

Skywalker Keeps the Humanity in Automated Soundtrack Mastering

The “Coda” Automated Media Ecosystem Preserves Creative Intent While Automatically Generating Hundreds of Versions and Language Packages

**765**

different language audio packages automatically created

Weeks of work cut to

REAL TIME

Improved audio formats for

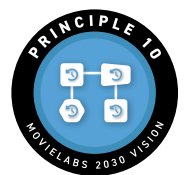
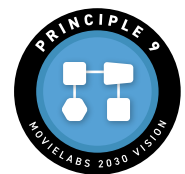
EVERY LANGUAGE

SUMMARY

With the explosion of streaming services that operate on a global footprint, the mastering pipeline has been stressed like never before. Each major movie or TV episode typically can have as many as 765 different language audio packages created from the US original, each of which needs to be quality checked, encoded and packaged for delivery. The process can take weeks and involves creative and technical teams working to make derivative versions which must preserve the original creative intent as much as possible.

The “Coda” Automated Media Ecosystem is a new extensible software platform from Skywalker Sound that automates the creation of soundtrack versions and cuts the deliverable process down from weeks to faster-than real-time. The system has already been used on premium Disney+ releases such as The Mandalorian and Moon Knight.

By automating the international and multi-format soundtrack processing from the highest original source mix format (often the Dolby Atmos mix), the automatically derived versions can be created with the same fidelity and attention to detail as the original language, improving the experience for all consumers in international markets. As the process is automated without the possibility of human error, the need to QC each pass is reduced, resulting in better results for consumers and considerable time and cost savings for content owners.



BACKGROUND

Modern movies and TV shows are played on a huge variety of devices and environments – from multiplex theaters to home cinema systems and headphones on mobile devices – and the soundtrack must be mixed specifically for each combination of playback system, environment, and codec. The versioning process begins when the final creative sound mix has been signed off by the director. For a US motion picture release, this will usually be the US theatrical mix in an object-based surround format such as Dolby Atmos. Then the versioning process can start, which sometimes requires a team of sound mixers to translate the creative intent into different output formats for home media object-based audio, 5.1/7.1 surround, stereo, and near-field mixes (for headphone and home

entertainment uses). Each mix is completed for corresponding international versions, which often includes the added complication of mixing in different dubbed audio tracks for dialog. The large matrix of versions must also include descriptive audio and censored releases. Each newly mixed version must be QC'd by another team and leveled appropriately for local market regulatory loudness requirements and with correct metadata for the package.

The entire process occurs during the finishing stages of post-production when final picture and sound are ostensibly locked. There is intense pressure to deliver, often with time and budget constraints imposed on the teams, which sometimes results in use of shortcuts. For example, international markets may be limited to stereo or surround mixes without object-based audio, or release in some markets may be delayed, with media not available 'day and date' in all formats/markets simultaneously. These compromises are often necessary to meet deadlines and budgets, but can result in lower quality deliverables for some markets, devices or distribution technologies.

The complex logistics of these final mastering steps requires moving multiple iterations of granular files around the world—all with subtly nuanced differences involving versions and languages—so that they come together in dozens of different deliverable packages. The complexity of this process increases the chance of human error, especially under the stress and constraints of producing content for today's global markets. In addition, each show has unique creative needs with different studio policies to contend with – the result is a succession of 'snowflake' workflows, different for every production.

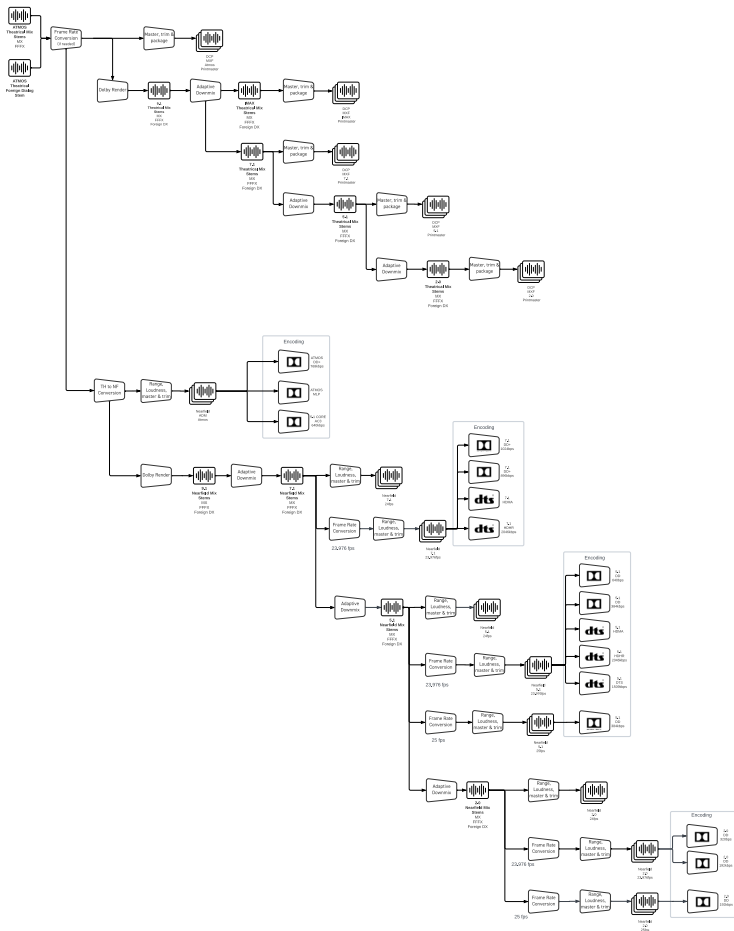


Figure 1: Each language deliverable package requires multiple steps of technical audio and adaptation tasks (including downmixing, loudness correction, editorial conforms, framerate conversions, home theater adaptation, and audio file formatting) by a mixer using specialist rooms and equipment, followed by encoding and packaging. The process can take 1-2 weeks

SOLUTION

Skywalker Sound developed new internal workflows and tools to cope with these challenges in its primary sound services business. That automated software solution, which made its own internal processes more efficient, has been matured and is now offered to the market as a separate product for others to implement at scale. The result of seven years of development, the 'Coda' automated software mastering system takes the complexity and manual attention out of producing mastered media packages and transforms a process that took teams of people weeks of work into an automated workflow that runs the entire process in hours.

Coda takes the QC'd master audio file, in the highest fidelity mix available, and runs a series of automated, simultaneous tasks to create all required deliverables from that master consistently in faster than runtime of the content – i.e. A 30 minute show can be processed, creating all versions in less than 30 minutes. The processes are optimized and GPU-accelerated so that it can be run multiple times if outputs must be regenerated to

accommodate technical or artistic changes to the final source mix. This flexibility takes enormous pressure away from the delivery teams and gives creatives the time to make needed changes without compromising the team's ability to generate the full array of highest quality deliverables. All derived versions can retain the original mix (e.g., full object-based audio), avoiding quality gaps and the shortcuts that sometimes compromise previous versions in some markets.

The system automation includes a new feature, dubbed Intelligent Workflow Creation, that calculates the optimum processing workflow chain for each job based on the preferred outputs and deliverables of each content owner, maintaining the flexibility to deliver multiple variable packages while removing the complexity of extensive manual process programming for each new show/film. By examining the input files and the required output deliverables, Coda intelligently creates a unique and dynamic script to run workflow tasks in the optimum order to ensure the best quality deliverables for the requested outputs. At the end of the workflow, Coda deposits and delivers new assets programmatically to a host of destinations as defined by content owners. This allows for an automated media finishing pipeline from final sound mix to delivery to distribution, differentiated for each piece of content and the specific needs of each production.

While the Coda system is impressive software automation, its true power lies in the intelligence built in for each audio processing task, developed and tuned by Skywalker Sound's unique talent pool over a nearly 50-year history of best practices and workflow innovation.

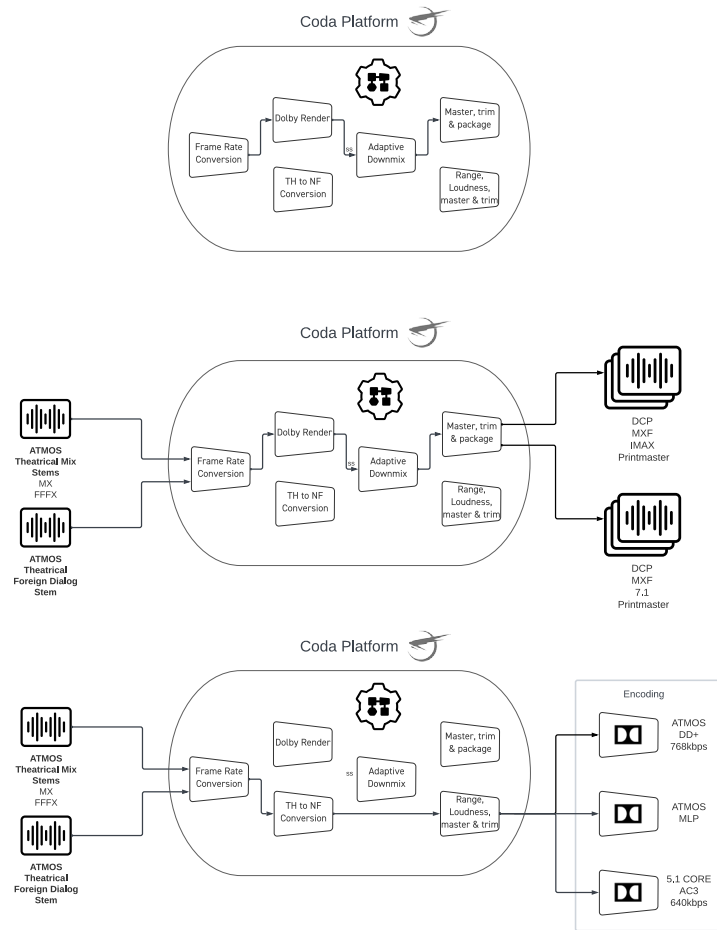


Figure 2: By using Coda, Skywalker can initiate all the deliverables simultaneously. The Intelligent Workflow Creation Manager selects the correct sequence of automated tasks to deliver the correct media packages with all required adaptations for the specific market package.

“Without a modern AI tool such as Coda, it would be practically restricting to meet these requirements with 100% accuracy within the budgeted time. I would much rather divert that time and effort into the creative process and give the director and sound crews that extra breathing room to make the project really shine. Now that we have such a tool, I would not deliver a show without it.”

CHRIS BARNETT, RE-RECORDING MIXER/SOUND SUPERVISOR, SKYWALKER SOUND

ARCHITECTURE

The Coda platform is constructed as a series of containers incorporated in a Kubernetes cluster. Each container represents a module, whether it be processing, orchestration, file manipulation, or any necessary third-party library functionality. While the various containers serve distinct functions, they all communicate via a common API, defining internal Skywalker functions as well as intramural library functionality, allowing an orchestrator to submit job commands via payloads, as well as keep track of progress.

The modular nature of the Coda platform (including the orchestrator itself) allows for seamless scaling upwards and downwards based on the volume of submitted jobs and processing required for the jobs. Similarly, as the Platform API and UI are also containerized, operator access and job submissions both scale as needed to meet demand.

The intelligent orchestrator analyzes the operator’s job submission dynamically to build a unique workflow of processing modules based on the declared source assets and desired output media. Once the workflow is constructed, the modules required in the workflow are spawned and given prescribed processing instructions and temporary storage locations. The modules and all associated resources are expunged when the operator job has completed successfully, with log entries from each process written to a persistent location for future reference and data analytics.

BENEFITS

As Coda reduces human involvement in the post-production and mastering process, the largest and most immediate benefits are time and cost savings – from potentially hundreds of hours of human work (some with high end talent) to zero, saving hundreds of thousands of dollars per production. In addition to the lower costs, Coda also gives teams days of extra time—a critical resource especially at the end of production—which can be used to speed up a release or iterate on the creative mixes themselves. Studios implementing Coda also save costs and time on QC (since derived versions are QC'd upstream), reduce time and resources spent sending, finding and chasing deliveries of different versions of media, and decrease risks of errors introduced by these more manual and complex processes. These changes lower the stress of teams delivering the final outputs at the very end of production, a time when extra vendors and resources are thrown at projects to ensure delivery timelines are met, often resulting in budget overages. More importantly, creative time can be maximized by minimizing technical pipeline needs. Coda offers immediate and predictable results which can be scheduled and run without human intervention or oversight.

Explicit Cost Savings:

- Mastering in post-production, int'l dubbing and studio ops now automated
- Reduction in stage/mixer/editorial/support costs on each project
- Lower costs and reduced need to perform QC (fewer versions made manually)
- Reduction/elimination of costs related to fixing problems identified by current manual QC process
- Reduced costs in studio ops services: labor, real estate, specialty equipment management, scheduling, infrastructure, error reduction

As production resources (people, infrastructure, creative rooms, mixing equipment and software licenses) continue to be in short supply, studios and post-production houses can streamline operations by trusting the Coda automation to handle the rote tasks, all while freeing creative talent to focus on the primary objective of adding more value to the content. In addition, with fewer transfers, systems and vendors involved, there are fewer risks of security mishaps, breaches or misconfigurations, enabling an automated cloud-based workflow to be more secure as well. Consumers worldwide also benefit by getting high-end processed audio in every language, mixed for the specific format on every device, adding value to the content and the delivery platform (streaming, pay tv, home media, theatrical performance).

Implicit Savings:

- Mastering & finishing time per title (projects finish 1 to 2 weeks earlier with automated mastering/finishing)
- Streamlining of studio group logistics: fewer operators, immediate fixes, less redundant data
- Reduced need for spillover mastering to third party/external vendors
- Security improvement with the ability to keep more pre-release content in-house.

Content Improvements:

- Less time required for technical mastering = more time for creative work = better content, more languages
- High-end audio processing for every language
- Improved audio formats for every language—regardless of regional spend capacity (i.e., everybody gets immersive audio track ATMOS)
- Enhanced brand value for international projects with correct formats for each venue in each language
- Increased soundtrack quality control from creative mix

PARTNERS

Coda was developed and created in-house by Skywalker Sound technology teams, with the exception of third-party libraries required for production deliverables (e.g., Dolby and DTS SDKs for encoding). The system is built in an extensible model so the core automation and scripting engine can be driven by APIs to enable Coda to integrate into more complex asset management systems and drive other non-audio media workflows.


ALIGNMENT WITH MOVIELABS 2030 VISION PRINCIPLES

The 2030 Vision includes software-defined workflows that can be dynamically generated based on the assets and tasks to be processed. The Skywalker Coda platform demonstrates how these concepts can be implemented today with a workflow that changes dynamically based on the tasks required for specific assets and productions.


This use case demonstrates the following 2030 Vision Principles:

PRINCIPLE 1

Skywalker's Coda platform can be deployed by studios and post-production providers on private or public cloud infrastructure. As assets are uploaded by Coda they may receive a unique URI so all processes can discover them in the cloud system. Only 1 copy of all files is created to minimize duplication. All resources are then completely torn down upon completion of the job.

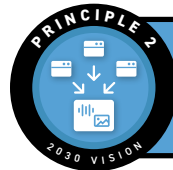


All assets are created or ingested straight into the cloud and do not need to be moved




PRINCIPLE 2

Coda hosts a range of audio processing engines which operate in different sequences depending on the specific workflow for that production and/or studio. All the processes are containerized and operate directly on the cloud copies of the assets. Coda is extensible so additional audio and other media processes can be added in the future.




Applications come to the media




PRINCIPLE 9

The Coda platform not only demonstrates the value of automated workflows created using open APIs and common data formats but it dynamically creates those workflows using intelligent analysis of the assets it's provided. (a software compiled workflow). This is a great example of how standardizing data and metadata formats can provide the opportunity for automation to make our lives easier and by exposing APIs, other applications can also interface with Coda and expand their mutual capabilities.



Media workflows are non-destructive and dynamically created using common interfaces, underlying data formats and metadata



PRINCIPLE 10

Time is the currency of productions, and the value of that currency increases the closer to release. By changing an old time-consuming manual process to a new automated system Coda demonstrates how software defined workflows can dramatically open up more time for productions, especially in that critical last few weeks of post.



Workflows are designed around real-time iteration and feedback



The Coda tool does an incredible job of maintaining the mix balance and creative intent of our ATMOS mixes through all delivery formats and effortlessly meets the stringent loudness specs of today's streaming services."

SCOTT R. LEWIS, RE-RECORDING MIXER, SKYWALKER SOUND

NEXT STEPS

Skywalker learned with the development of Coda that true modularity and containerization is critical in developing systems for media assets. From digital signal processing to databases, components need to be continually upgraded and modified. With monolithic applications, there is huge risk of inadvertent effects when modifying a seemingly isolated component. However, in a containerized environment, API and payload validation allows for programmatic understanding of module version and processing abilities.

Skywalker makes the case that firms providing content creation software and tools would greatly benefit from offering programmatic entry (APIs) into their applications, as well as SDK tools to assist with the inter-exchange of proprietary file formats when building platforms like Coda.

Similarly, there is a need for a federated identity structure, where individuals, firms and studios may interlink their particular identity provider with others, allowing roles and true identities to propagate across the industry. This is described in Principle 6 of the 2030 Vision. Applications that deliver interoperable software-defined workflows that span organizations, like Coda, require this common production user identity.

MOVIELABS PERSPECTIVE

The Skywalker Coda case study demonstrates the benefits of an automated process that handles repetitive tasks imposed on creatives by legacy workflows, i.e. downmixes which require little creative interpretation. By focusing creative talent on the artistic work that creatives love and trusting the deliverables to automation, Skywalker has enabled a dramatically faster workflow that is less prone to human error, unlocking considerable time and resource at a critical and stressful window at the end of post-production. As Coda is extensible through API, it also provides a window into a future where other applications and non-audio workflows can be processed by the same intelligent workflow manager through common interoperability. That is an enticing proposition for the 2030 Vision's focus on software-defined workflows and lays a great groundwork for others to build upon and around.