information as we do not know when or where there may be a crossover. From what it looks like now, it looks like IMF will stick around for quite a while. With DCP's intrinsic disadvantages eliminated, IMF is being widely adopted by the film and television industry for its ability to simplify distribution workflows and lower costs and time-to-market; and there are a lot of discussions around how it can be shaped into ways that can serve different, specific purposes. This paper was written with the mission of introducing, to fellow audiovisual archivists, a new format that will be the new DCP. It would be interesting to see how much change will take place in just a few years' time.

It looks like more work will be needed to ensure that IMF is reliable for preservation, not just to save storage space. Buzz Hays said:

> The one thing that we do not have right now, which we do need, is an extension of IMF that is a superset of IMF – which gets back to the notion of 'here is the set of all of the raw materials that were used to create the IMF (which is only a small part of the data created for a production) and it's imperative to create links to all of the original source material so that any portion of the content can be reproduced at some point in the future'. Currently, IMF has no reference to the source material that it came from. We need to take the same model and extend it back up to where the material came from so that you can take that and the metadata, meaning we will be able to recreate the IMF at that point because IMF is just a placeholder for the final formats.

Regardless, IMF is important now because the streaming channels require so many versions. There is a lot of interest especially among streaming platforms that do not create master files until it is requested; they can be created on-the-fly in the cloud. As far as production goes, IMF is an important technology adopted by vendors and studios across the industry as a measure to effectively save storage space and make versioning easier and cheaper. As such, it would be helpful for audiovisual archivists to be aware of this format type no matter which type of institution you belong to. Tim Schwab from Culture House Productions shared that their first IMF experience will be this year, because this is the first year that they will be working with big streamers. The IMF is created and the delivery is managed by a third-party post-house, so they will be minimally involved in the creation of the file itself as most production companies would be. It can be assumed that IMF creation and management will remain relevant only to major studios or large post-production facilities for now if not continually, but smaller archives and independent moviemakers would inevitably have to understand the structure of the format to effectively communicate with streaming platforms when the chance arises.

IV. Unlocking the Potentials of Cloud Technology in Audiovisual Archives

1. If Not Linear Tape Open Technology, What Do We Have?

When media enterprises reach their threshold, they have three options for long-term media

archiving: on-premise digital magnetic tape-based archival storage in a Linear Tape-Open (LTO) library; the cloud— either public or private; or some hybrid combination of the two. Before choosing, these enterprises must consider the following four important factors: cost, performance, accessibility and security. Additionally, in some situations, the operational model of a business must be considered as a part of any evaluation. Specifically, how a business operates will determine whether it is better served by making an upfront capital investment in storage technology or by choosing a recurring monthly operational expense to archive media assets.

LTO technology, developed jointly in 1998 by Certance (now Quantum), Hewlett Packard, and IBM with the intention of creating an open format to deliver multiple product and media sources to the market, still remains primary solution for media preservation across archives. It remains an industry standard specifically in the entertainment world as well, as it enables data compression, track layout, and error correction to maximize capacity and performance while maintaining data integrity. Despite the wide use of LTO tapes across archives, there are concerns that arise with its stability since the fact remains that they are magnetic media on polyester base. However, for now, there seems to be no better alternative for longer-term strategy than LTO tapes since cloud services do not yet readily provide fixity reports to users on demand. According to Kathryn Claypool, one aspect of Paramount Pictures' policy is to store two copies of the same file on two separate tapes. Their system operates in such a way that, if one of the files results in an error against the database during a regular healthcheck, then the other copy— which presumably works

without errors— would be utilized instead. Subsequently, a new copy is made after the 'failed' file is disposed. In the event of a disaster, such as a flood, the problem becomes much more complex. A Disaster Recovery (DR) system in which tapes containing copies of the file are packaged in boxes and have been shipped off to another location that is unaffected by the disaster. While such system provides for the minimum level of safety, the caveat is that one will never be able to guarantee that the tapes will consistently remain in good condition and still requires hardware to read those tapes.

Another issue with LTO tapes is that they work two generations back only. What ends up happening when one has a restoration or archiving issue, for example when a collection has LTO-3, often times archives will purchase the LTO-5 as it is slightly easier to find. An LTO-3 drive is difficult to find, and if found, may not even work at all. An LTO-5 drive will play LTO-3 because it goes back for more than two generations. The newer ones will not. Hays noted:

> Studios realized a long time ago that manufacturers of LTO tape generally advise the longevity of tapes to be about 30 years, but technically speaking, it is good for 30 years if you never use it. The life span shrinks quickly when in use. As a result, most of the studios will only hold onto LTO tapes for about five years, some ten. And as obvious as this may sound, they are constantly retrieving everything in the archive and making copies of the files onto the newest LTO format until it's time to repeat the process a few years later. And the new tape format is still based in fragile magnetic media. So here we have this leap-froggy situation, since there is not much of a longer-term strategy.

But if not Linear Tape Open Technology, what do we have? While this chapter explores cloud technology in detail, I suggest that we should take a moment and think about the kind

of physical digital storages that can ultimately provide better preservation than LTO tapes, because cloud technology alone will not suffice.

2. Cloud and NDSA Levels of Preservation

Data storage on cloud is one feature of cloud computing, which uses a large shared pool of computing resources to add additional computing power. It reduces manual effort on the part of the IT team, which reduces costs and significantly shortens the time needed to start using new computing resources. Cloud services can provide easy, automated replication to multiple locations and access to professionally managed digital storage and integrity checking. As a result, bit preservation (durability) of digital information can be at least as good (or better) than can be achieved locally. In addition, major cloud service providers like Amazon and Google very recently started incorporating theories or concepts that once only resonated with the archiving community, such as the NDSA Levels of Preservation. On why this model was adopted into Google's services, Hays said that while the initiation of the NDSA Levels predates his tenure at Google, he "was talking with some of the storage team, and they said that due to the adoption of the levels in other industries, it made sense". He said that "it is important to note that there are far more stringent requirements and practices in order to deliver eleven-nines of durability on the Google Cloud platform, but the NDSA Levels are a great way to get people used to a much more complex system without a steep learning curve," and that "the NDSA Levels of Preservation is well understood by archivists so it made sense to adopt the model as it helps with the 'trust' factor when moving traditional archives to a new storage medium". He further explained

that "by 'new', we are talking of spinning discs, but with significantly higher durability than can be achieved on a local server setup - trust is still a real issue".[68]

3. Deployment and Service Models

There are different cloud deployment models available; public, private, hybrid, and community. In most cases, cloud services are operated within the domain of public Internet by companies like Amazon and Google. Public cloud storage services, such as those offered by Amazon, Microsoft and Google, have captured the attention of audiovisual archives, largely because of pricing. However, determining the true cost of cloud archiving can be tricky. First, an institution must decide whether it needs fast access to data stored in the cloud or it can wait a longer time to retrieve its content. Amazon Web Services illustrates the point. Its low-cost Glacier storage ($0.004 per gigabyte per month as of this writing), which is intended for archiving applications, typically can take three to five hours for a standard retrieval. Its more expensive S3-IA (Simple Storage Service-Infrequent Access) service ($0.01 per gigabyte per month as of this writing) promises low latency and high throughput.

Other costs, such as the per gigabyte price of data retrieved from storage, must be factored in.[69] Public cloud services also can provide geo-spacing which protects archived data

[68] Buzz Hays, email message to author, "Hi from Soojin (NYU student)" March 2, 2020.
[69] Here, the cost would include making retrieval requests and the cost of bandwidth to connect an enterprise's site with the cloud.

against the ravages of localized natural disasters, such as tornadoes, hurricanes, floods and earthquakes, by distributing data across multiple geographies.

Some larger organizations have also found value in operating private clouds inside their own data centers, where similar economies of scale begin to apply. Hybrid clouds combine on-premises infrastructure, or private clouds, with public clouds so organizations can reap the advantages of both. In a hybrid cloud, data and applications can move between private and public clouds for greater flexibility and more deployment options. Finally, community cloud is a cloud service model that provides a cloud computing solution to a limited number of individuals or organizations that is governed, managed and secured commonly by all the participating organizations or a third-party managed service provider. It is a hybrid form of private clouds built and operated specifically for a targeted group. These communities have similar cloud requirements and their ultimate goal is to work together to achieve their business objectives.

It is common for mid-sized and large audiovisual archives to have multiple workflows, each with its own storage requirements. As a result, many organizations have chosen to leverage the value of a hybrid archive solution based both on cloud and on-premise storage resources. A hybrid strategy can also offer the archives the benefit of taking advantage of the best of both archiving alternatives.

A good example of a hybrid model would involve cloud bursting, an application deployment model in which an application runs in a private cloud or data center and bursts into a public cloud when the demand for computing capacity spikes, scales up processing power by leveraging more CPUs and GPUs as needed in the cloud to accomplish a compute-intensive task. It offers a good use case that illustrates how archives can take advantage of the best of both storage options. For instance, a hypothetical media enterprise that wishes to use cloud-based artificial intelligence tools, such as speech-to-text functionality and facial recognition, to generate useful metadata describing archived video assets, can take advantage of processing power in the cloud and continue to use on-premise storage to achieve its specific cost, accessibility and security goals. In this hybrid application, a media enterprise can temporarily push a portion of its archive—most likely lower-resolution proxy files—to the cloud where AI processing can create additional metadata. The metadata can then be retrieved and married to the on-premise archival copy, while the version of the metadata in the cloud as well as the proxy files can be discarded. The process can be repeated as budget allows, until the entire LTO-based library is updated with enhanced metadata.

There are also three different service models that users can choose from; IaaS (Infrastructure-as-a-Service), PaaS (Platform-as-a-Service), and SaaS (Software-as-a-Service). IaaS refers to cloud-based services, such as pay-as-you-go storage, networking, and virtualization. This provides the access to resources offered by the provider. It can include, for example, virtual or physical devices. The most common examples are Amazon

Web Services and Microsoft Azure. IaaS helps enterprises create and manage their data as they scale, paying for the server space that they used to develop hardware or software. PaaS refers to cloud-based platform services that provide developers with frameworks that they can use to create custom apps. The model provides programmers with a platform that is used to create software delivered via the Internet. All servers, storage, and networking can be controlled by the business or a third-party provider. The most known examples are Windows Azure, Google App Engine, Openshift. PaaS makes development, testing, and deployment processes fast, easy, and cost-effective. SaaS is cloud-based software available for purchase on a subscription basis. It is mostly used for apps that need both web and mobile access. SaaS products do not need to be downloaded and installed on individual devices. Most of them can be run directly from web browsers. It should be taken into account that customers are not responsible for hardware and software updates.

4. **Object Storage**

With object storage, data is bundled with customizable metadata tags and a unique identifier to form objects. Objects are stored in a flat address space and there is no limit to the number of objects stored, making it much easier to scale out.

The metadata tags are a key advantage with object storage as they allow for more accurate identification and classification of data. Using a search application, users can easily search for a specific object, even if the data itself is not easily searched (such as an image, or

media clip, or data set). For storing unstructured data, object storage is the most desirable option. Search capabilities and unlimited scale make object storage ideal for unstructured data, which, according to the research firm IDC, will take up 80 percent of world's data by this year, 2020 – the year expected by IDC to see data growth of up to 44 times since 2009.[70] The below remark by Hays highlights the importance of interoperability between cloud systems and legacy Media Asset Management (MAM) systems adopted by archives:

> Object storage is the only option that can effectively store this data at scale. It simplifies the file management, although it has complicated various workflows which are reliant on Portable Operating System Interface (POSIX) file systems.[71] Most of the M&E software we deal with has not yet been written in a cloud native form, and file system translation has not been modernized as a result. This makes it more complex in the near term since file system translation is required. When dealing with massive numbers of files, object store is better at quick retrieval and security. Workflows are complicated by the fact that both structured and unstructured data sets are involved, much of this is due to MAMs. We are working toward a model where there is no centralized MAM application, but rather applications reference the relevant structured or unstructured databases directly within the application where it is required, thus avoiding constant migration of different data types into and out of MAMs.

Apart from the cost, fixity check is another factor that makes audiovisual archives hesitant about diving into cloud storage. With object storage system, fixity of files can be managed through the traditional method of archiving – as in, LTO tapes and robots, institutions could go and run a fixity check to see how a file has changed in any way. However, within the cloud, the so-called erasure encoding which uses object storage system, enables the user to save parts of a file in multiple locations rather than as a full complete file. More

[70] Jean-Paul Isson and Minha Hwang, *Unstructured Data Analytics* (Somerset: John Wiley & Sons, Incorporated, 2018), 65.
[71] The Portable Operating System Interface (POSIX) is a family of standards specified by the IEEE Computer Society for maintaining compatibility between operating systems.

specifically, the file is split and encrypted into separate chunks and remain disparate. These separate chunks of a file can be merged back together at any time to recreate the same file; but at this point in time, doing so incurs an egress charge. Paul Charron from Paramount said:

> You cannot really put something in a deep storage and run a fixity check on it without incurring cost. Paramount for example, runs millions of fixity checks every year, or every month. Basically, a lot of files. This is every single frame, not just assets. And they are able to get immediate reports that show if the checksums haven't changed. The big cloud storage systems are checking all these bits, and they guarantee you a number like the eleven-nines durability in which they are saying that the chance of losing something is almost one in a trillion, but they cannot give you back immediate reports. But then for audiovisual archives, they cannot help but worry if they are safe right now. And you just would not know – unless you pay that charge.

A system that either reports back what the cloud was seeing when they checked it, or, figure out a way that allows users to ask, without incurring that egress charge, to check the files and make sure there are no problems at the current moment would be ideal. But as for now, it is not supported by major public cloud service providers. If something fails, which service providers assert that it is practically impossible, what happens is that because the bits are spread out, the system will detect the ones that are broken, and figure out which bit they are exactly so as to grab a copy of that particular bit from another file separated geographically, and replace the existing bit. This, as a result, saves space. Buzz Hays further explained that "the way a file is stored in the cloud is not at all the same as it is when stored in our own computers". In the cloud, as soon as a file is uploaded, it is broken up into thousands of pieces called 'shards'. Each shard is stored in a different place and has

its own hash and the collection of shards which constitute a document or a frame, and that whole collection of shards has its own MD5 hash. There are multiple hashes which operate simultaneously. Therefore, the integrity of the whole file is constantly checked in and out. Fixity check on cloud happens on a low, completely automated level. In the case of a failure, although extremely unlikely to happen, the user is not notified directly as the files are repaired or replaced by duplicate data automatically – which is exactly what archivists fear the most. He added that "this is a part of the trust exercise that we need to go through". Technical issues such as this is another reason why audiovisual archivists should be acquainted with terminologies such as object storage or unstructured data. How a file gets saved is a crucial piece of information in this context, and archivists will have more power in their hands when choosing storage options.

5. Potential Useful Features for Archivists

Buzz Hays said that "cloud is typically used as a DR copy in many cases, although some archives are beginning to migrate their data into cloud as the primary source with data tape as the DR copy". Cost is a significant factor, but this seems to be changing rapidly. The biggest benefit customers are finding with cloud is not so much the storage as it is the intelligence that is derived through the AI/ML tools. This, combined with the ease of transcoding on the fly creates significant new opportunities for monetization of media archives. He added that "there is still work to be done on preservation formats, but for now, the hybrid model works quite well".

Another useful and concerning feature for archivists is the cloud's function of automatically sending less used files to deeper storage, thereby maximizing cost vs. access. For Google Cloud, this is not available as a product *per se* but there are enough ML/AI implementations that enable such feature. As of now, it is not a product. Instead, it is a monitoring function that Google has in the analytics - but it is not to say that they cannot come up with smarter ones. In addition, studios cannot easily deaccession items simply because they do not know what can be deleted permanently.

One of the most pressing issues for audiovisual archives is discoverability, as archiving and preservation in the feature-length movie (and commercial television) world is around sales. Enhanced discoverability can help sales and marketing teams to make informed decisions for business. Kathryn Claypool noted that "if there is a cloud system that had complete awareness of what all the files were, then the computer will tell you what makes sense to keep with enhanced discoverability". In relation to this, she added that "it would be helpful if cloud, at the ingest point, enabled us to build in metadata that records the intricate relationship between items beyond just parent-child relationship and also categorize each asset so that we can distinguish the originals from the derivatives." This would mean that in order for the computer to let you know what to throw away, you have to first provide the criteria on what is considered savable and what is derived from something or not. She added that "the caveat is that that the status and hierarchy change continuously, so the key would be to keep everything updated and build into the format the metadata and connect

the conformance and linkage". While this kind of system or function has not been realized in cloud services yet, conversations around building a highly detailed metadata system that allows archives to easily narrow down items will need to continue as the volume of items in audiovisual archives grow. Deaccessioning is an extremely challenging task for archives; most inventory managers feel wary of permanently deleting items in their collections as they might prove to be useful or even necessary in the future. If cloud-based storage systems allow you to identify straightway the family structure of a certain item, and distinguish unique copies versus the rest, archives would be able to effectively save up on storage costs by only keeping what cannot be thrown away.

Computer storage capacity greatly resembles a physical space in that it costs money, and it begs questions such as: Should the data be compressed? What is the long-term plan to perpetuate the data? Migration? Data redundancy at multiple different geographic locations?

V. Interviews with Professors from Film Schools: Archiving and Preservation, a One-Person Job?

In this day and age, just what is a movie studio? It used to be the place where all the departments – the camera, lighting, directors, editors, and more – were centralized. This has changed over time. There are now efforts around centralizing the production process so that the data is sitting in one place, people work on it, go away, and come back to work on

it, somewhat like a studio system. Unfortunately, issues pertaining to building longer-term strategy has been left up to archivists who may have brilliant ideas but no resources to implement them. In an era where technology is advancing at a pace one can hardly keep up with, archivists must constantly update and share knowledge with the community so that there is some sort of measure ready, although imperfect, that will help scholars, students, and audience to access today's movies in the future when there is another new set of dominant digital format, interface, and storage. However, is it really just archivists that are responsible in ensuring that digitally created movies are still accessible decades from now? The specific elements that need to be considered when forming a digital material such as the container format or codec mean that the ideal moment to begin archiving for digital materials is at the moment of its inception, which calls into question the responsibilities that filmmakers, or even further, film schools, have. Raising awareness around the importance of archiving and preservation is vital for the long-term, so that organizations will understand the level of attention and budget required to maintain a robust archive. To that end, strengthening education so that those who were trained can pass on their knowledge will be ideal. If film schools were to include minimal training on archiving and preservation in their film production curriculum, at least to a degree where filmmakers are able to understand the inherent fragility of digital files, archiving and preservation could be regarded as an obvious step in the filmmaking ecosystem. The responses given by some Professors of film schools in the US when asked whether there is an interest for their schools to educate student filmmakers on the basic elements of moving image preservation, and if possible, what they should teach, and how the school archives and preserves their

student works, were somewhat disappointing but expected.

Professor Michael Burke, also the Associate Dean of the Kanbar Institute of Film & Television at the Tisch School of the Arts, New York University, said that while preservation is something that is always on his mind, "not much is done"—which also applies to educating filmmakers about moving image preservation.[72] However, digital preservation practice (having the facility and equipment to keep digital copies of student films) has allowed the school to "maintain high-res versions of all screeners given to the school," according to Caitlin Stickels, Evening Manager at the Digital Media Library in the department.[73] The works that enter school-run student film festivals: Sight & Sound Showcase, New Visions and Voices, and the First Run Film Festival are received in ProRes 422/MOV and copied to the digital server for preservation.

Professor Ira Deutchman at the School of Arts Film Program (former Chair of the Film Program from 2011-2015) at Columbia University wrote that "while there is both interest and desire for Columbia to be active in preservation and restoration work, right now there is no formal program, and not much that is being actively done". He added that "the subject comes up now and then, but the school does not have funding or bandwidth to take it on right now". In terms of student work storage and preservation, he believes "that the department is now storing digital files of student work, but this is very recent, so there is no archive from all the many years of earlier work".[74] The school's effort to store digital

[72] Burke, Michael. Interview by Soojin Park. Interview Type. Broadway 721, Dean Michael Burke's Office, November 14, 2018.
[73] Caitlin Stickels, email message to author, "Re: MIAP student (Soojin) needs help (again)!" December 11, 2018.
[74] Ira Deutchman, email message to author, "RE: iradeutchman.com Inquiry Form" November 21, 2018.

files of student works, according to Donggyun Han, a student in the school's graduate program, said that "all students submit DVD copies as well as MOV files of their final work on a server, as all of them—with very rare exceptions to do with insurance reasons—are played at the school's annual film festival".

Some responses were more detailed. R. Patrick Lile, Business Administrator of Operations for Northwestern University's Radio-Television-Film Department and documentary producer and director, said that while he believes "it is not necessary that all students or filmmakers know and understand archival and storage, archiving and preservation is a very necessary part of the film industry" and "is certainly a useful tool to have a basic understanding". He added:

> Contemporary archiving and even the term archiving continue to be redefined. Modern digital filmmaking and how it is archived be it Vimeo or a hard drive is different than the work a trained archivist who knows how to preserve and restore film. While our students in the MFA in Documentary Media program use archived material in some of their work, we do not offer courses in archiving/preservation of film.[75]

Professor Caroline Frick at the University of Texas at Austin's Department of Radio-Television-Film, who is well known for her research in moving image archiving, also the founder and Executive Director of the Texas Archive of the Moving Image, said:

> Preservation is not something that the majority of filmmakers/film students think about (much like most professional filmmakers). We have one of the largest programs in the

[75] R. Patrick Lile, email message to author, "FW: Question from NYU grad student" November 20, 2018.

> country with thousands of undergraduates and a fairly large MFA cohort. The MFAs take courses on media history but nothing to do with preservation. It is, per usual, an afterthought.[76]

Dino Everett, Archivist of the Hugh M. Hefner Moving Image Archive at USC's School of Cinematic Arts, when asked to which extent he agrees on the importance of teaching filmmakers the basic elements of archiving and preservation, answered "a hundred percent". He said that "it could begin with basic training on film handling, because it will help them gain understanding that while they are creating their work in digital, realizing that movies can be tangible would help instill a sense of value to the works and help the students realize that they need to be handled with care". After citing file naming conventions, file formats, and aspect ratios as some of the things that could be taught, he also added that the discussion does not only apply to film schools as "the fact that every field of education is using born-digital files like CAD drawings and PDF files means that universities in general should start building robust digital preservation systems and teach students how to handle them." His reply also sparked interesting thoughts around copyright ownership of student works; he said that "while students retain the underlying rights of their works, meaning that they do not need USC's permission to sell or distribute a feature based a short they made while in school, the rights to the work itself belongs to USC".[77] This can be found on the school's website: "Intellectual property in any work produced at USC that uses substantial University resources is protected under a USC copyright and

[76] Carole Frick, email message to author, "Moving image preservation questions from NYU grad student" November 6, 2018
[77] Interview with Dino Everett, November 29, 2018.

owned by USC".[78] Although this means that the revenue generated by USC's student works goes to the school (and it is only on rare occasions that student shorts make money), the school can preserve the works without having to rely solely on taking risks based on fair use, particularly the ones in more urgent danger of deterioration.[79] Copyright ownership do not necessarily mean physical ownership, but when the work is no longer accessible, what is the point of owning the rights?

The kind email responses and in-person discussions with the above educators, archivists, students, technicians, and filmmakers draw the conclusion that the interviewees are aware of the attention that must be given to archiving and preservation. However, in general, it does not seem like much thought has been put into educating filmmakers on the fundamental principles of archiving and preservation. For many of them, especially who are neither archivists nor documentary filmmakers, the idea seems to be perceived as something detached from the reality; some of them also kept mistaking the question as asking whether or not film schools should have a program dedicated specifically to moving image preservation.

Conclusion

Like Timothy Robert Hart and Denise de Vriehe have noted in "Metadata

[78] "Copyright at SCA," January 30, 2020, https://cinema.usc.edu/admissions/copyright.cfm.
[79] "Exclusive Rights in Copyrighted Works," December 2, 2018, https://www.copyright.gov/title17/92chap1.html#106.

Provenance and Vulnerability", the fact that "digital objects currently being created often go through several modifications makes it difficult to identify the original or authentic copy of the object", and this also means that "care must be taken at the time of creation and curation of any digital objects because although some metadata are typically generated automatically, many elements that will play a pivotal role later must be created manually".[80] What is often taken for granted is the fact that digital information also exists in a physical space, like hard disk drives or flash drives. The details of what constitutes archiving and preservation methods may have changed, but the basic logic remains the same in that information is recorded onto a surface. However, it is equally true that digital media is not only more unstable but more complex than what most archivists have worked with. Handling, archiving, and preserving digital contents is very different from doing the same with film. All this cannot be done alone; when the students, filmmakers, and scholars in the future inherit our contents, they would be able to understand the context and reason behind those contents only by reading the documentation you produced, that explains the how and why of the choices you made. We do not even have to look that far ahead. Every day, there are new posts being uploaded on internet forums with titles like: "Adobe Premiere destroyed my short film!", or "My external hard drive that contained all my RAW files crashed". By sharing our findings, we must work together to feel somewhat comfortable in this game of "forever catching up to the present".

[80] Timothy Robert Hart and Denise de Vries, "Metadata Provenance and Vulnerability," Information Technology and Libraries 36, no. 4 (December 1, 2017): 24–33.

Bibliography

"4K, HDR, HFR: Calculating the Storage Impact in Media Workflows," 2016. http://landing.quantum.com/rs/561-AAR-658/images/Tom_Coughlin_Whitepaper_2016.pdf.

"2018 THEME Report." MPAA, 2017. https://www.motionpictures.org/wp-content/uploads/2019/03/MPAA-THEME-Report-2018.pdf.

"A Guide To The Interoperable Master Format (IMF)," 2019. http://www.telestream.net/pdfs/datasheets/App-brief-Vantage-IMF.pdf.

"A Third of Americans Live in a Household with Three or More Smartphones," May 25, 2017. https://www.pewresearch.org/fact-tank/2017/05/25/a-third-of-americans-live-in-a-household-with-three-or-more-smartphones/.

"ACES Primer, Glossary and Quick Start Guides - Get up and Running with ACES!" ACES Central, August 12, 2018. https://acescentral.com/t/aces-primer-glossary-and-quick-start-guides-get-up-and-running-with-aces/1433.

"AMWA Application Specifications," 2019. https://www.amwa.tv/specifications.

"Archive EXchange Format Home." Accessed December 5, 2019. http://www.axf.io/history.html.

Arrighetti, Walter. "Review on ACES Colour Spaces," June 16, 2016. https://acescentral.com/t/review-on-aces-colour-spaces/171.

"AXF – Archive EXchange Format." SMPTE, December 2012. https://slideplayer.com/slide/10798342/.

Batista, Jay. "Software Infrastructure Global Viewpoint - IMF Supports Interoperability and Increases Performance," December 2018. https://www.thebroadcastbridge.com/home/category/software-infrastructure/entry/11410/using-imf-supports-interoperability-while-increasing-performance.

Bosque, Ulrich du. "Transforming IMF with OPL," August 16, 2019. https://www.thedpp.com/news/transforming-imf-with-opl.

Butts, Tom. “IMF: Unleashing The Benefits Of File-Based Content,” February 23, 2018. https://www.tvtechnology.com/news/imf-unleashing-the-benefits-of-filebased-content.

Chang, Annie. “SMPTE Standards PDA Webcast: IMF (Interoperable Master Format),” January 16, 2013. https://www.youtube.com/watch?v=bmhv36hmSP4.

Coughlin, Tom. “4K, HDR, HFR: Calculating the Storage Impact in Media Workflows,” 2016. http://landing.quantum.com/rs/561-AAR-658/images/Tom_Coughlin_Whitepaper_2016.pdf.

Crisp, Virginia, and Menotti Gabriel Gonring. *Besides the Screen*. New York: Palgrave MacMillan, 2015.

“CST-RT021-Part2:2013,” 2013. http://www.cst.fr/wp-content/uploads/2015/03/CST-RT-021-C-2013-Part2-EN-Mezzanine-format-specification.pdf.

Devlin, Bruce. “MXF AS02 and IMF: What’s the Difference and Can They Work Together?,” July 2015. https://www.dalet.com/news/mxf-as02-and-imf-whats-difference-and-can-they-work-together.

Dinur, Eran. The Filmmaker’s Guide to Visual Effects: *The Art and Techniques of VFX for Directors, Producers, Editors and Cinematographers*. Focal Press, 2017.

“DPP and SMPTE Have Got Together, but What Does This Mean for You?,” October 2018. https://www.tedial.com/blog/article/dpp-smpte-together/.

Ellis, John. Visible Fictions: *Cinema, Television, Video (2nd Ed.)*. London: Routledge, 1992.

Eriksson, Kennet, Björn Isakson, and Kojo Mihic. “IMF – Interoperable Mastering Format, for Local Markets,” August 22, 2018. https://www.linkedin.com/pulse/imf-interoperable-mastering-format-local-markets-bj%C3%B6rn-isakson/.

Follows, Stephens. “Film vs Digital – What Is Hollywood Shooting On?,” 2016. https://stephenfollows.com/film-vs-digital/.

Grotticelli, Michael. “Distribution & Delivery Global Viewpoint,” November 2018. https://www.thebroadcastbridge.com/home/category/distribution-and-delivery/entry/12383/netflix-mandate-prompts-increased-interest-in-the-interoperable-master-form.

Hart, Timothy Robert, and Denise de Vries. “Metadata Provenance and Vulnerability,” Information Technology and Libraries 36, 36, no. 4 (December 1, 2017): 24–33.

"How IMF Can Benefit a Facility Versions Matter," 2017. https://www.dalet.com/sites/default/files/2018-03/Dalet_IMF_WhitePaper_2017.pdf.

"IMF - SMPTE ST 2067," Tektronix, 2019. https://www.abacanto.net/wp-content/uploads/2019/02/TEKTRONIX-TIPSTRICKS-158-IMF-ES.pdf.

"IMF: A Prescription for Versionitis," March 7, 2016. https://medium.com/netflix-techblog/imf-a-prescription-for-versionitis-e0b4c1865c20.

International Cinematography Summit at the Academy (Tech Level: Low). Academy ACES, 2018. https://www.youtube.com/watch?v=PGXwnMLT9YU.

"Interoperable Master Format - Business Situation (by Marquise)," n.d. https://documents.pub/document/imf-interoperable-master-format-business-situation.html.

"Interoperable Master Format (IMF) Specification," February 19, 2011. https://wikileaks.org/sony/docs/05/docs/IMF/IMF_Specification_V1.0.pdf.

"Is 'Visible Fictions' Still Relevant? By John Ellis," October 19, 2016. https://www.rug.nl/research/icog/research/conferences/research-colloquium-film-studies-fall-2016/visible-fictions?lang=en.

Isson, Jean-Paul, and Minha Hwang. *Unstructured Data Analytics*. Somerset: John Wiley & Sons, Incorporated, 2018.

Jackson, Dom. "The Interoperable Mastering Format," May 2017. https://www.smpte.org/sites/default/files/section-files/BBTB2017-W04%20Dominic%20Johnson%20-%20IMF.pdf.

Jarrett, George. " Interoperable Master Format for TV: The Solution to Versioning Woes?," May 2017. https://www.ibc.org/manage/imf-for-tv-the-solution-to-versioning-woes/1975.article.

"Library of Congress Formats Statement-2020," 2019. https://www.loc.gov/preservation/resources/rfs/RFS%202019-2020.pdf.

"Library of Congress Recommended Formats Statement 2019-2020," 2019. https://www.loc.gov/preservation/resources/rfs/.

Lucia, Cynthia, and Rahul Hamid. *Cineaste on Film Criticism, Programming, and Preservation in the New Millennium*. Austin: University of Texas Press, 2017.

"MEDIA COMPOSER 2019," May 2019. https://connect.avid.com/rs/149-WFZ-676/images/avid-media-composer-2019-tech-paper-final.pdf.

"MXF Archive and Preservation Registered Disclosure," December 2018. http://www.digitizationguidelines.gov/guidelines/rdd48-2018_published.pdf.

"MXF— the Material EXchange Format," July 2002. https://tech.ebu.ch/docs/techreview/trev_291-devlin.pdf.

O'Falt, Chris. "Ang Lee Deserves More Respect For the Ambitious Digital Experiment of 'Gemini Man,'" October 10, 2019. https://www.indiewire.com/2019/10/gemini-man-ang-lee-interview-promise-digital-hfr-3d-cinema-1202180316/.

"Open Source." Accessed December 2, 2019. https://www.imfug.com/open-source/.

Press, Gil. "IoT Mid-Year Update From IDC And Other Research Firms," August 5, 2016. https://www.forbes.com/sites/gilpress/2016/08/05/iot-mid-year-update-from-idc-and-other-research-firms/#7e68efec55c5.

Provenzi, Edoardo. *Color Image Processing*. Basel; Beijing ; Wuhan ; Barcelona ; Belgrade: MDPI, 2018.

"Quality Control for File-Based Workflows," October 24, 2016. https://www.telestream.net/pdfs/general/Quality-Control-for-File-Based-Video-Workflows-25W608940.pdf.

Ravell, Dean. "Blurring the Lines between Film and TV," November 3, 2015. http://www.btglifestyle.com/blog/2015/11/03/blurring-the-lines-between-film-and-tv/.

"Recommendation ITU-R BT.2077-2," June 2017. https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2077-2-201706-I!!PDF-E.pdf.

Roettgers, Janker. "Cord Cutting Will Accelerate in 2019, Skinny Bundles Poised to Fail (Report)," April 22, 2019. https://variety.com/2019/digital/news/2019-cord-cutting-data-1203194387/.

Roettgers, Janko. "Inside Intel Studios, the World's Largest Volumetric Capture Stage," October 7, 2019. https://variety.com/2019/digital/features/intel-studios-volumetric-capture-holograms-ar-vr-1203358126/.

Sepulveda, Elsa. "La CST-RT021 Aboutit à La Publication Du Standard « Cinema Mezzanine » à La SMPTE," December 4, 2016. http://www.mediakwest.com/cinema/item/la-cst-rt021-aboutit-a-la-publication-du-standard-cinema-mezzanine-a-la-smpte.html.

"SMPTE Standards Ecosystem," 2015. https://www.smpte.org/sites/default/files/images/1664_SMPTE_Wallchart_v5c.pdf.

"SMPTE TSP-2121 Application DPP ." Accessed December 2, 2019. https://www.smpte.org/technical-specifications/tsp2121-app-dpp?action.

"STANDARDS REPORT 2019," March 2019. https://www.smpte.org/sites/default/files/Final%20Outcome%20Report-WhitePlains-Mar%202019.pdf.

"Technology Fact Sheet: IMF," September 2018. https://tech.ebu.ch/docs/factsheets/IMF_v2.0.pdf.

"THE EVOLUTION OF MEDIA CREATION." MovieLabs, 2019. https://movielabs.com/production-technology/.

"The Interoperable Format (IMF) in Film Preservation," October 2019. https://search.proquest.com/openview/5ed0858b9cf37009acd7dc43211c1198/1?pq-origsite=gscholar&cbl=29010.

"THE NIELSEN TOTAL REPORT 2018," 2019. https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/q3-2018-total-audience-report.pdf.

"TV KEEPS US ENTERTAINED OUT-OF-HOME," May 15, 2019. https://www.nielsen.com/us/en/insights/report/2019/tv-keeps-us-entertained-out-of-home/.

"Understanding High," November 2015. https://www.smpte.org/sites/default/files/users/user29811/wp-uhd-update-v3.pdf.

Weiss, S. Merrill. "AXF-Archive EXchange Format: Interchange and Interoperability for Operational Storage and Long-Term Preservation." SMPTE Motion Imaging Journal 123. SMPTE, January 2014. https://doi.org/10.5594/j18371XY.

Wells, Nick, Oliver Morgan, Jim Wilkinson, and Bruce Devlin. *The MXF Book*. Oxford: Routledge Ltd, 2006.

"What Is Happening with IMF? The File Format, Not the Int. Monetary Fund by Annie Chang," 2018. https://www.youtube.com/watch?v=X02BdknxIxY.

"Workshop on Preserving Movies with IMF," November 15, 2019. https://www.youtube.com/watch?v=7oyJFhgIrzw.

Zhu, Kevin. “Internet-Based Distribution of Digital Videos: The Economic Impacts of Digitization on the Motion Picture Industry,” Electronic Markets 11, 11, no. 4 (December 1, 2001): 273–80. https://doi.org/10.1080/101967801753405562.