The Nineteenth Annual Interactive Audio Conference
PROJECT BAR-B-Q 2014

Group Report: Metadata = Money

Participants: *A.K.A. "Metadata Actuaries"*

Philip Nicol, Dolby
Ty Kingsmore, Realtek
Sergio Liberman, Freescale
Jack Joseph Puig, Waves
Ted Kao, DTS

Marcus Altman, Dolby
David Roach, Optimal Sound
Bobby Lombardi, Gobbler
Andy Rumelt, Cirrus Logic
Bryce Toth, Beats By Dre

Facilitator: Jim Rippie, Invisible Industries

Problem Statement

Experiencing media is shifting from strictly linear to a more personalized multi-stream object oriented approach to audio and video consumption. Current metadata content and tracking methods are inadequate for this expansion in experience. Technological advancements in the past decade have presented opportunities to better personalize user experience, but its effect in the consumption of audio and video has been lacking. Existing methodologies do not provide metadata preservation solutions for the following:

- Artist's artistic intent.
- Rights holder tracking and detection for content and processing algorithms.
- Device and playback environment.
- Derivative work products, including in-band and out-of-band metadata and content.
- Synchronized bidirectional workflows, including biometric tracking.
- Identification of simultaneous tracks.

We propose recommendations to accomplish these goals, via the flow of metadata through Production, Transmission, Render Device, Monitor Device, Browser, and Application. Within each of these components, short term and long term recommendations will be explored, as well as opportunities moving forward.

Opportunity Statement

The resulting effects arising from our proposed recommendations will benefit many parties within the metadata stream, including content owners, aggregators and distributors, technology providers, and consumers. For content owners and distributors, our recommendations will help aid in piracy mitigation. Opportunities also arise in Big Data, as the flow and preservation of metadata will be improved as part of the production/consumption and can be personalized to the intended user or environment.

What is Metadata

Metadata shall be defined as "data about data". The purpose of metadata is to facilitate the discovery and exchange of relevant information.

Examples of stream-based metadata that we are covering include:

- Sample Synchronized Metadata: (*e.g.: Fader information)*
- Frame Synchronized Metadata: (*e.g.: DRC, object position*)
- Scene Synchronized Metadata: (*e.g.: Average loudness)*
- Container Synchronized Metadata: (*e.g.: MP3 tags, copyright information)*

Examples of device metadata include:

- Speaker or microphone configuration
- Transducers' Response and Limitations
- Effects processing capability such as calibration, EQ, compression, etc.
- Synthesize capabilities
- User Interface capabilities

Examples of sensor metadata:

- Biometrics (heart rate, temperature, pulse)
- Room acoustic response
- GPS
- Mood
- Facial expressions
- Speech and environmental noise (Always On/Listening): key words, sirens
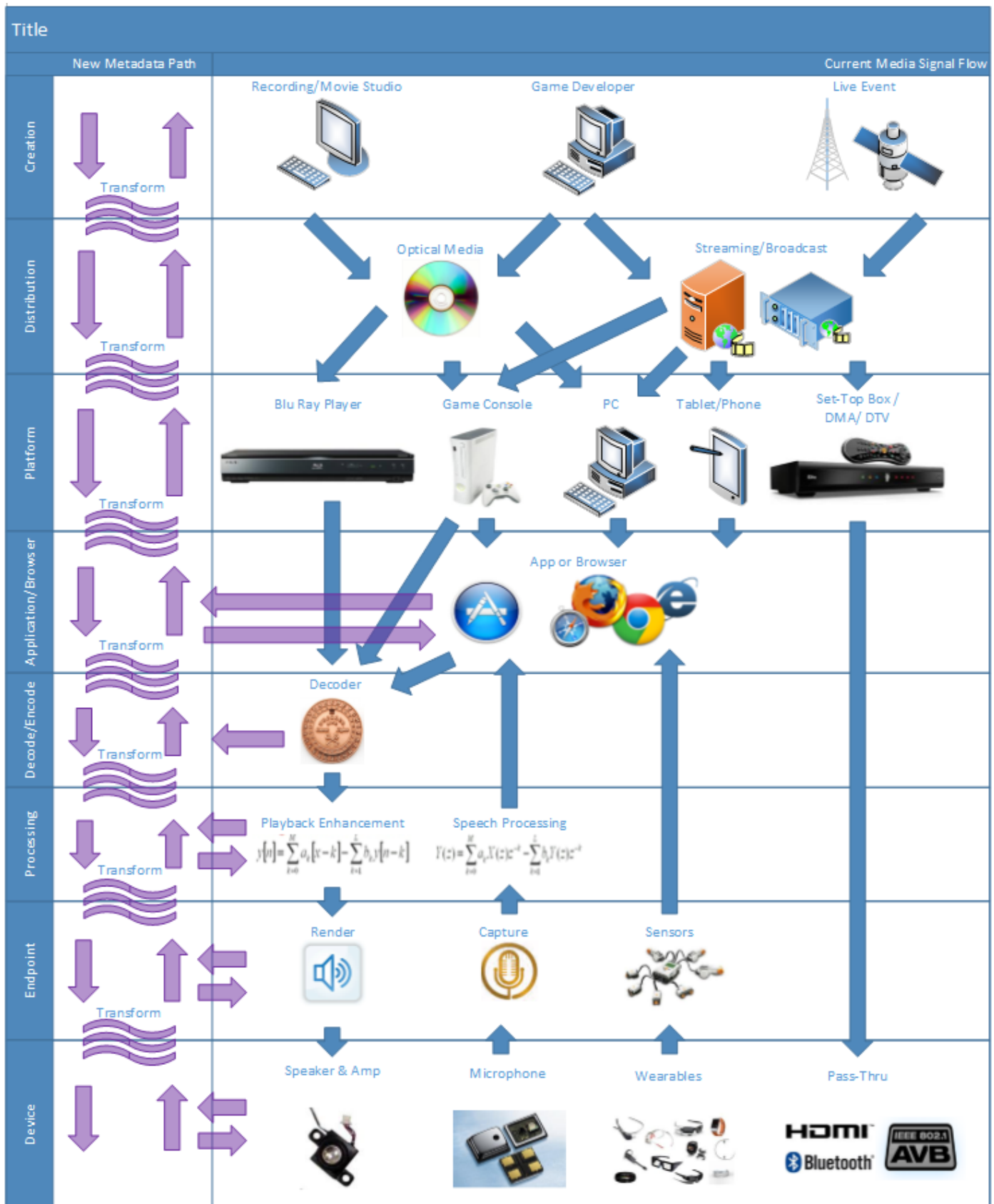
This diagram shows in blue the current signal flow for many different combinations of devices. The violet arrows show the recommended path for metadata. The metadata transforms occur as the metadata is passed from stage to stage, adding or removing metadata along the way.

Recommendations

Production:

The production stage of the data stream represents content creation and preparation for distribution or transmission. It is here that the content's intrinsic properties can be captured and passed along as metadata.

Short Term Recommendations:

- Ensure the preservation and interoperability of metadata throughout the production process and into the transmission process. Metadata created at the stage of production shall be preserved throughout the entire process, down to consumption. Interoperability throughout the different stages of this process shall also be preserved so that metadata can be carried from stage to stage.
- Production tools shall provide metadata input at different levels of granularity, including sample, frame, scene, and container.
- Production tools shall be able to simulate or preview different types of end points. This endpoint modeling can be used to better prepare content for end user consumption.

Long Term Recommendations:

- Once established, these mediums for the preservation and interoperability of content metadata will provide opportunities for live & interactive content production and control.
- The content creation process shall see a convergence of interactive and static production tools.

Opportunity:

The recommendations of the production phase will help to drive premium experiences and the new revenue opportunities associated with them, while also ensuring a more consistent & personalized user experience altogether. They will also provide a more effective method in tracking content ownership and aiding in piracy mitigation.

Transmission:

The transmission stage of the data stream represents the delivery of content and metadata from the distributor to the consumption device.

Short Term Recommendations:

- Ensure the preservation and interoperability of metadata throughout the transmission process and into the render stage of the process. Metadata generated at the point of creation shall be effectively transported to the render device, and shall be preserved for any other points of interoperability along the transmission path.
- If rendered or transcoded anywhere along the transmission path, metadata should be augmented to reflect that operation, and all essential metadata must be preserved.
- Example devices that operate as part of the transmission stage include devices such as DMAs and set top boxes.

Long Term Recommendations:

- Ensure that from the endpoint, metadata can also be passed upstream. The transmission stage shall include frameworks for the preservation and interoperability of metadata bidirectionally, always preserving essential metadata in the uplink and downlink.
- Data aggregated from an IoT environment shall be able to be included in the metadata uplink and the transmission stage must ensure the preservation and interoperability of this newly acquired data as part of that uplink.

The recommendations of the transmission stage will help in delivering premium and personal consistent user experiences. With the framework for bidirectionally flow of metadata in place, content can be rendered and adjusted for playback based on personal preference or environment. As a result, IoT data collected further downstream can be streamed upward or used during rendering. Furthermore, content ownership information can be passed down the data stream to aid in piracy mitigation. The revenue opportunities arise in this mitigation, as well as in the application of the transmission of this metadata.

Render Device:

The render device will render the content for consumption by the end user. It is at this stage that the metadata received from earlier stages in the process will be used to render the content.

Short Term Recommendations: Engagement with stakeholder organizations

- Ensure that the attributes of the endpoint device are available to modules earlier in the signal processing chain. Some examples of such attributes include:
    - Number of speakers
    - Type of headphones
    - Listening environment
    - Biometrics
    - Head-tracking information
    - Playback device capabilities
- Ensure that metadata can flow through the entire media stack *(e.g.: from demux to decode to process to render, and also to smart amps and DSPs).*
- Report processing capabilities that are present and the state of their parameters.

Long Term Recommendations: Execution

- Evolutions in OS architectures shall provide support for metadata exchange throughout the media stack.

Opportunities:

The recommendations provided for the render device will aim to improve current user experience. By using metadata from both upstream and downstream of the render stage, the device can render the content differently. This provides for opportunities in a more personalized and premium user experience while also bringing consistency to the content delivery. This improvement in user experience can result in many new revenue opportunities, including what applications can be derived from metadata. Because the render device serves a crucial role in content delivery, metadata can provide the device with the content ownership data required for playback, and can mitigate piracy in this manner.

Monitor/Record Device:

The monitor or record device will monitor and sense current state and environment of the playback device. It is at this stage that metadata meant for streaming upward is aggregated and transmitted.

Short Term Recommendations:

- Captured metadata should be time-stamped for synchronization.

- Ensure that attributes of the endpoint device are available to modules later in the signal processing chain. Some examples of these include sensors, microphones, GPS, etc.
- Ensure the accessibility of metadata between render and monitor stages.
- Ensure support for user generated capture. Metadata must be able to reference other metadata from a parent-child relationship (*DJ remix, user ratings, game play events, etc*)

Long Term Recommendations:

- Sonic profiling based on environment and hardware for optimized and personalized playback.

Opportunity:

The recommendations provided for the monitor/render device will aim to provide more premium and personalized user experiences by upward transmission of metadata. Revenue opportunities arise in this metadata and its application. Further opportunities arise from the application of metadata acquired as part of the monitor device include the protection of content ownership, little data aggregation, and technology intelligence.

App / Browser:

Similar recommendations apply to an App or Browser as part of the media stack.

Short Term Recommendations:

- HTML5 must support standardized bidirectional flow of metadata.
- The application must support bidirectional flow of metadata.

Long Term Recommendations:

- Execution on bidirectional flow of metadata in both applications and in HTML5.

Opportunities:

Bidirectional flow of metadata is crucial for applications and browsers to utilize the metadata to drive improved and personalized user experiences, and ultimately provided opportunities for new revenue. These recommendations also provide means for big data collection and content ownership protection.

Out of Scope Material:

As the above-presented topic is quite broad, a more focus approach was taken. Below lists topics covered in discussion but not included in this report.

- How to define the level of synchronization and how to implement it.
- Deciding when metadata is sideband, in-band, or out of band.
- How to define monetization model.
- Define low level interfaces to transfer metadata from one domain to another.
- How an OS will pass metadata through the media stack.
- How these recommendations will influence copyright law.

Other reference material:

- Project Bar-B-Q 2013: Using sensor data to improve the user experience of audio applications
- Project Bar-B-Q 2006: Providing a high level of mixing aesthetics in interactive audio and games
- Project Bar-B-Q 2000: Multichannel audio