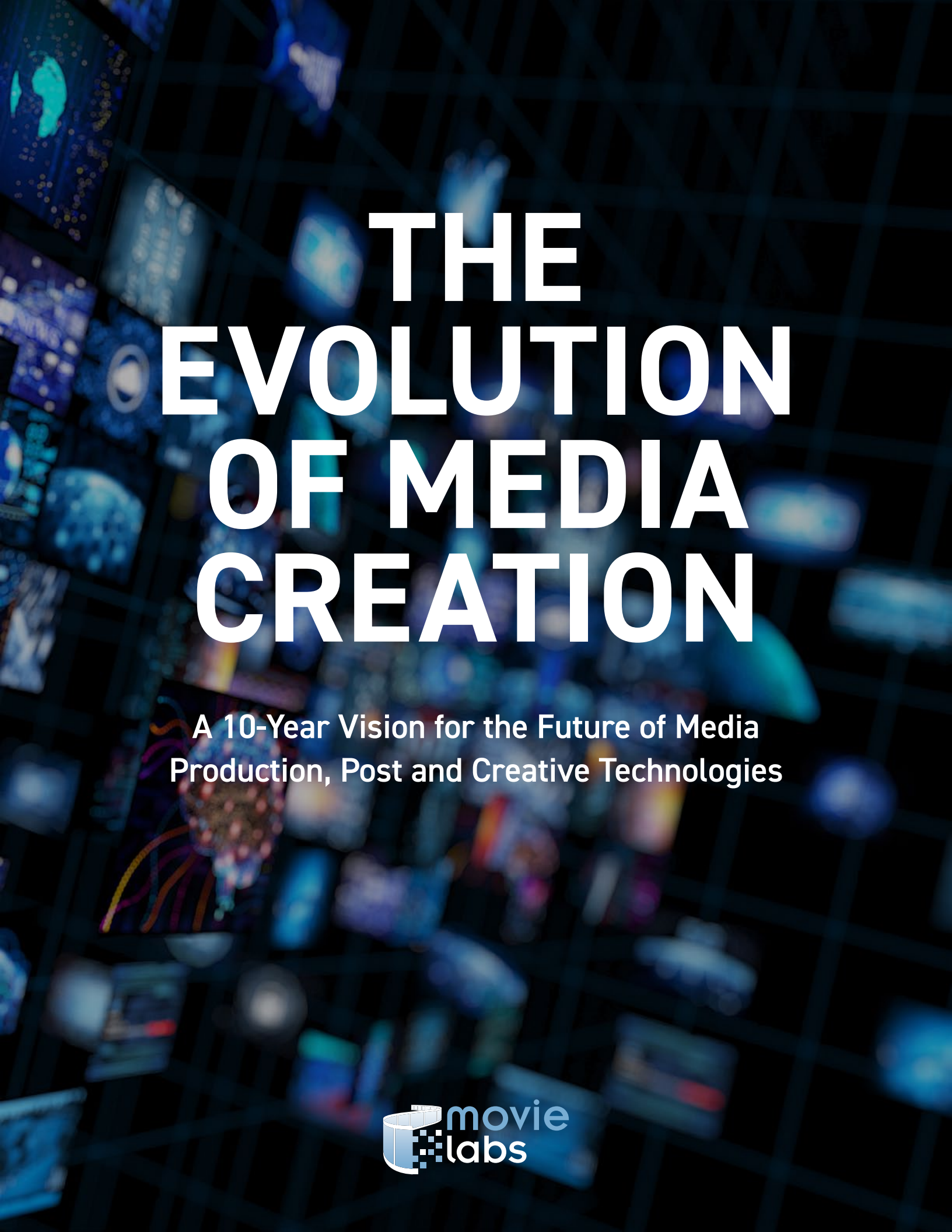


THE EVOLUTION OF MEDIA CREATION

A 10-Year Vision for the Future of Media Production, Post and Creative Technologies

movie labs

Motion Picture Laboratories, Inc. (MovieLabs) is a nonprofit technology research lab jointly run by Paramount Pictures, Sony Pictures Entertainment, Universal Studios, Walt Disney Pictures and Television and Warner Bros. Entertainment.

MovieLabs enables member studios to work together to evaluate new technologies and helps the industry develop next-generation content experiences for consumers, reduce costs, enhance security and improve workflows through advanced technologies.

CONTENTS

EXECUTIVE SUMMARY

MEDIA CREATION BY 2030

The media industry has seen tremendous technological advances and achievements in storytelling over the last 10 years. Still, major challenges and inherent inefficiencies in production and distribution workflows threaten to limit future ability to continue innovation, reduce production time and expand output. Historically, storytellers and creatives have been the beneficiaries of new technological innovations. Within this document, MovieLabs and its member studios have laid out a bold vision for the future state of technology in filmmaking during the next 10 years, with a call to action for the industry to collaborate by appropriate means to achieve shared goals and continue to empower future storytellers and the creative community. We describe future technological advances that will enable seismic changes in media workflows with one objective in mind – *to empower storytellers to tell more amazing stories while delivering at a speed and efficiency not possible today.*

Creatives are constantly challenged to achieve more with less time and stretched budgets. Within our 2030 Vision, we have included ways to use emerging technologies to provide more "time" to filmmakers to make the creative choices they want to make and to be able to iterate more times to reach their ambitions for every production. That does not mean every technology is applicable to every title, and filmmakers will independently decide what to take from this paper and how to apply it to their individual productions. HDR is a similar use case – a new technique and workflow that filmmakers can choose to use to create dynamic realism or deliberately dial back if they want a muted or flat look. Not every filmmaker will want to utilize every new technology we describe here, but they will have more options and tools from which to choose.

There are 10 foundational principles in this document, each an element of the 2030 world we envision, and each accompanied by an overview, examples and broader implications

for the future. By publicly laying out these principles and the supporting documentation, we hope to lead an industry dialogue with relevant stakeholders and fuel innovation to encourage companies and organizations to help deliver on these ideas for the benefit of the ecosystem. At a high level, these principles envision a future in which

1. All assets are created or ingested straight into the cloud and do not need to be moved.
2. Applications come to the media.
3. Propagation and distribution of assets is a "publish" function.
4. Archives are deep libraries with access policies matching speed, availability and security to the economics of the cloud.
5. Preservation of digital assets includes the future means to access and edit them.
6. Every individual on a project is identified and verified, and their access permissions are efficiently and consistently managed.
7. All media creation happens in a highly secure environment that adapts rapidly to changing threats.
8. Individual media elements are referenced, accessed, tracked and interrelated using a universal linking system.
9. Media workflows are non-destructive and dynamically created using common interfaces, underlying data formats and metadata.
10. Workflows are designed around real-time iteration and feedback.

Many of these trends have already started and may be delivered earlier in our 2030 time horizon, but we are extrapolating from edge cases today and looking for how to enable these future use cases to be available for all productions, regardless of budget. By laying out this comprehensive view of the future, we hope to avoid the fragmented and piecemeal approach to the digital migration that largely resulted in a change of storage medium but no major improvements or efficiency gains in workflow. By approaching innovations like cloud-based workflows in a systematic and intentional way, the industry can enable changes that will improve the creative process for all participants.

This white paper further explains the principles outlined above and articulates practical use cases that demonstrate how the principles will come together to promote benefits for all. MovieLabs plans to publish additional documentation to apply these principles in real-world scenarios.

The goal of this white paper is to describe a vision for the future of media creation that enables new content experiences limited only by our imaginations. It aims to start a broad technology conversation among industry stakeholders to create efficiencies, expand output and save creatives time to do what they do best. The ultimate improvements may result from a variety of means, including industry best practices and standards and/or innovation at individual companies. Given the 10-year event horizon, it is premature at this point to opine on the means to implement all improvements. This will evolve over time during the course of discussions with relevant stakeholders. Importantly, during this entire process, each company involved will make its own completely independent business decisions. Each will unilaterally determine how it plans to innovate and, most importantly, with whom it plans to do business. For legal reasons, no party involved in the process will discuss with whom they plan to do business or on what terms.

SECTION 1
INTRODUCTION

Over a 10-year time horizon, many things in our fast-changing world will evolve – new technologies will be created, new processes defined, new partnerships formed – but in the year 2030, we can be sure that people will still enjoy watching audiovisual content. We can debate the form of that experience – it could still be enjoyed on a 16 x 9 screen, or maybe inside a headset, or projected on the environment surrounding the viewer. But one thing will be certain: audiences will still enjoy stories well told with technology that stimulates the senses and creates magic.

With this 10-Year Vision paper, MovieLabs draws a picture of that future, describing how media experiences will be created and prepared for distribution. But this is not an academic exercise; this future vision becomes a north star to guide the next 10 years of innovation. Articulating what the studios collectively want to enable technologically will help identify gaps in current technology, standards and processes and guide creation of a roadmap for how the industry will innovate to deliver on that vision. This paper should be read by technical leaders at vendors, software and hardware producers, and other studio suppliers to understand where studios want to go in enabling new technical workflows.

> **"We describe a 10-year horizon to ensure that we can move technologies from early experimentation and prototyping through scale phases and into maturity.**

We should be clear that the various technical innovations we expect are all developing at different time scales and many are available in early forms today. We describe a 10-year horizon to ensure that we can move technologies from early experimentation and prototyping through scale phases and into maturity. Once mature, these technologies can be made available as standardized processes and workflows for creators of productions on

all scales on movies and TV shows. As we move through the next decade, we fully expect to adapt our Vision to accommodate new technologies that may accelerate beyond our expectations, and to threats or opportunities that may emerge.

This document and the underlying principles were developed by MovieLabs with contributions from technology leaders at major studios. We extrapolated current technology trends and expectations for future growth and innovation. Our scope focuses on scripted, non-interactive entertainment TV and movie content and is therefore not aimed toward video games, live events, game shows and other non-scripted experiences that have different content creation requirements and workflows – although we recognize that many of the key principles carry over into those use cases.

Today's scripted entertainment is characterized by five distinct workflow processes that have evolved over the last 100 years. It's worthwhile to benchmark these processes now because they will very likely change fundamentally over the next 10 years in ways that they have not previously.

Preparation and R&D – Pre-greenlight activities including developing a script, packaging talent, budgeting production scenarios and developing visual presentations required to pitch and greenlight the project. It also includes R&D of new technologies that might be used in the project.

Preproduction – Covers the steps after greenlighting involved in defining detailed plans and processes for production. It includes virtual production and previsualization ("previz"), which are used to plan more efficient principal photography and ensure the seamless combination of physically and digitally produced elements.

Production – The steps in capturing and creating content on set, on location, in animation, in VFX, etc. It includes lights, cameras, sets, talent, grips, green screens and huge media files.

Postproduction and Mastering – Often the lengthiest part of the creation process, this includes steps such as editing, adding visual effects (VFX), mixing/editing audio, color grading and creating dozens of international masters.

Distribution – Preparing and delivering numerous variants of the content to the owner's distribution partners for onward delivery to consumers. Delivery includes theatrical distribution, physical media (optical disc), pay-TV services, broadcasters and over-the-top (OTT) internet services.

In 10 years' time, entertainment productions could be produced in very different ways and today's workflow steps could be rearranged or inverted dramatically.

WHAT COULD MEDIA CREATION LOOK LIKE IN 2030?

By 2030, new technologies, especially real-time photorealistic graphics, will create a more iterative creative process where the phases of early production merge together into

1. **Pre-Photography** – Script development; production design, including setting up much of the visual FX assets in full fidelity; CG scenes; previz of some or all of the scenes; and creation of everything except the live action, with an ongoing process of R&D and refinement of tools and workflows.
2. **Principal Photography** – Similar to today, but with the option of physical actors interacting in real time with digital assets and virtual characters. Photography on some shows will expand to include volumetric capture of entire scenes, use of light field cameras or capture of multiple simultaneous camera angles.
3. **Final Postproduction and Mastering** – Color grading, final sound, VFX finishing and editing.

Starting at the beginning of a production, major movies will be previsualized with photorealistic real-time engine (RTE) pipelines with such quality that a director or cinematographer can make lighting, shooting, performance and even editing decisions in an environment that seamlessly blends the physical and digital worlds. These changes unlock many more real-time iteration opportunities for creative decision-making at any point before, during, and after production. The artist will have flexibility to try new ideas and experiment throughout the creative process, seeing her/his work in close to final render quality generated by a real-time ray-tracing engine.

We can also envision principal photography where directors are freed from the constraints of limited sets and a select number of cameras, to a stage filled with hundreds of small cameras and sensors capturing entire light fields, scenes and performances simultaneously from every angle. These "volumetric capture stages" have emerged in the past few years for immersive media projects and experimentation, but have not yet become commonplace in studio productions. However, in the future, more innovative and elaborate setups that capture whole light and space volumes will be more prevalent. A new step in post-production would enable directors, DPs and editors to select which of these many camera angles should be presented to the audience.

We can expect that media creation workflows will be cloud based with every file (from first script, to camera captures, VFX assets and audio tracks) stored in the cloud. Applications that process the media will be cloud native, meaning the application directly communicates with cloud storage and runs on a virtual workstation hosted in the cloud rather than a workstation sitting beside the users. These cloud-based workflows can be inherently more secure than on-premise workflows since cloud-native production offers the opportunity to use a security architecture tailored for cloud workflows. Applications running on virtual workstations eliminate the need to duplicate and move files between different vendors, store assets on vendor sites and permit the content creators to maintain sovereignty over their assets.

“We can expect that media creation workflows will be cloud based with every file (from first script, to camera captures, VFX assets and audio tracks) stored in the cloud.

These changes in content creation workflows can also be expected to impact consumer delivery and experiences of that media. By 2030, we expect the media elements that comprise finished movies and shows to increasingly remain separate as they are prepared and distributed, then be bonded together into consumer experiences by linked metadata. We can expect future consumers to enjoy media experiences that adapt the playback experience to them by rendering out a uniquely optimized version of the content package for the specific display (brightness, color, contrast, resolution), audio (speakers performance, proximity, placement) and environment (ambient light and reflections, noise levels) that the consumer is in. By using edge computation and rendering of targeted media elements, the industry can avoid creating thousands of delivery masters for every discrete consumer device, bandwidth and playback capability, and instead deliver uniquely what is best for that consumer at that time.

With audio experiences, we also expect innovation. In the past 10 years, we have moved from channel-based audio (5.1, 7.1) to object-based audio, in which individual audio objects can be placed and tracked in 3D space around the audience. During the next 10 years, we expect these trends to continue to include potentially millions of audio objects with a granularity down to audio for individual particles. This flexibility will require new AI-based tools to handle the complex modeling and interactions of objects. The result will be rich sound fields calculated by the interaction of sound elements on the screen as they are realistically

rendered out based on the echo, reverberation and distortion of their location and their interaction with other sounds in the field. This level of fidelity is similar to the particle physics used in today's game engines and the ray-tracing used to computationally calculate the individual light rays in a scene and how they move around, interacting, refracting and reflecting across objects in the scene. In fact, the audio and video technologies may combine to calculate both light rays and sound fields simultaneously based on specific environments. A new wave of creative tools will need to be developed to allow the production team to then adapt and manipulate this calculated reality to tell the stories they want to tell.

We can also expect many mundane or repetitive production tasks to be accomplished through artificial intelligence–enhanced tools, automation, bots and processes. While it's hard to predict every possible use case for such tools, we expect advances to speed creatives' work by animating background VFX characters, doing first-pass color grades with various looks for directors to pick from, and automating rotoscoping, background removal, and "world-building" that generates realistic 3D environments for directors to walk through and edit. Our objective with this vision is to enable a world where AI-based improvements can be efficiently developed and deployed at scale.

"We can also expect many mundane or repetitive production tasks to be accomplished through artificial intelligence–enhanced tools, automation, bots and processes.

While not all titles will benefit from all of these new capabilities by 2030, these advances will be increasingly accessible to creative teams of any scope, budget or schedule. As always, there will be financial considerations, tightening schedules and high production expectations that are common to all forms of content. The tools, processes and technologies that enable a high-budget movie to do amazing things today will eventually trickle down and enable lower cost productions to benefit from the same innovations in the future. We expect our 2030 future Vision to apply to everything from blockbuster movie franchises down through episodic productions, documentaries, independent films and even student films.

Today, some of these developments may seem like fantasy, but examples and early prototypes of all these technologies are already appearing. To unleash their full potential and enable their value to be commonplace by 2030, the industry stakeholders (studios, storytellers, independent production companies, post-production and VFX companies,

application/technology providers and infrastructure providers) need to start working on these innovations now. They must establish new processes, procedures and standards. There is much to be done to create these new media workflows and realize the benefits of reduced production costs, shortened production schedules and productions freed from the constraints of today's workflows.

SECTION 2

CHALLENGES IN REALIZING THE CREATIVE VISION TODAY

When the movie industry moved from videotape-based workflows to file-based workflows, many of the processes did not change; only the storage media did. We moved from duplicating and shipping videotapes to duplicating and shipping data tapes/hard drives, from inserting edits on tape to editing digital files, and from broadcasting in modulation to broadcasting in bits. As a production moves through its workflows and between the storytellers, executives and vendors, there is considerable duplication and movement of files, all of which is time consuming, difficult to track, prone to human error and subject to potential charges for egress and data movement. Furthermore, every time a file is moved, it runs the risk of being misplaced, leaked or corrupted or losing its vital metadata. Precious time is wasted recovering from such problems. As these files continue to increase in size and complexity, these problems are exacerbated. Our Vision needs to find a way of simplifying these steps while securing the process and maintaining the flexibility inherent in globally distributed workflows and workforces.

Modern digital creation workflows have struggled to keep pace with collapsing timelines for delivery and exploding VFX shot counts. Traditionally, network TV episodes were delivered weekly and produced on a staggered production timeline. At any moment, different episodes were at different stages and the post-production demands were consistent. However, the model for OTT and some cable distributors to deliver whole seasons of TV series at the same time has disrupted this cycle. For example, post-production may be postponed until principal photography is complete for all episodes, thereby creating a spike in demand for post-production services as 13 or more episodes appear at once, all needing color correction, sound mixing, finishing, dubbing, QC and encoding simultaneously. This results in a drought or flood scenario for vendors. The industry needs solutions that enable more elastic post-production resources, most likely through leveraging global talent availability, perhaps with AI-based automation to accelerate the processes.

Increased demand for content will require reuse of production assets. With the massive growth in scripted content and popularity of franchises, brands and universes, production elements are used in multiple features, across marketing material (e.g., multiple campaigns, press kits, and co-marketing material), ancillary material (e.g., bonus features, supplemental material, and XR experiences) and related works (e.g., video games and animated spin-offs). Each new deliverable requires a creative team to create, review and approve, often concurrently with the feature's production and post-production. While ancillary material could hypothetically reuse CGI and other VFX assets to create other media types (game engines, XR experiences), reuse is currently not easy. Often, assets are recreated, draining resources and contributing to production delays. Late supplemental media lowers the impact of the title and franchise. A robust platform to store, find, share and archive production assets could enable reuse in sequels and spin-offs such that these assets are not recreated for each title.

Movie and TV productions are faced with significant security challenges today. The bigger the audience, the more likely that someone will try to steal the product. In common with all cybersecurity, this is an ever-escalating battle of technology, with defenders needing 100% success and attackers needing just one way in. Cybercriminals are making good use of cloud technology, and many of the tools they need are available on the dark web as SaaS services today, with the threat of more powerful tools like quantum computing coming. Managing cybersecurity is a complex task, and the aspects of cybersecurity that are unique to our industry make it even more complex. The production environment is fluid. Multiple vendors and sometimes sub-vendors are used, each with different security capabilities. As cloud workflows take over from traditional workflows, traditional perimeter security solutions are not well suited and we will need a new cybersecurity approach that is designed expressly for dynamic cloud infrastructure, shared compute models and software-defined workflows.

> **"In common with all cybersecurity, this is an ever-escalating battle of technology, with defenders needing 100% success and attackers needing just one way in.**

Identity management, the basis of access controls, is becoming increasingly difficult to implement without getting in the way of users. Video production employs a large number of people in development, management, crew, marketing, VFX, post and so on. These people

are employed by studios, production companies, agencies and vendors. Many are individual contractors who could be working on multiple productions simultaneously. It becomes difficult to track everyone, and even harder to have effective security management. One studio IT team estimates that an individual working on multiple productions across different studios can have up to 50 different applications and systems to access, each with its own minimum password requirements, multi-factor authentication methods and expiration schedules. This creates a nightmare of usernames and passwords for team members to manage – often by reusing them or writing them down, thus negating the gains of securing access. As workflows become more distributed, effective identity management becomes even more important. The industry must find a way to simplify identity management and do so in a way that does not frustrate users.

Managing the media elements used throughout productions has become an enormous burden – a typical tentpole movie can include hundreds of thousands of discrete media elements, including CGI assets, audio stems (sound effects, music, dialogue tracks), lookup tables (LUTs), composited layers of video and physical assets such as props, some of which have both physical and CGI manifestations. Management of these assets spans multiple systems; sometimes, they are even still managed using "pencil and paper" systems, which are archaic by today's standards and for most other industries that have benefited from huge gains in productivity by automating their production systems. In addition, each of these media elements may have a specific creator with specific contractual license terms that constrain usage. When the asset becomes part of the final content, the ability to identify where an asset was used and the rights associated with it may be lost. Usage and licensing rights are often tracked in contract documents, guild license terms or spreadsheets, and often in a format that does not facilitate tracking compliance and other obligations. We have an opportunity at this time to move teams and processes into the digital age using modern technologies and enable tracking and linking of our complex asset and workflow systems.

Despite advances in computing technology, rendering of complex VFX-heavy scenes in productions is still not possible in real time with full ray-tracing. Some scenes require hours or days to render a single frame, even using massive render farms filled with parallelized and powerful servers running billions of complex ray-tracing calculations. These "offline" renders constitute an enormous cost in modern filmmaking in money and time. The time required to render limits productions to a multi-day loop (render, receive feedback, iterate, render, receive feedback, iterate, etc.). With up to 30% of the budget of a major Hollywood movie going to render costs, any additional efficiencies can have a huge impact on the profitability of titles. While the number of render passes can be minimized, certain

rigging, lighting, animation and simulation decisions can only be made when a scene has been rendered at final output quality with ray-traced rendering. Ultimately, productions will need to migrate to a world where these renders can be achieved with photorealistic quality in real time for rapid iteration cycles and allow creatives more time for their work.

> **“Continuing to scale the content output from our industry requires dealing with some of the inefficiencies that are inherent in embracing unique creative workflows.**

Finally, the industry takes great pride in the fact that the creative process is a culmination of experiences and workflows and each production is a unique combination of talents, processes and tools. Creatively, every production should be a “snowflake.” However, from a systems perspective, “every production is a snowflake” creates a nightmare of complexity for studios managing and finding files, configuring tools and authorizing processes. Continuing to scale the content output from our industry requires dealing with some of the inefficiencies that are inherent in embracing unique creative workflows.

The MovieLabs 2030 Vision foresees an industry that collaboratively addresses these challenges, increasing efficiency, quickening production cycles, lowering costs, improving security and increasing profitability while creating opportunities for storytellers to realize their vision and delighting audiences.

SECTION 3

THE 10 PRINCIPLES FOR THE FUTURE OF MEDIA CREATION

There are 10 fundamental principles in the MovieLabs Vision, grouped into three sections. Each will be explained here in greater detail. Many of these principles are interrelated and become a mesh when enacted together. In future outreach, MovieLabs expects to work with industry stakeholders to demonstrate the benefits of applying these principles to practical applications. Within this section, each principle is described in three parts:

1. an overview,

2. practical examples of its application and

3. its broader implications for the industry.

SECTION 3.1
A NEW CLOUD FOUNDATION

The first group of principles deals with a **New Cloud Foundation** for the industry. It is important at this point that we define what we mean by *cloud* as it can have numerous interpretations. We define cloud here as not just the hyperscale cloud providers[1] known today, but as any internet-accessible storage platform that can be used as a common space for collaboration and exchange of data. This may indeed be provided by a hyperscale cloud service, a niche cloud provider serving specific use cases, a corporate data center with firewall access or even a near-set storage system used as a staging point to a larger cloud system. The fundamental point is that the nearly unlimited storage and compute, the pay-as-you-go (*opex* instead of *capex*) business model, and the workflow benefits of a "single source of truth" will make the cloud an integral part of our industry in the future and remove the duplication of assets seen today.

PRINCIPLE 1: ALL ASSETS ARE CREATED OR INGESTED STRAIGHT INTO THE CLOUD AND DO NOT NEED TO BE MOVED

OVERVIEW

In our vision of the 2030 creation process, all assets, from the first script to every captured file, every computer-generated asset and all associated metadata, will be stored immediately upon creation in the cloud.

Acquisition devices (cameras, microphones, sensors, script supervisor systems) could directly connect to the cloud, transferring assets seamlessly. These envisioned "cloud-native" acquisition devices would send encrypted, uncompressed files, and simultaneously

[1] We use the term hyperscale cloud providers, but they are often referred to as "public" cloud providers; however, there is nothing inherently "public" in their offering and the word may get in the way of understanding the role they fulfill.

a proxy, straight to cloud storage. That process would change production workflows at a fundamental level.

It is probably worth clarifying that the definition of *cloud* could extend to on-set uses. Many cloud operators offer small footprint local instances of their storage and compute which could be used on set or on location. By designing workflows using microservices operating in containers, they can work consistently regardless of whether they are operating on local infrastructure or thousands of miles away in a large data center. This local instantiation of the cloud can handle services, like fast playback of video for review, for IP acceleration and control of upload traffic to the main cloud (especially in areas where bandwidth is unpredictable or unreliable).

EXAMPLES

Currently production "dailies" are available next-day for directors, producers and executives to review. But if capture devices can stream directly from the set, even in remote locations, a number of processes can start immediately—dailies become "immediates"—offering live and remote viewing experiences that could be viewed on a desktop in real time without the need to travel to stages or remote shoots.

Editors and colorists could do a first-pass edit or color grade during photography from anywhere in the world and quickly provide feedback to directors on-set before they move on to shooting or setting up the next shot or location.

Uncompressed camera files could be placed in the cloud, where they are available to VFX providers who need the original plate. Later, those files are used for final conforming and compositing. Ultimately, they become part of the archive record. Meanwhile, proxy files, whose smaller size makes them more practical to use for many production processes, can be made available in the cloud, allowing those processes to begin immediately after capture.

The cloud files represent a "single source of truth" for reference purposes and enable everyone in production to see and understand the latest version of an asset, which helps with version control issues. An audit or changelog can track how and by whom that asset has changed over time.

IMPLICATIONS

Sufficient bandwidth to enable camera-to-cloud capture is an issue in 2019, especially in remote locations, and there will continue to be a balance going forward between the ever-increasing size of files and the speed of compression innovation and expansion in wired/wireless access technology. And yet the benefits of having all assets stored in the cloud from creation means the studios will continue to push the boundaries of cloud-ingestion technologies.

> **"For this Principle to be realized, the industry needs hardware, software and cloud vendors to work together to design cloud-integrated systems that can securely create, encrypt, validate and store captured assets in cloud object storage over advanced data communications networks.**

For this principle to be realized, the industry needs hardware, software and cloud vendors to work together to design cloud-integrated systems that can securely create, encrypt, validate and store captured assets in cloud object storage over advanced data communications networks.

Although ubiquitous connectivity (wired and wireless) continues to grow, it will be strained to keep up with the enormous on-set data requirement. Forecasts indicate that 5G will allow speeds of up to 1.4 Gbit/s and Wi-Fi will continue to increase in speed with the latest 802.11ax specification at up to 11 Gbit/s – both enough for a 1080p video feed, but insufficient on their own for the expected increase in file sizes. We can expect the volume of data captured to continue to increase with increases in resolution (4k to 8k and beyond) and higher frame rate capture. New on-set capture technologies, including volumetric capture, point clouds and light fields, could see an even more dramatic growth in the volume of data coming from set. For example, a light field generated with an array of ninety-six 2K cameras would generate hundreds of gigabits per second of data, beyond what any foreseeable network can handle. These data would undoubtably need to be compressed or preprocessed in some way before uploading to the cloud.

PRINCIPLE 2: APPLICATIONS COME TO THE MEDIA

OVERVIEW

While today huge digital media files need to be moved between facilities, by 2030 we can envision a high-bandwidth, low-latency, cloud-enabled production world where these files do not move and software tools *come to the content* instead of the other way around. Ultimately, this requires fast cloud storage and processing, which can serve perhaps hundreds or thousands of virtualized workstations where artists' software tools can access assets without the need to move them. The artist will need only a screen, an input device and sufficient internet connectivity to stream a virtual desktop; the heavy lifting of compute processes will be handled in the cloud and revisions to files handled remotely.

While cloud technology might move files within its infrastructure (e.g., temporary caching of specific files at the edge closest to the artist to ensure minimal latency), from the user's perspective, files are not moved. That is, files are moved only for convenience, not necessity.

EXAMPLES

In the future, a globally distributed team of VFX artists can be allocated multiple shots from a major movie and each one assigned a step in the process, as if they were all working locally. Instead of each user copying the master plates and the CGI asset libraries, they would instead be provisioned access to a virtual workstation, already preloaded with their required applications, software licenses and the media files they will need to access. As the compute can be co-located with the storage, there is no egress of the master files; they remain in the cloud, and only streamed desktops leave the cloud.

Likewise, an audio mix could be created, with a mixer receiving streamed access to the proxy video and full resolution audio elements to create a final mix, even though none of those files need to be locally resident. The output of the mixer's work would be a metadata file that describes the final mix, ready for a final render/composite/packaging step.

Most creative processes could be achieved in a similar fashion with a remote desktop connection generating descriptive metadata. The most extreme case will be remote color sessions and DIs since the assets used are typically uncompressed final camera files that

need to be displayed on color-critical monitors at native resolution, in 16-bit depth, with a wide color gamut and without any compression artefacts.

IMPLICATIONS

Currently, creative talent often needs to live and work in cities close to the productions and their media files. If the media is cloud-native and can be streamed anywhere, talent is no longer tied to the production locale, which can also be anywhere in the world. Remote production talent in currently inaccessible markets can be called upon, perhaps for extremely specialized tasks, opening new avenues for creativity and a new pool of talent to address current shortfalls in specialist talent. And the move of applications to the cloud greatly expands the set of devices and locations from which a creative can access content and tools.

> **"If the media is cloud-native and can be streamed anywhere, talent is no longer tied to the production locale, which can also be anywhere in the world.**

Studios can have confidence in their data security as workflows will no longer require external vendors to maintain copies of these files. Issuing and revoking access to assets can be done online by producers quickly and with instant effect. There will be no "shadow" or "rogue" copies of files, alleviating concerns about version control, where different versions of files have different edits. In addition, the vendors need not invest in on-site storage equipment and compute infrastructure; they require only large bandwidth pipes to access the internet. Hopefully, these changes will lower the barriers to entry for new vendors and increase competition for services.

PRINCIPLE 3: PROPAGATION AND DISTRIBUTION OF ASSETS IS A "PUBLISH" FUNCTION

OVERVIEW

Traditionally, propagating assets to the next stage in production or distribution of finished assets involves packaging and delivering files to another party, often using the internet as a distribution path. In our model, however, the files are already resident in the cloud and therefore do not need to be moved. A core tenet of our approach is that anything from

production assets to the final finished content can be shared via the cloud – but the decision to do so rests with the current controller of the asset. During production, it may be the director who decides a shot or cut is ready to be seen by others, or the VFX artist who decides a scene is ready to share with the VFX supervisor. In distribution, it may be the executive who approves a release to consumer distribution. For the system to operate, all users must have the confidence to trust that even though storing files in the cloud means all assets *could* be visible to others, the controller's *action* to actively publish the content down the line is what actually makes it visible and accessible to the recipients. Also, the workflow security system could limit access to specific uses, e.g., viewing, editing, etc., based on which cloud applications are allowed access via this "publishing" workflow.

Content protection systems have been developed over the last 20 years to protect media from internet piracy, but in our future vision, delivery will be predominantly via the internet and even unfinished production masters will be stored in the cloud already. For final delivery to consumers or distributors, the finished production can be made available in the cloud in a distributor staging area and a "dynamic package" created with a manifest containing the media that the distributor is licensed to receive. Future AI applications could even automatically interpret bilateral agreements and prepopulate these manifest files using smart contracts.

EXAMPLES

With content captured directly to the cloud, the director or DP can decide when, to whom and for what purpose (e.g., viewing or editing) the dailies (or what could be called "immediates") are published.

The "VFX pull," a process in which VFX providers traditionally "pull" to their local storage the full resolution plates they need to work on, could be redefined in the future workflows. We may see a "VFX push" of permissions, which would enable the VFX provider's artists and applications to access and work on the master plates stored in the cloud. The plates would not move, but they would be available to that vendor for the length of their agreement.

Content owners could publish all finished video and audio elements to their cloud "staging area" along with a manifest file that references how they all link together. Each distributor can access appropriate parts of the package in accordance with permissions associated with

their contracts. For example, a Spanish pay-TV operator may receive permissions for a high bit rate encode with Spanish audio, an edit that meets the local broadcast regulations and an optional English subtitling track. A digital cinema operator may get permissions for a DCP package and specific trailers optimized for that time/date playback license. Meanwhile, a global premium movie retail store could retrieve multiple regional edits from the consumer HDR grade with all audio options and a package of bonus features and interactive menus.

Content owners could go one step further and provide pre-encoded VBR files, which could be pre-cached on distributors' CDNs, ready for the initial spike in traffic on release day when the media is "released to the web."

By providing dynamic packages of content, the studios can mix and match assets and update and augment elements within delivery packages as other related content is made available. If there was ever an issue with rights clearances, for example, on a music asset in a specific territory, it could be instantly changed by adjusting the manifest file, where an audit log would clearly show which distributors have access to that media.

In addition, an automatic update notification can be sent down the line to recipients in case an item has changed. In the case of production, these could be updated versions of shots or files that need to be automatically rippled down to all users who have access to that variant and need to know it has changed.

IMPLICATIONS

This principle requires producers, executives and content owners to have established a basic level of trust in the cloud and its security protections. We have more work to do, as an industry, to alleviate any lingering concerns about placing unreleased content assets in the publicly accessible cloud and all stakeholders need to work together to ensure any concerns are addressed.

The requirement on the industry will be to work together to extend the work done in IMF, and potentially, other interchange formats such as DCPs and improved descriptive metadata, such as that found in the MovieLabs Digital Distribution Framework (MDDF), will be required to enable a much more efficient global digital supply chain. We may need future packaging formats that can contain real-time engine elements that can be rendered

at home, light field or point cloud 3D asset packages and/or standardized XR formats for delivery of immersive narrative pieces.

PRINCIPLE 4: ARCHIVES ARE DEEP LIBRARIES WITH ACCESS POLICIES MATCHING SPEED, AVAILABILITY AND SECURITY TO THE ECONOMICS OF THE CLOUD

OVERVIEW

Content archives contain intellectual property potentially supporting billions in future revenue streams. Therefore, archiving and preservation of media assets is a vital, although often underappreciated, function.

Creating and storing media assets in the cloud obviates the need to support legacy archiving hardware (e.g., continuously obsoleted tape formats) and standard media formats to preserve files. Those issues will fall to the cloud service providers themselves in the future.

This principle recognizes another key advantage of the cloud for storage: subject to economics and cost viability, the media need not "go anywhere" when it is archived. Currently, archiving can be seen as putting assets in a place where unauthorized people cannot access them and nobody can destroy them. The drawback is that it is difficult to retrieve and repurpose those assets. In the future, a cloud-based archive can be indexed and made readily available to authorized users for monetization, cross-referencing for future productions, remastering and education (of both people and algorithms).

"This Principle recognizes another key advantage of the cloud for storage: subject to economics and cost viability, the media need not "go anywhere" when it is archived.

Accessibility will no longer be about who *physically* has access, but who has been *assigned* access through policy. These policies will vary between content owners based on how they wish the asset to be used. Policy decisions in a cloud future will include factors such as speed of recovery, costs associated with deep asset retrieval and what is kept in fast

"online" storage (e.g., proxies and metadata) versus the deep archived master files.

EXAMPLES

At the end of a production, the content owner could archive the media with the push of a button, assigning archive policies to that media. Behind the scenes, the cloud might demote files to a lower cost and slower performance cloud storage tier, with all assets remaining indefinitely online and available for future productions, remastering or redistribution at a later date.

> **"Future AI/ML bots could crawl massive data depositories looking for reference material to use for inference…**

Future AI/ML bots could crawl massive data depositories looking for reference material to use for inference – for example, highly specialized data for learning the distortion effects of a legacy lens.

New productions wanting access to reference material used on earlier movies or others in the same canon can quickly search and find original concept designs, actual shots or VFX models used on old titles.

Understanding and controlling which assets are kept in which cloud storage tier can be complex now and will likely get more complex as millions of new files are added weekly. However, cloud storage tiers are likely to evolve to become "self-optimizing" by automatically moving less used files to deeper and deeper archive tiers to maximize cost versus access. Machine learning tools in the future could use pattern analysis to predicatively retrieve and preposition deeply archived files so they can be ready for reuse as they are required.

IMPLICATIONS

As archive access will likely be slower and more expensive relative to active projects, a robust proxy that can be easily accessed in fast-access storage tiers will be required, not just for video files, but all elements of the archive, including 3D assets, audio elements and volumetrics. For example, a new specification for a 3D asset proxy could specify a high resolution, high dynamic range rotating turntable video of the asset that shows the asset

from every angle and explodes to include how the elements, including mesh, textures and rigs, were applied to create the final asset.

Before any studios trust a third party to store more than 100 years of archives, they need to get assurances that their content will be secure; retained in bit-perfect form; protected from hostile action, system failures and natural disasters; and accessible indefinitely. This means trusting foundational elements of cloud infrastructure such as the underlying at-rest encryption, the key management system (so content owners may never lose access to their encryption keys) and protection from future digital attacks (rogue actors, viruses or even electromagnetic pulses) that could corrupt the archives.

“Before any studios trust a third party to store more than 100 years of archives, they need to get assurances that their content will be secure; retained in bit-perfect form; protected from hostile action, system failures and natural disasters; and accessible indefinitely.

In addition, there are still issues with cloud economics that mean it can be hard to predict future costs based on current estimations for how often those archived files may be retrieved, re-indexed or moved to faster tier storage, all of which will need to be resolved before mass migrations of data can occur.

PRINCIPLE 5: PRESERVATION OF DIGITAL ASSETS INCLUDES THE FUTURE MEANS TO ACCESS AND EDIT THEM

OVERVIEW

Celluloid/polyester film has a number of enviable traits, not least of which is that it will always be readable by future generations by simply holding it up to a light source. The same is not true of digital storage media created even a short time ago. It might require discontinued devices capable of reading that data. Data might be stored in a file system that is no longer supported. It might use a file format no longer maintained. Or it might

require an application that is now defunct or orphaned by progress. Therefore, as principle 4 addresses access to future media files, this principle addresses the need to ensure that we can continue to open and use those files in the future.

> **"The essence of an archive is storing for perpetuity that which cannot be recreated ever again.**

The essence of an archive is storing for perpetuity that which cannot be recreated ever again. Storage in the cloud (Principle 1) solves many problems of devices and file systems. But it may also be necessary to protect against defunct applications by seeking out interoperable file formats, open standards and potentially open source code to ensure we can continue to open and edit files in the future. Future archiving standards may include not just archiving the assets created by an application, but also archiving the application itself to provide perpetual access, and even archiving the virtual machine that was running the software (with its specific I/O and interface requirements). In such a way, future emulators could "rehydrate" the exact machine and allow a future user to open and edit the media or asset. However, that may still not be enough to upgrade that asset to something that can be used by modern applications. Achieving that goal likely requires industry support for both open file standards and basic open source software to read those files.

EXAMPLES

Camera RAW files are used on some productions as the "master" from principal photography, but those file types are proprietary to each camera manufacturer and need to be DeBayered using proprietary algorithms. Rather than archive just RAW files that may be hard to edit in the future, it may make sense to also archive copies of any proprietary software required to access the files and to use an open file format (e.g., OpenEXR or a yet-to-be-created standardized RAW format) as an archival format for all camera files that can contain the full resolution, dynamic range, frame rate and color space from the originals.

In addition to archiving each element in the workflow, the workflow itself could be archived – including every application, file and metadata used throughout. This sort of comprehensive archive could include too much information to be efficient as an archive, but may be useful for future workflow optimizations, as a way to look for opportunities to improve efficiencies in future productions.

IMPLICATIONS

As innovation occurs around new tools, equipment and techniques that sometimes use customized data and metadata to enable rapid innovation, it will be important to bear in mind the ongoing need to protect vital media assets for future generations of filmmakers, consumers and studios.

SECTION 3.2
SECURITY & ACCESS

Now that we have established a cloud-based storage model for future media, data and metadata, we will address how those assets will be accessed and secured such that the required work can be done with users (creatives, vendors, executives) barely noticing that their work is secured. The next three principles deal with an approach to ensure a seamless security system that is constantly adapting to future threats.

PRINCIPLE 6: EVERY INDIVIDUAL ON A PROJECT IS IDENTIFIED AND VERIFIED AND THEIR ACCESS PERMISSIONS EFFICIENTLY AND CONSISTENTLY MANAGED

OVERVIEW

In our 2030 Vision, access to every asset can be authorized to a specific individual for a specific task and a specific duration. To that end, we need an industry-wide means to identify and validate every industry person who has access to a production asset (a single "Production User ID"). Writers, producers, colorists, studio executives and anyone else involved in the process would have a unique identity that would be used to determine what they can control, access or edit. Very few people need access to every asset on a production, so restricting access to certain assets or types of assets can immediately provide a more secure environment. A colorist may not need access to VFX assets, but does need final composited frames; a dubbing artist may need two weeks' access to the final English master and the script, but does not need access to the final audio stems; and so on. A creative working on an asset may limit access to a small number of specific users, and then increase who has access as the asset progresses toward completion. We envision a system, therefore, in which a user on a project can be identified and authenticated and be assigned a specific task and access to specific files for a specific duration.

The writer or initial production team might initially establish rights and permissions during the pre-photography stage for each user. As studios, agencies, talent and vendors are attached to a project, the list of approved partners grows, including those who can, in turn, delegate rights to others. Likewise, if a particular user is removed from a project, his/her permissions to access secure media files can immediately be removed. Given that the removed user has no local media files (per Principle 1), there is no danger of orphaned or rogue content being accessible on a previous crew member's device.

"As studios, agencies, talent and vendors are attached to a project, the list of approved partners grows, including those who can, in turn, delegate rights to others.

EXAMPLES

A producer is working on two movies for two different studios. When she logs into her web portal to view assets and workflow tools, the platform separates the confidential assets from each movie so there is no danger of assets being shared between studios.

An editor leaves a post-production company to move to a competitor, and much like his corporate email is removed as he walks out the door, his access and permissions to sensitive media files are also immediately withdrawn.

An independent creative is typically provided multiple apps when starting any production to manage access to physical and digital assets and handle timecards, payroll and other operations support. Using a standardized production ID, the creative need only remember one username and password combination to access all the relevant apps. Ultimately usernames/passwords may be phased out, as they will in other industries, in favor of other authentication systems that are more secure and easier to use, such as biometrics.

IMPLICATIONS

The entertainment industry is likely unique in that it releases products after they have been conceived, created, touched, enhanced and financed by an astonishing number of independent parties who collectively contribute to the end result. Enabling individualized and fine-grain access controls to media assets in the way we describe has always been

desired, but never before achieved. There will therefore be resulting work with all of the guilds and the creative community to ensure the new system is well understood, the value of it correctly conveyed and the system ultimately embraced.

A considerable amount of coordinated IT infrastructure and industry agreement is needed to enable this section of the Vision. IT demands are not just at studios but at production companies, vendors, software and tools providers and guilds and their individual members. This coordinated work will deliver payback in industry-wide efficiency gains and a more secure and auditable workflow. Users benefit by having one set of credentials that will follow them between productions. There is also a potential benefit in linking this single-user system to a blockchain to associate these users with their work via a distributed ledger.

This model could allow accurate tracking and auditing of any particular person's involvement in a production. Their production profile (e.g., IMDB or resume) could be verified as correct and automatically updated. Likewise, studios and productions will be able to accurately pay out any fees or required back-end revenue shares with full knowledge of who contributed what to the production. Of course, respect for privacy would be a necessary element of this approach.

PRINCIPLE 7: ALL MEDIA CREATION HAPPENS IN A HIGHLY SECURE ENVIRONMENT THAT ADAPTS RAPIDLY TO CHANGING THREATS

OVERVIEW

By 2030, workflow must be secured using the best cybersecurity technologies that exist at each moment in time. This requires continuously refactoring to stay ahead of ever-emerging malicious tools that exploit vulnerabilities and penetrate systems. Advances in technology, such as quantum computing, may mean that critical parts of the security system occasionally have to be replaced. Our 2030 Vision therefore includes a philosophy of cybersecurity built into the core architecture. Enterprise-centric security solutions that protect infrastructure are insufficient.

The new security architecture must be designed specifically to protect cloud workflows and individual assets rather than the infrastructure the workflows run on. The result will be more secure and less intrusive in the workflow and provide content owners and

creatives with control not only over security, but also over the integrity of the workflow. For example, finer granularity of security will also allow creatives working on a project better control of which individual assets they make available and to whom as those assets become ready for wider review or for the next stage in the workflow.

> **"The new security architecture must be designed specifically to protect cloud workflows and individual assets rather than the infrastructure the workflows run on.**

"Security by design," the approach we are recommending, means designing systems where security is a foundational component of system design. The approach takes malicious practices for granted and makes no assumption as to the trustworthiness of users and services or what an attacker may or may not do.

It is critical that C-suite executives and all stakeholders actively support the security standards through policy, financial backing and internal promotion and by appreciating the enormous damage that can be done to profitability for every stakeholder if they are breached.

The MPAA's Content Security Best Practices currently requires assessment of service providers, such as brick-and-mortar post-production facilities or cloud-based solutions, against a set of industry-accepted security best practices. The guidelines reflect today's approach to cybersecurity in the broader enterprise world and are not yet envisioning entirely cloud- and software-based workflows. In addition, thinking on cybersecurity has evolved beyond perimeter defense (since it often fails to resist skilled attackers) and will continue to evolve over the next 10 years to include principles such as Zero Trust and predictive threat detection. While the practice of performing third-party assessments is not expected to go away in the next decade, there is hope that the validation aspects become less burdensome through the use of advanced cloud security tools and techniques and validation of security components built into the applications and processes that are the building blocks of cloud workflows.

Many of the breaches that are discovered daily across all industries fall into two categories:

1. Problems in the selection, implementation, configuration and user training of security mechanisms.
2. Vulnerabilities in the technology, primarily software, that is used by many enterprises.

Cybersecurity in the 2030 workflow will be built around the principles of a *zero trust network* model, in which nothing inside or outside the organization is automatically trusted and instead everything and anything is verified before a connection or access is permitted. UltraHD offered the industry the opportunity for significant improvement in the way content is protected for consumer delivery, and cloud workflows offer the same generational opportunity to change the way production is protected.

A separate document goes into more depth on the cybersecurity approach that will be needed to underpin this Vision.

EXAMPLE

A future cloud-based security architecture would no longer rely on access control lists – instead, there would be cryptographic protection of assets (e.g., a frame, a shot, an audio stem or whatever the production decides). Where access control lists fail, when an attacker gains greater privileges than they should have, cryptographic protection adds a layer of security that protects assets even when the attacker *can* access them.

> **"This Principle highlights the difference between a facility using cloud resources to augment or replace its infrastructure (a "hybrid cloud"), and *production in the cloud*, when the entire workflow is itself protected regardless of the infrastructure it is running on.**

This principle highlights the difference between a facility using cloud resources to augment or replace its infrastructure (a "hybrid cloud"), and *production in the cloud*, when the entire workflow is itself protected regardless of the infrastructure it is running on. That does not mean each facility would drop its own perimeter security systems, but rather they can be focused on their own IT infrastructure instead of the job of managing content owners' security on their behalf.

IMPLICATIONS

Zero trust networks require engagement by every contributor to the workflow ecosystem: network equipment vendors, cybersecurity tool providers, application providers and service providers, and the organizations that use the systems. One challenge will be to implement zero trust in a way that is frictionless for users. Single sign-on provides users seamless, secure access to the media and tools they need to do their work.

Today, an attacker will typically breach the weakest link in the chain, gain trusted status and then hop from one system to another inside an internal network. Building workflows around zero trust principles prevents someone with access to one system from gaining access to another without reestablishing their authorization. The zero trust principle requires trust, such as user credentials, to be verified at each hop. Fortunately, authentication is a global problem and new standardization initiatives, beyond the entertainment industry, will simplify the implementation and make the new security model transparent to users once they are authenticated.

PRINCIPLE 8: INDIVIDUAL MEDIA ELEMENTS ARE REFERENCED, ACCESSED, TRACKED AND INTERRELATED USING A UNIVERSAL LINKING SYSTEM

OVERVIEW

A fundamental requirement for producing content is the ability to locate everything used in the production, whether it is camera frames, documents, sound files, CGI models or final rendered video. Further, it is essential to understand how those files all *relate* to each other. It is within these relationships and modifications to files that creative processes occur. In future systems, these media components will all be in the cloud (Principle 1), so a cloud-appropriate linking mechanism will be necessary to understand these relationships. Perhaps this would be through a new universal link, much like a URL is used to direct a standard browser to find a webpage even though that browser does not store the physical locations of every website in the world.

> **“It is within these relationships and modifications to files that creative processes occur.**

A linkage system would serve two critical functions:

1. Links can be resolved to a physical location. An application using the link is directed to the appropriate storage and does not rely on it being in the same location where it was last accessed. Objects can move within a cloud or between clouds transparently to the user. The cloud can perform whatever optimizations make sense, such as caching at the edge, without impacting users or systems.
2. An asset could carry the links to other assets associated with it. For example, a camera frame file can be linked to the metadata description of the frame, and the metadata description would be linked back to the camera frame file. A camera frame can also be linked to the dialogue file recorded at the same time and vice versa. Links between objects are bidirectional.

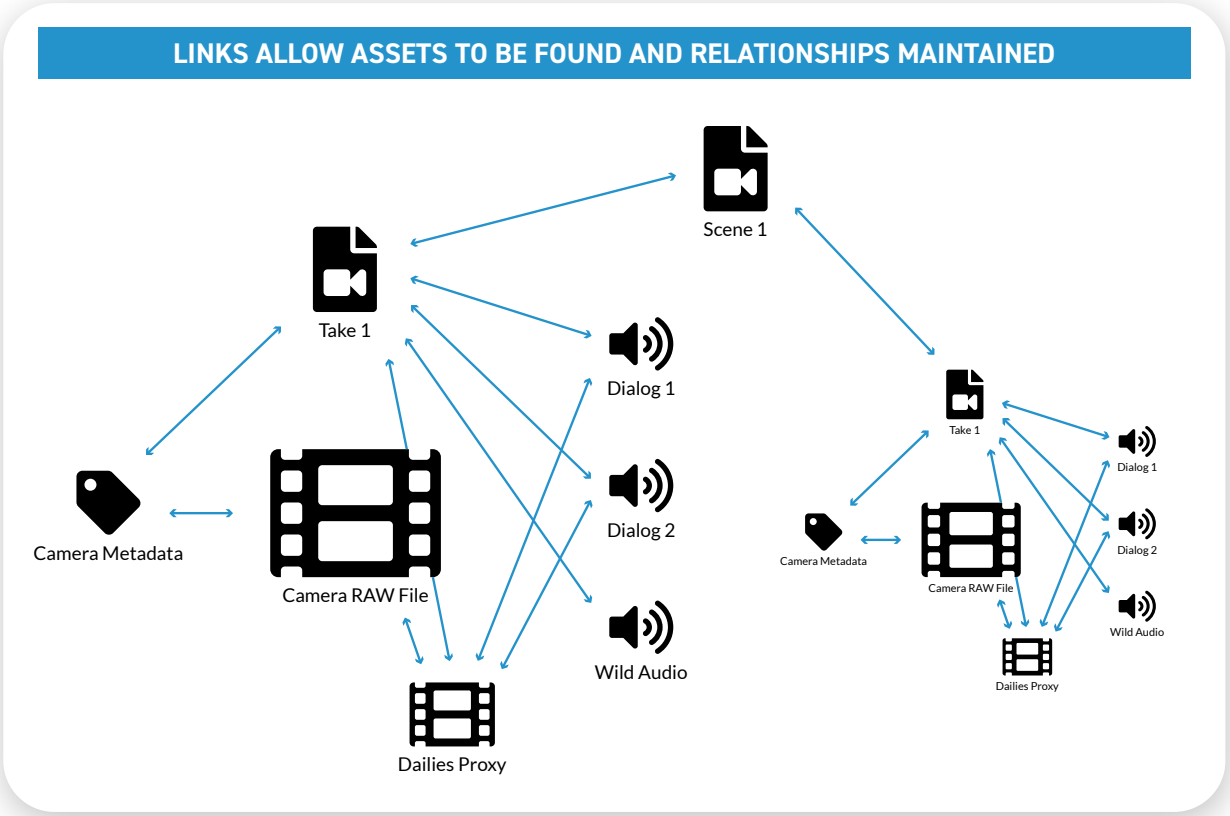


Figure 1: *Unified linking maintains the relationship between assets*

In Figure 1, the scene is linked to the takes and each take is linked to the assets that were created during the take. These links are maintained through production, post and distribution and into the archive. Even years after a production has wrapped, it will be possible to know which files made it to various cuts of the content and retrieve the appropriate files.

Organizing and maintaining files as links allows workflows to be established in which all components remain independent throughout production and final video files are only "flattened" (rendered, composited, mixed) during the very last step of a production. Before that final step, a truly non-destructive workflow can be established whereby creative tools are tracking changes and edits in open standard metadata files (CDLs, EDLs, mix sessions) but not actually modifying the source media, enabling full creative freedom to keep making changes right up until final render/publish.

EXAMPLES

In today's world, a production typically uses a VFX company whose render farm is in its private data center and the applications access the production's files using a file system. Rearranging the files on the file system, or moving them to the studio-managed cloud, currently breaks access, which then must be manually reestablished. With this principle, the files could be moved between storage or cloud providers freely; applications would use the link, which would be seamlessly updated and access would be uninterrupted.

The linking system could also accommodate a "one-to-many" scenario, for example, to enable edge caching of critical files at various locations around the globe, but without each carrying a new location – the link could be the same for each copy of the file and applications could resolve the link to the closest local copy to them.

In current productions, processes occur in physical locations (with cameras, sets, props, actors) and in the digital world (digital characters, simulations, environments), but the two are kept very much separate. By 2030, this link system could bridge the physical and digital worlds, allowing digital assets to appear rendered perfectly in the physical world and physical objects and plates to seamlessly appear in the digital world. The greatest challenge here is to develop an open system of interfaces whereby lighting (location, intensity, direction), 3D camera positions/tracks and 3D locations of characters and objects can be freely passed between physical and digital environments within the linking system.

So, for example, lighting and camera positions can be set in virtual production in 3D space before principal photography. If any changes are made to light or camera positions/tracks during actual photography, those changes are seamlessly transferred back to the digital representation of the scene, keeping the two in lockstep.

IMPLICATIONS

For this to work, applications and devices across the industry need to support the use of links to create and access files, in much the same way filenames are used today. As DNS enabled an explosion of web content and lowered barriers to entry, likewise a linked approach could enable an explosion of simplification across the industry when it is natively supported by partners.

The use of distributed ledgers (such as blockchains) is frequently suggested as a solution to a broad range of problems. Despite the hype, there are genuine use cases in the media and entertainment creation and distribution ecosystem for such blockchain technologies. The innovation we hope to engender in this principle could allow an industry-wide blockchain that can be used to track and link all the components within a finished piece of media with their agreed contract terms, all expressed as a series of smart contracts.

We believe there will be considerable savings compared to today's VFX workflows. Ensuring via linking mechanisms that digital and physical camera and object locations will always match makes it unnecessary to perform processes such as match-move and rotoscope, or to calculate lighting references or camera tracks. These processes are time consuming and labor intensive, so reducing or eliminating them completely will save productions considerable time and money.

SECTION 3.3

SOFTWARE-DEFINED WORKFLOWS

Building on the first eight principles, which provide a common core of cloud infrastructure and a consistent system to track assets and assign projects to real people, all in a secure environment, now we can move on to iterations in the actual workflows that power our industry. The next two principles pertain to software-defined workflows, i.e., dynamic and modifiable connected services that operate on off-the-shelf hardware instead of specialized, custom devices. These applications are delivered in modern companies today as microservices (although 2030 software systems may have evolved still further). Regardless of how they are delivered, our objectives are to encourage new creative tools and workflows that allow more creative expression and faster innovation.

PRINCIPLE 9: MEDIA WORKFLOWS ARE NON-DESTRUCTIVE AND DYNAMICALLY CREATED USING COMMON INTERFACES, UNDERLYING DATA FORMATS AND METADATA

OVERVIEW

As technology advances at a blistering pace, the demands to constantly redesign production pipelines to accommodate new technologies are becoming untenable. Furthermore, dependence on an agglomeration of legacy tools results in a fragile environment, susceptible to failures that can ripple throughout an entire production process. This principle mitigates these issues by establishing standardized building blocks for workflow processes with common data file types, descriptive metadata and interfaces for applications to interface with those systems. By adopting a modular methodology for production pipelines, creatives can quickly construct and adapt workflows from these building blocks. The blocks will each

have their own defined minimum data, metadata, input formats and output formats and will easily communicate with each other using consistent underlying data systems.

We also include the notion of non-destructive workflows; that is, whenever possible, the original asset is maintained in its original state. Production processes layer modifications that are described in metadata files. In that way, the original assets can always be retained and any changes or enhancements regressed back by peeling away layers of modifications.

We can envision an industry *interface layer*, likely in the form of a series of standardized application programming interfaces (APIs)[2]. Workflows would consist of processes exchanging assets data and associated metadata through this interface layer with other

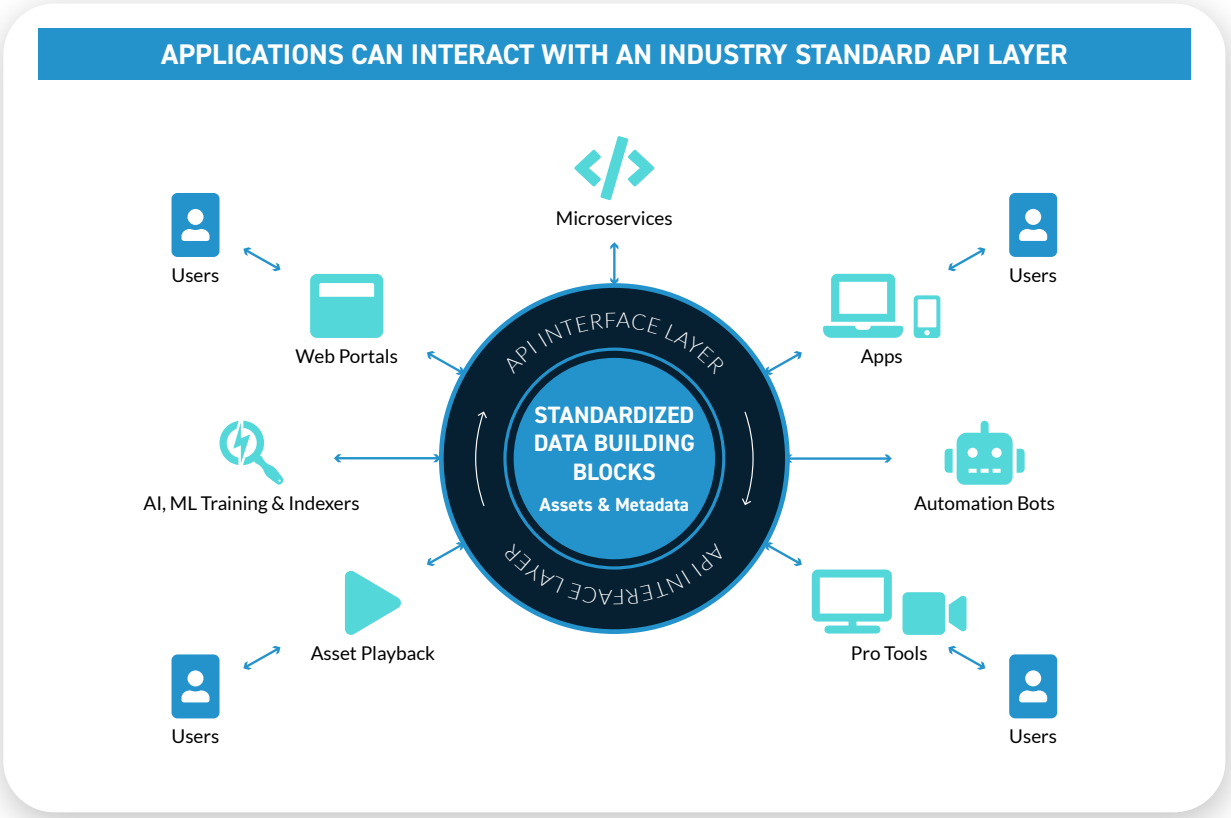


Figure 2: *Standardized building blocks contain agreed common data and metadata for key processes in production. The API interface layer abstracts that information such that any application, tools, portals or service can interface through the API layer with any other tool without needing to know about it in advance.*

[2] Such industry-wide agreed and open APIs can be developed in a secure manner, as has been done by the banking industry with the global BIAN network. See https://bian.org/

processes. This model would support a marketplace-style environment where providers could compete to offer the best components and/or services as modules that plug into a specific workflow/pipeline. Content creators can select, mix and match any of these services to design a workflow or swap them out without having to redesign their pipeline from scratch.

Such a system would drastically shorten the "pipeline development" phase of a production, allowing rapid prototyping facilitated by intrinsic interoperability between the specific technologies with the environment. This would also empower the production team to remain agile with its technology choices. The production could change out any part of the workflow, from camera manufacturer to editing system, without an interruption to the production.

Each major process used in production – from on-set and dailies down to dubbing and mastering – uses distinct file types and metadata as inputs to the creative processes. The resulting work may be altered assets (e.g., a composited image) or altered metadata (e.g., an editorial decision list or EDL). For the interface layer to work most effectively, we will need to describe both the standardized asset files and a minimum amount of metadata. Similarly, there are certain standardized outputs (e.g., data, metadata) for each production subprocess, which are contained in the data storage. Each content creator may have their own additional sets of metadata that they will want to track on a per-project basis, and the system will need to accommodate such ancillary datasets.

Over the years, the list of standardized metadata and data for the building blocks may grow, but for now, there are enough similarities in production processes that an initial set of data and an extensible metadata schema can be developed across studios and productions to start to bring some order to the chaos.

EXAMPLES

A new company is formed that creates a new niche production tool to track color management information from on-set. Instead of creating its own API, the company plugs into the existing interface layer, which now quickly and easily allows any content creator to integrate the tool and immediately begin ingesting information.

A VFX vendor delivers element packages back to the studio for archive. The interface layer

would understand the various elements (3D models, textures, etc.) and extract the metadata from each asset to present to the studio's databases for further data processing. Because this metadata extraction has been standardized, the studio does not need to create its own normalization of the data, as the VFX building blocks already contain typical data fields for each asset.

We cannot hope to predict every nuanced data field that may be required for future productions, but by defining an extensible schema, we may accommodate a new camera technology in the capture building block and allow it to interface with legacy software applications with no prior knowledge of the new technologies or file types.

> **"The standardized building blocks coupled with an industry-wide interface layer enables the best of both worlds: economies of scale from consistency, and freedom and creativity for each production.**

IMPLICATIONS

The standardized building blocks coupled with an industry-wide interface layer enables the best of both worlds: economies of scale from consistency, and freedom and creativity for each production. By creating an industry standard interface layer, we can ensure that any number of web applications, creative tools, bots and other software that understand the underlying files, their structure and associated metadata can be developed.

By creating standards for consistent media and metadata nomenclature, hierarchy and storage interfaces, we can define what is the same for every production (scripts, production notes, video files, audio files etc.) and where they can be found in the cloud.

By 2030, we may have an entirely object- and metadata-based storage system, which means filenames become irrelevant. In the interim, however, an early step to building-block workflows would be to at least normalize the naming systems with an open data model so that productions could be consistent in how they describe foundational pieces, such as a scene or take, and how they name a 3D asset's mesh versus its other component pieces.

We can envision physical studio facilities adapted to support these building blocks. They would have the ability to configure equipment and software to perform certain functions in a room one day, and the next day, rapidly modify the configuration using different building blocks to perform a different function.

> **Rather than predict where artificial intelligence will have the biggest impact in the production and distribution of media, we are focused on creating an environment that enables those AI tools to work most effectively.**

Rather than predict where artificial intelligence will have the biggest impact in the production and distribution of media, we are focused on creating an environment that enables those AI tools to work most effectively. By providing the industry with structured datasets from these building blocks, which can be provisioned for read-only access, AI and ML bots crawling for data and looking for optimizations to workflows and processes will be able to much more easily resolve what data they are scouring. These automations can enable considerable efficiencies throughout production by reducing mundane and repetitive tasks. However, they need to work from a structured dataset.

PRINCIPLE 10: WORKFLOWS ARE DESIGNED AROUND REAL-TIME ITERATION AND FEEDBACK

OVERVIEW

Currently many creative processes happen without real-time feedback. For example, VFX renders can take 24 hours or more to create final composited frames. This makes iteration slow. However, in late 2018, the industry saw dramatic improvements in the quality of video game engines with the addition of hardware-accelerated GPU-based ray-tracing on affordable workstation and cloud graphics cards. In the future, a new suite of filmmaking tools will evolve from today's game-creation engines – so as not to confuse these tools with game creation processes, we refer to them as real-time engines (RTEs). These new tools plus the new cloud foundation principles will dramatically change the creation and economics of filmmaking, potentially upending the sequence of creation, and enable new

workflows at preproduction, production, postproduction and potentially even delivery of filmed media in the future.

Today's (2019) game engines are increasingly used to previsualize sequences of films (usually the complex action-packed scenes). In some cases, entire movies have been previsualized on a game engine system. Game engine renderers have also been used to create final cut renders on some movies. Beyond these examples, there are many more opportunities for these real-time iterative workflows in the future.

EXAMPLES

Preproduction

By 2030, traditional camera-based productions will look increasingly like the workflow for the animated features of today, that is, a world without unneeded setups or unproductive production days. With animation today, the processes are iterative. The movie is first storyboarded out scene by scene. Then it is performed with "scratch audio," and finally with increasingly advanced animatics. At all times, the director can arrange and rearrange the scenes, timing, characters and dialogue. With subsequent revisions, the movie is increasingly locked as it is, then performed with final voice talent, animated with full fidelity and rendered with realistic lighting. We foresee the early production steps for live-action movies using a similar approach and RTEs. The show, potentially with interim actors, can be designed, animated, correctly lit and edited and have accurate camera angles set in the pre-photography stage and potentially before it is greenlit. The quality of the title can iterate in this pre-photo stage and assist a production in making complex decisions and spotting potential issues before they reach a critical state and move to principal photography.

Principal Photography

We do not just foresee RTE tools being used in pre-photography; we can also expect to see those tools being used during production as content is being shot with traditional cameras. By using RTE tools combined with XR technologies, directors will be able to see photorealistic versions of digital characters or objects interacting on-set with physical actors and objects. By looking through the camera lens or via head-mounted displays (HMDs), the cast and crew will no longer see green screens or stand-in representations of digital characters. Two physical actors could act opposite each other despite being thousands of miles apart. A director and cinematographer would be able to make lighting and camera decisions with

> **A director and cinematographer would be able to make lighting and camera decisions with absolute confidence, knowing how the final output of the scene will look with both digital and physical elements blended seamlessly together.**

Postproduction

The process of postproduction and VFX in 2030 could be shorter than it is today because the real-time engine will take away the pain of waiting for lengthy and expensive offline render farms to finish before artists can see with confidence the results of their work. The production could also deliver final plates to the VFX vendors with digital objects or characters already composited in, mitigating much of the work required in current postproduction processes. For some smaller productions, these RTE-rendered scenes may be fine to use for final rendered pixels. These time and cost savings could be used to reallocate budget elsewhere, perhaps forward to pre-photography, or to allow more time for iteration and improve the final product.

Distribution

We can also envision scenarios in which the rendering step occurs on the consumer's device or, just before that, at the edge of the internet during distribution (using perhaps a cluster of GPUs in local neighborhoods with burstable capacity to handle shared graphics compute in a contended way, in much the same way that bandwidth is contended among neighbors in a local zone now). This enables a world of dynamic media that adapts to the consumer's unique playback environment. If consumers are engaging in more immersive entertainment experiences (like video games), then it is possible that the finished form of the media is not a single piece of narrative video, but an entire CGI environment that contains the storylines, objects and characters the director wants to use. That experience could change and react to the audience and their viewing device (for example, to match the native display resolution, color gamut, and frame rate) and not be fixed every time it is consumed – much like a video game today.

Because all the required assets will already be securely in the cloud (Principles 1 & 7), linked

to each other in asset packages (Principle 8) and designed with a new policy approach to publishing (Principle 3), we can then confidently unlock the power of edge compute to deliver these new experiences.

> **"...the adage "we can fix it in post" may change to "we can fix it before we shoot."**

IMPLICATIONS

These examples illustrate how the adage "we can fix it in post" may change to "we can fix it before we shoot." We can also expect changes in the structure and scheduling of major productions, with perhaps less time being devoted to postproduction; instead, those people, budgets and time will be shifted forward to much more robust and fully formed visualizations in pre-photography.

However, the range of additional media that could be captured may mean postproduction evolves into new processes, such as selecting final camera angles and performances from a range of viewing angles that were captured on set.

A new open standard real-time engine package would need to be developed that could deliver the range of digital assets to the consumer to be rendered in real time as they watch the experience.

The impact of the RTE will be considerable and broad, but we can see the need for standardization of some core components of the rendering, translation and packaging of digital assets, and that will require broad industry collaboration across tools, vendors, GPU providers and creatives.

SECTION 4
BENEFITS

The MovieLabs 2030 Vision lays out a world where technology empowers storytellers to realize their vision with higher quality and in speeds only dreamt of today.

Content owners will have the ability to manage more projects simultaneously with more predictability across production back ends. They could see a movie before it is even made and make critical decisions about budgets, marketing and even the viability of a title from early renderings of the content. Transmedia content experiences for big franchises can be developed simultaneously with assets shared and reused from cloud storage and an active library for all legacy content always available and retrievable.

Every participant in the ecosystem benefits by being clearly identified with their work, appropriately credited via a robust tracking log to ensure that every element in a title is correctly licensed and cleared, even through various distribution channels.

These foundational changes shared across the industry will create opportunities for innovative companies of all sizes to develop novel tools, services and workflows that will further accelerate improvements for all stakeholders. The common foundation will also provide stability and predictability for studios while allowing filmmakers to continue to define and refine their workflows for each project.

In this future, providers of media software tools need not build their own user identity systems and authorization back ends. Nor would they need to handle all file operations and complex search algorithms to find files and metadata that was previously buried within proprietary formats. Manufacturers of capture technologies or workstations would not need to develop their own cloud back end, permissions and security systems, but could instead adopt industry standards for security, cloud access, APIs and metadata nomenclature.

And the content-making process will be inherently more secure. Rogue copies of prerelease content will not exist, only provisioned users will be able to perform functions on the

content and every asset will always be secured. MovieLabs seeks to apply the technologies described in these combined principles to ultimately benefit all participants in the ecosystem by further empowering creatives to achieve more with their time and budgets.

SECTION 5
CONCLUSION & NEXT STEPS

This report takes threads of technology that exist today or are emerging and weaves them into a tapestry of what the future may offer. Many of these technologies come to market at the pace that innovation allows, regardless of what this industry does. However, MovieLabs believes there is a considerable opportunity to progress the visions in this paper to be realized sooner, bringing the enormous promised benefits in the near term.

Achieving that will require two things:

1. A clearly delineated plan with a workback schedule that identifies the projects the industry needs to deliver in the next 3, 6 and 9 years to achieve these goals.
2. A coordinated approach in which the research, explorations, learnings and proposals of each group can be shared within the community to refine the approach and innovate even faster.

The future for all participants – from directors to actors, tools providers and lawyers – can be improved if the industry executes this shared Vision, but success will be best achieved if all potential beneficiaries work together appropriately to advance cross-industry goals.

APPENDIX A
GLOSSARY OF TERMS

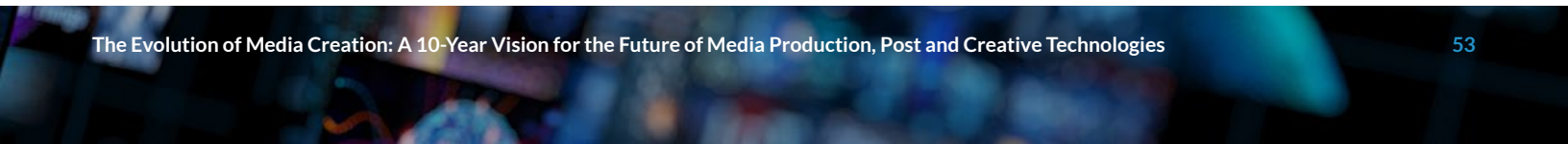
The terms below are referenced in this white paper and therefore defined below. Note that many of these terms have various definitions, but those provided below are how we have defined them for the purposes of this document.

ACES Workflow	Academy Color Encoding System enables a color-accurate workflow with interchange of high-quality motion picture images regardless of source
Animatics	An animated storyboard
Application Programming Interface or API	A clearly defined set of subroutines with a standardized communication protocol to initiate, communicate and derive results from the program
Audio Stems	A combination of audio tracks grouped together and used in the final mix of a piece of content
Blockchain	An implementation of a distributed ledger used to secure and track transactions
Color Decision List or CDL	A metadata file that describes the global, regional and intra-frame modifications to color, tint, temperature, etc. made by a colorist to set the visual look of a piece of content
Content Distribution/ Delivery Network or CDN	A geographically distributed network of proxy servers and their data centers. The goal is to provide high availability and high performance for web content by caching frequently accessed content close to end users
Cloud	We define cloud in this paper as not just the hyperscale cloud providers, but any internet-accessible storage platform that can be used as a common space for collaboration and exchange of data
Compositing	Overlaying and blending different media elements together to create a single image – used, for example, to insert VFX elements within a piece of video ("the back plate") such that they look seamless with the original content.
Conform or Conforming	The final step in editing, when the EDL is used to replace the proxy files with the original camera master files to create a cut of the content with the highest quality source video

Dailies	The video output of each day of shooting with synchronized sound. Typically, dailies include only content of interest to the director.
DCP or Digital Cinema Packages	A packaging format developed by the studios to securely package all the elements of a movie (video files, audio assets, playlist) for distribution and playback in theaters
DeBayering	A step to take the RAW output from digital camera sensors (expressed as a pattern of red, green and blue subpixels) and create usable video files
Digital Intermediate or DI	A legacy term for film being scanned into a digital file format for the application of digital color processes. Now most projects are created digitally anyway, but the phrase is typically applied to the theatrical color pass and the master creative "look" used by all subsequent color grades
Distribution	Preparing and delivering numerous versions of the final content masters to the content creator's distribution partners for onward delivery to consumers
DP, DoP or Director of Photography	The creative responsible for all cameras, lenses and shots during principal photography and whose approval is typically required for the subsequent DI as well
Editorial Decision List or EDL	A metadata file that contains an ordered list of reel and timecode data representing where each video clip can be obtained and its in/out cuts, transitions and other elements in order to conform the final cut. EDLs are created by offline editing systems
High Dynamic Range or HDR	Dynamic range refers to the difference between the darkest (blackest) point in a piece of video and the brightest highlight point. High Dynamic Range typically refers to capture systems capturing >14 stops of light, a workflow to maintain that dynamic range and final graded video with low to zero black levels and >1000 nits of brightness
Interoperable Media Format or IMF	An evolution of the DCP, an IMF package is a SMPTE standard for providing a single, interchangeable master file format and structure for the distribution of digital master packages between businesses around the world. The IMF package can contain all possible distribution variations for a piece of content
Media Elements (MELs)	The components of a finished piece of media, which could be video (down to individual frames), audio assets (music, dialogue, dub, sound FX) or CG assets (animated characters, CGI objects)
OpenEXR	An open image file format developed by Industrial Light & Magic for use in storing digital elements at the highest dynamic range so they can be inserted into scenes with varied contrast levels
Light Field Camera	A light field camera, also known as a plenoptic camera, captures not only the color and intensity of light (as with traditional cameras), but also the angle at which light rays are hitting the sensor(s). Although this increases the amount of captured data, the spatial information can be used for a variety of purposes, such as creating a depth map and adjusting picture parameters including depth of field, focal distance and Bokeh at a later time

Lookup Tables or LUTs	A format for predefined color transforms that change the look of an image, often with a specific target display in mind
Offline Renderer	A tool or process to composite and render disparate media elements (CG, video, audio) to a video in non-real-time (i.e., at less than the frame rate of the finished content). Currently, it can take hours for every frame of video
Over-the-Top or OTT	Distribution of video via broadband IP connections (a historic reference to cable systems, in which video was delivered as the primary signal but IP traffic was delivered as an auxiliary data stream)
Point Cloud	A set of data points in 3D space, typically created by scanning an object or scene and allocating a set number of data points to represent the scene
Postproduction & Mastering	Often the lengthiest part of the creation process, including steps such as editing, adding visual effects (VFX), audio mixing/editing, color grading and the creation of dozens of international masters by combining the various finished media elements together
Preproduction	Includes script development, production design, casting, budgeting and other steps required to greenlight and prepare a piece of content
Previsualization, Previz	An early production process that uses rapid prototyping to enable creative teams to visualize shots, sequences or whole movies before they enter principal photography
Principal Photography	The capture of media, typically on set or on location. The term is used for live action captured with a camera
Production	The steps in capturing and creating content on set, on location, in animation, in VFX, etc. Includes lights, cameras, sets, talent, grips, green screens and huge media files
Public Cloud Provider	A provider of cloud services (storage, compute, database services, etc.) that are generally available for any individual or company to sign up for. We prefer the term *hyperscale cloud providers* to describe companies providing these services
RAW File	A file created by a camera as a direct stream of the data from the image sensor in its "raw" format that needs to be interpreted, typically by a DeBayering process, to make recognizable images on a screen. RAW files can be large, are sometimes uncompressed and are proprietary to each camera manufacturer, unlike a standard compressed image format (JPEG, MPEG, etc.)
Real-Time Engine (RTE)	A tool or process to composite and render a disparate media element (CG, video, audio) to an output device in real time (i.e., at or higher than the frame rate of the finished content). Referred to more commonly in 2019 as game engines because their original function was the real-time rendering of game elements and physics
Real-Time Engine Package	A proposed new packaging format to collate and describe all the MELs within a piece of media, where they move and how they should be rendered to the screen. RTE packages are required by real-time engines to accurately describe a frame or sequence and a finished piece of content

Rotoscoping	A process in VFX to extract one video element from another in a video (for example, to remove a foreground character from the background they were shot against)
Variable Bit Rate or VBR	A video compression system that encodes a video file in multiple different quality levels and bit rates to enable playback that adapts to a consumer's device characteristics or bandwidth constraints and is also used to enable rapid start-up of streamed content
Virtual Production	A relatively new process combining VFX elements in real time during principal photography to enable creatives to see and interact with both physical and digital worlds simultaneously
VFX, Visual FX, Computer-Generated Imagery (CGI)	Terms largely used interchangeably now to refer to assets generated by computer artists or simulations rather than captured by cameras
Volumetric Capture	An additional step during photography that adds capture of data from different angles that describe where characters and physical objects are in 3D space. Can enable many more camera angles than a traditional camera and can be used to interpolate additional 3D angles and computational images/sequences not possible with a physical camera
XR	A catch-all phrase to encompass the general technologies of augmented reality (AR), mixed reality (MR), virtual reality (VR) and other immersive video formats like 360° video
Zero Trust Network	The zero trust model is a security principle that says organizations should not automatically trust anything inside or outside an organization, and that anything and everything must be verified before access is granted



APPENDIX B
ACKNOWLEDGMENTS

This paper would not have been possible without the invaluable support and contributions of the following executives:

Anthony Anderson, *Universal Studios*

Shadi Almassizadeh, *Disney*

Bill Baggelaar, *Sony Pictures*

Hanno Basse, *20th Century Fox (former)*

Richard Berger, *MovieLabs*

Annie Chang, *Universal Studios*

Susan Cheng, *Warner Bros.*

Vicky Colf, *Warner Bros.*

Daniel De La Rosa, *Sony Pictures*

Miles DelHoyo, *Warner Bros.*

Daphne Dentz, *Warner Bros.*

Eddie Drake, *Disney*

Don Eklund, *Sony Pictures*

Anthony Guarino, *Paramount Pictures*

Jim Helman, *MovieLabs*

Arjun Ramamurthy, *Disney*

Horst Sarubin, *Universal Studios*

Craig Seidel, *MovieLabs*

Jason Shea, *Disney*

Adam Slohn, *Warner Bros*

Spencer Stephens, *TechXMedia*

Mike Trainotti, *Paramount Pictures*

Mark Turner, *Entertainment Technologists*

Jamie Voris, *Disney*

Michael Wise, *Universal Studios*

MovieLabs is a jointly run industry lab that enables member studios to work together to understand new technologies and drive the right ones to adoption. We help set the bar for future technology and then define specifications, standards, and workflows that deliver the industry's vision. Our goal is always to empower storytellers with new technologies that help deliver the best of future media.

Email: info@movielabs.com
Twitter: @ MovieLabsNews
Website: movielabs.com

HEADQUARTERS
222 Kearny St, Suite 300, San Francisco, CA 94108