

MetadataQ:

Automated, Frame-Level Intelligence for Video, Delivered Inside **Avid** and **Grass Valley**

An effortless way to make your video searchable, understandable, and easier to work with directly inside the systems your teams use every day.

MetadataQ is Digital Nirvana's AI-driven metadata automation platform that analyzes video and audio content frame by frame and writes structured, time-coded metadata directly into Avid and Grass Valley environments. It combines speech analysis, natural language processing, and computer vision to describe what is being said, what is happening visually, and how content is structured without introducing a new interface or changing existing workflows.

The Reality of Working with Video at Scale!

Modern media organizations generate and ingest enormous volumes of video every day. Live feeds run continuously. Multi-camera productions create hours of raw footage per shoot. Archives grow faster than teams can review or enrich them.

Across industries, the challenge is not the lack of video, it is the lack of structured understanding at the moment decisions need to be made.



In news and broadcast operations

Hundreds of live feeds may be ingested daily. Without real-time transcription and segmentation, identifying relevant events across feeds remains slow and manual.



In sports workflows

Key moments occur within seconds. Without play-by-play scene detection and frame-level tagging, editors must review long recordings to locate highlights, which delays turnaround and limits media reusability.



In entertainment and archives

Decades of content exist with minimal metadata. Finding specific scenes, themes, or appearances often requires manual review, leaving large libraries underused.



In AI and data-driven initiatives

Raw video lacks the structured annotations needed for training and analytics, making it challenging to convert footage into usable datasets.

Addressing these challenges requires breaking video down into structured, time-aligned elements that systems and teams can actually work with.

What Production and Editorial Workflows Require at Scale?

As content volumes increase, solving discovery, turnaround, and reuse challenges requires more than faster logging or incremental tooling. It requires a consistent way to understand video at a granular level and make that understanding available inside the systems teams already rely on. Video must be broken down into time-aligned elements that reflect what is being said, what is happening visually, and how the content is structured, so teams can work with the material efficiently without manually reviewing entire recordings.

What this looks like in practice



How MetadataQ Processes Video?

MetadataQ is designed to fit into existing broadcast and production infrastructure without adding operational complexity. It follows a clear, auditable process that ensures metadata is accurate, time-aligned, and usable inside editorial systems.

From ingest to usable intelligence



Ingest

Video is received from live feeds, ingest pipelines, camera media, or archived assets through existing Avid and Grass Valley workflows. Both continuous streams and file-based media are supported.



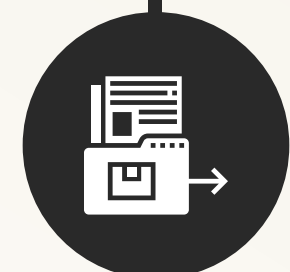
Analyze

Audio and video are processed using AI models tuned for broadcast and production environments. Speech is analyzed for dialogue, speakers, and terminology. Video frames are examined for faces, objects, logos, on-screen text, and scene changes. This data is then combined to create meaningful descriptions of what is happening within the video (e.g., police scene outside a residence; news presenter addressing audience; World Series victory parade; Cardi B speaking with Jimmy Fallon). Content is segmented into meaningful topics and scenes.



Time-Code Alignment

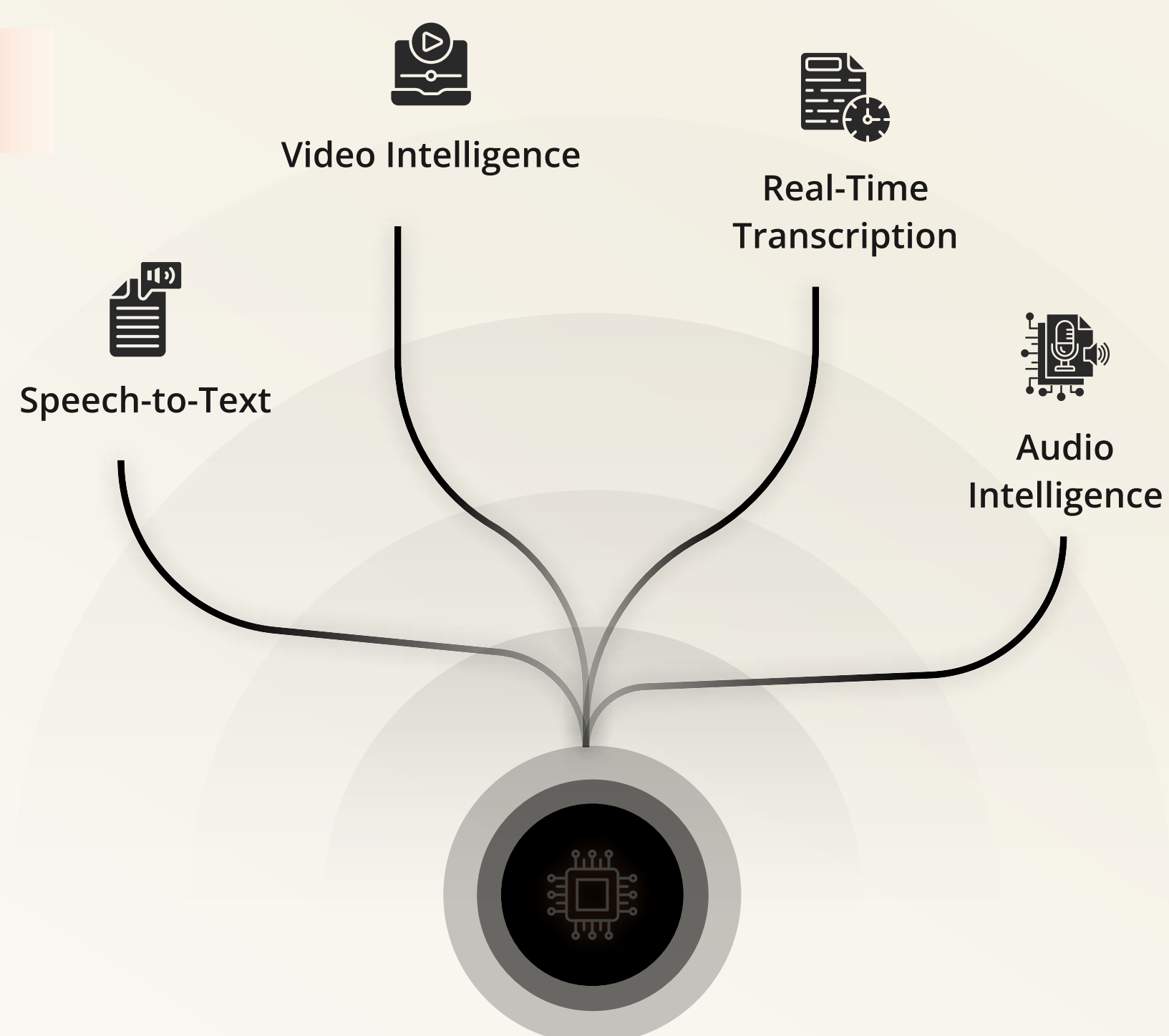
Every detected element is precisely synchronized with the timeline. Metadata is frame-accurate, ensuring editors can rely on it for navigation, search, and clipping.



Write Back

Structured metadata is written directly into Avid and Grass Valley systems as native markers, locators, and searchable fields. No secondary interface is required.

Once the need for time-aligned, structured understanding is established, the next question is how to achieve it reliably within real-world broadcast operations. The process must work continuously, support both live and file-based media, and integrate directly with existing systems. What follows is the operational flow that turns incoming video into usable, editorial-ready intelligence.



How MetadataQ Understands Video?

Processing video is only the first step. The real value comes from understanding what the content represents in editorial, visual, and contextual terms. This understanding enables teams to search for, navigate, and reuse video with confidence rather than relying on manual review.

MetadataQ examines both audio and video to extract meaning that aligns with how editorial and production teams think about content.

01

What is being said?

Spoken content is converted into structured, time-coded text that can be searched and navigated precisely. Speech is analyzed for dialogue, speaker changes, and terminology, producing transcripts that reflect real broadcast usage rather than generic language models. This makes it possible to locate quotes, statements, and discussions quickly within long recordings or live feeds.

02

What is happening visually?

Visual context is analyzed at the frame level to understand not just what appears on screen, but what is happening within the scene. Instead of producing isolated tags for faces, objects, or logos, visual elements are interpreted together and written as contextual descriptions that explain the moment in human-readable terms.

For example, rather than listing individual detections such as a person or a brand logo, the system can describe scenes like “a man holding a Coca-Cola bottle and cheering among the audience.” On-screen text such as lower-thirds, captions, and graphics is also captured and incorporated into this context. Scene changes and visual transitions are identified so teams can locate moments based on meaningful visual activity, not just keywords or speech.

03

How the content is structured?

Long-form footage and continuous streams are segmented into meaningful sections based on topics, scenes, and editorial shifts. Each segment is aligned to the timeline, making it easier to understand how content flows and where specific moments begin and end within a program or recording.

04

How teams work with this understanding?

All extracted information is written back into editorial systems as searchable, time-aligned metadata. Editors and producers can combine speech, visual elements, and structure when searching, allowing them to move directly to relevant moments without scrubbing through entire timelines.

The result is not just faster search, but a clearer understanding of what exists across live feeds, recorded media, and archives.

How Different Teams Use MetadataQ?

Once the video is processed and understood, its value depends on how easily teams can apply that understanding to real work. MetadataQ supports a range of editorial, production, and archival workflows by enabling content to be searchable, navigable, and reusable at scale.

News & Broadcast

- National broadcasters ingest large numbers of live feeds every day, often **exceeding 100 concurrent streams**
- Live content is transcribed and segmented in real time, creating searchable topics such as press conferences, rallies, or breaking events as they occur.
- Editorial teams can **clip and publish breaking news immediately**, while compliance teams filter and review sensitive segments without manual scanning.

Sports

- Sports broadcasters use **play-by-play scene detection** to identify actions and key moments as they happen.
- Entire sequences from possession to outcome are **time-aligned and tagged**, allowing editors to isolate and repurpose highlights within seconds.
- The resulting metadata can also be used by analytics and coaching systems, providing machine-readable descriptions of plays for performance analysis and training.

Archives & Entertainment

- Production companies and networks apply MetadataQ to large libraries of existing content, including decades of archived material.
- **Semantic search and scene-level descriptions** make it possible to locate moments based on actions, themes, people, or tone, rather than relying on episode titles or manual logs.
- Generated metadata can also be used to train internal AI models to recognize **custom objects, gestures, or visual patterns** specific to an organization's content.

AI & Machine Learning Data Training

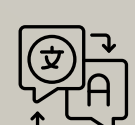
- **Sports analytics:** Extract labeled segments such as **"free throw," "corner kick," "penalty save,"** or **"fast break"** from thousands of hours of match footage to train performance analysis and prediction models.
- **Safety and research:** Identify and tag specific events such as vehicle accidents, **near-miss incidents**, hazards, or crowd behavior across large volumes of surveillance or traffic footage for AI model training.
- **Environmental and monitoring use cases:** Detect and label movement or activity from stationary cameras, such as human or animal motion in **forest surveillance** footage, to train **intrusion detection**, wildlife monitoring, or early-warning systems.

MetadataQ does not simply describe video it converts continuous media into structured data that supports editorial workflows, analytics, and next-generation AI systems.

Capabilities That Support Editorial and Production Workflows

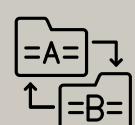
MetadataIQ provides a set of modular capabilities that work together to convert video into structured, usable information. Each capability can operate independently, but they are designed to reinforce one another within production, editorial, and data-driven workflows.

Speech & Language Understanding



Transcription and Translation

Multi-language speech-to-text with speaker identification and precise timecodes. Output supports subtitle creation and downstream localization workflows.



Vocabulary Adaptation

Language models can be tuned for domain-specific usage, improving accuracy for names, terminology, and broadcast-specific language.

Sports-Specific Intelligence



Play-by-Play Analysis

Sports action is described as complete sequences, with in and out points identified to accelerate clipping and highlight creation.



Event-Level Tagging

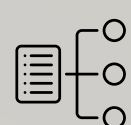
Individual plays, actions, and outcomes are time-aligned for reuse in editorial, analytics, and coaching workflows.

Content Structure and Context



Scene Segmentation and Summarization

Long recordings are broken into logical scenes or segments, each with a concise summary that reflects editorial intent.



Topic Classification

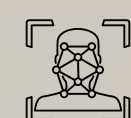
Content is categorized using standardized taxonomies, including IAB classifications, supporting editorial organization, marketing workflows, and ad tech alignment.

Visual Understanding



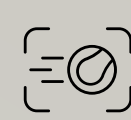
Scene Descriptions

Visual activity is described in clear, human-readable terms, explaining what is happening on screen rather than just listing detected objects.



Facial, Logo, and Object Recognition

Identifies recurring people, brands, and visual elements within frames, enabling search and tracking across large content libraries.



Optical Character Recognition (OCR)

Reads on-screen text such as lower-thirds, captions, signage, and graphics, making visual text searchable.

Content Compliance Monitoring



Sensitive Content Detection

Automatically flags profanity, nudity, violence, and other sensitive material to support editorial review and compliance processes.

Editorial Prioritization



Virality and Impact Scoring

Analyzes content signals to surface moments with higher potential editorial or audience impact, helping teams prioritize what to review and publish first.

Deployment and Integration

For metadata to be effective, it must integrate directly into production systems and scale securely across live and archived workflows. MetadataIQ is designed as a cloud-first platform with on-prem components for secure media handling, enabling high-volume processing without disrupting existing operations.

Integration Options



Avid

- Metadata is written directly into Media Composer, MediaCentral, and Interplay PAM/MAM
- Information appears as native locators, markers, and searchable fields
- Editors and producers work entirely within familiar Avid interfaces



Grass Valley

- Deep integration with Framelight X
- Metadata delivered via watch-folder and API-based workflows
- Timeline-level alignment supports archive enrichment and ongoing ingest

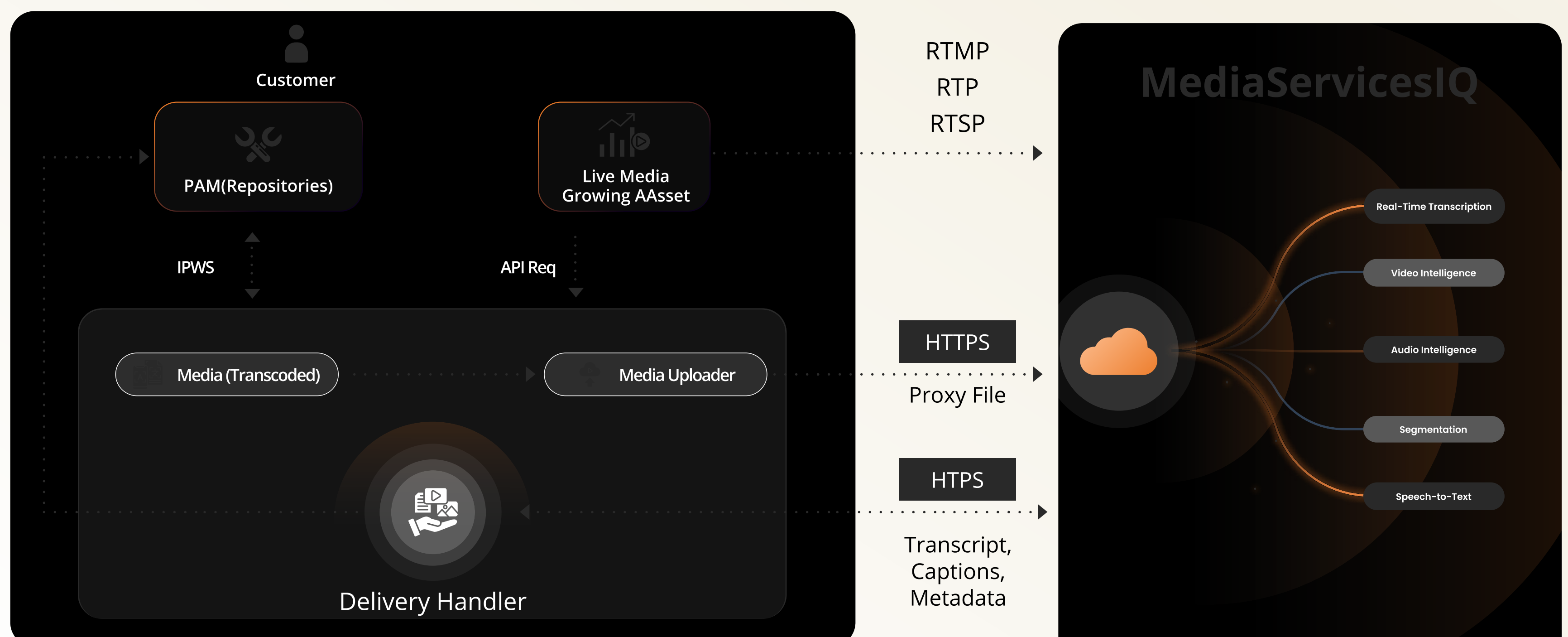
Open APIs

- Open endpoints support integration with other MAMs, CMSs, analytics platforms, and downstream systems
- Enables metadata reuse beyond core production environments

Deployment Model

- Cloud-first architecture for AI processing and scalability
- Lightweight on-prem components for secure media ingestion and control
- Supports real-time and batch workflows
- Media encrypted in transit; only metadata retained after processing

This architecture allows MetadataIQ to operate across live feeds, recorded media, and large archives while meeting operational and security requirements.



From Manual Review to Structured Intelligence

MetadataIQ changes how organizations work with video by replacing manual effort and fragmented metadata with a consistent, time-aligned understanding inside existing systems.

Before MetadataIQ

- Hours spent on manual logging
- Editors scrubbing long timelines
- Valuable content overlooked
- Compliance reviews are taking days.
- Metadata is fragmented across systems.

With MetadataIQ

- Automated, frame-level metadata generated in minutes
- Search and jump directly to key moments
- Hidden stories and reusable clips surfaced automatically
- Sensitive content was detected and reviewed quickly
- Metadata unified across platforms

The Value: MetadataIQ delivers measurable operational impact



Significant reduction in manual logging and tagging effort upto 90%



Faster editorial turnaround, supporting near-real-time workflows



Improved content reuse, helping surface value from existing libraries

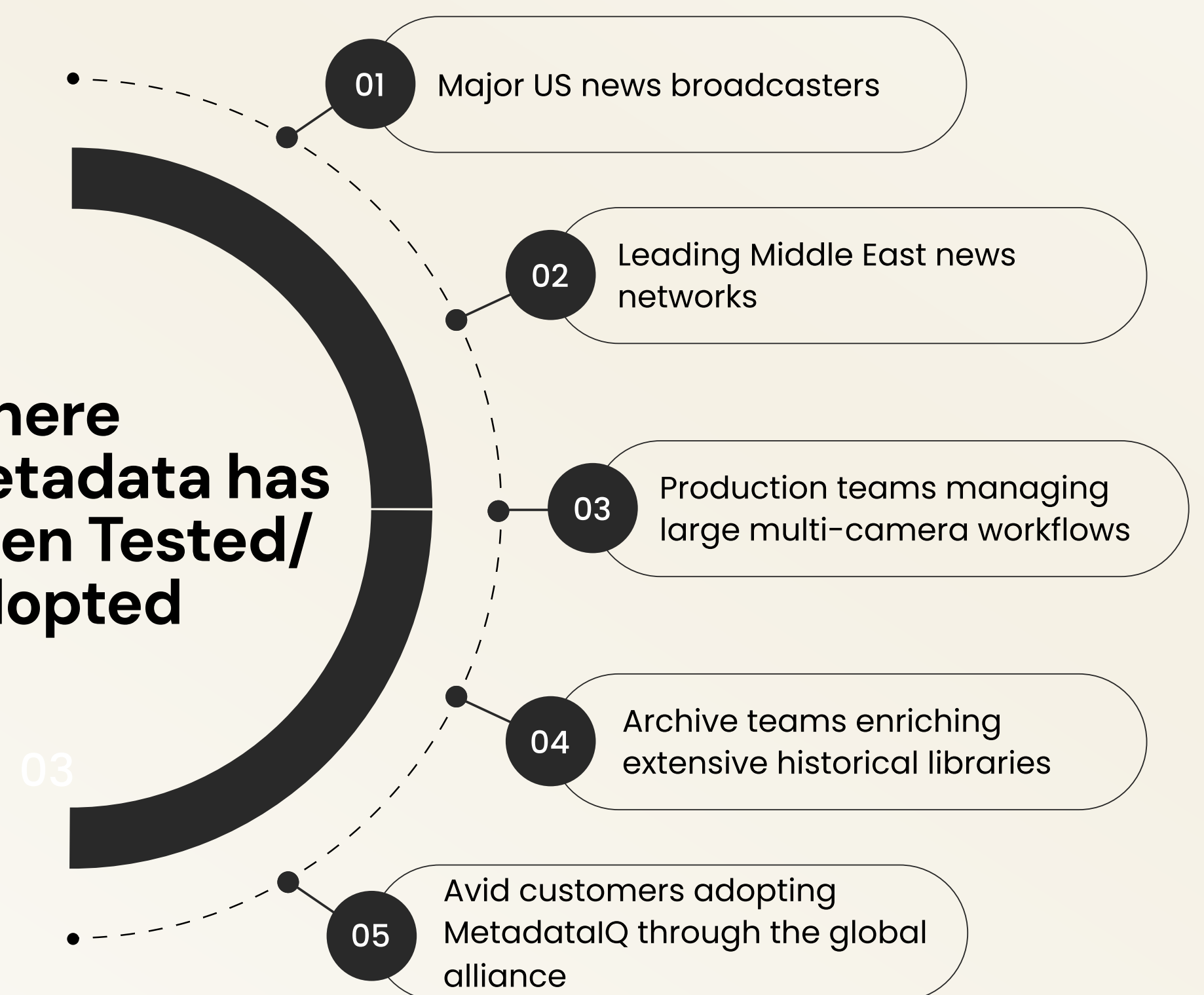


Direct enablement of AI and machine learning datasets from video



Context-aware compliance tagging with minimal manual intervention

Where Metadata has been Tested/Adopted



Built for Scale. Designed for Real Workflows.

MetadataIQ is designed for organizations that operate at scale, under time pressure, and inside complex broadcast and production environments. It is not a standalone experiment or a point solution. It is a production-grade system built to operate continuously, integrate deeply, and support real editorial and operational decisions.

This approach is best suited for organizations that

- **Ingest large volumes of live or recorded video daily**
- **Rely on Avid or Grass Valley as core production platforms**
- **Need faster discovery, turnaround, and reuse without adding manual effort**
- **Are preparing for AI-driven workflows that depend on structured video data**

MetadataIQ fits environments where accuracy, consistency, and integration matter more than novelty.

Who Should Consider This Workflow

- **News and broadcast organizations managing multiple live feeds**
- **Sports production teams operating under rapid turnaround cycles**
- **Entertainment and factual production teams handling long-form footage**
- **Archive and media asset teams responsible for large historical libraries**
- **Technology teams enabling AI, analytics, and automation initiatives**

About Digital Nirvana

Digital Nirvana has worked with broadcast and media organizations for more than two decades, delivering technology that supports monitoring, metadata automation, compliance, and media operations at scale.

The company's focus has always been practical: building systems that fit into existing workflows, perform reliably under real-world conditions, and help teams work more effectively with video.

MetadataIQ is part of a broader Digital Nirvana portfolio designed to help organizations understand, manage, and extract value from video across its lifecycle from live ingest and production to archive, compliance, and reuse.

Ready to explore how this workflow fits your environment?

[Request a workflow discussion](#)