VIDEO INTELLIGENCE FOR THE MEDIA & ENTERTAINMENT INDUSTRY

The AI Playbook To Unlock Success Within Your Content
| 01 | Introduction |
| 02 | Why the Media & Entertainment Industry Needs to Explore Video |
| 03 | Understanding Video Intelligence Through ‘Content Entities’ |
| 04 | Metadata Matters. However, It Doesn’t Solve Everything |
| 05 | Content Entities: Metadata to “Meaningfully” Describe Content |
| 06 | Building Video Intelligence Applications |
| 07 | Success Stories |
| 08 | Top Video Intelligence Applications |
| 09 | Content\textsuperscript{AI}: Building AI-Powered Content Value Chains |
The need to understand viewership trends and patterns is in rising demand. Technologies that emphasize viewership-based personalization are driving both engagement and monetization.

According to a Deloitte report in 2022, subscription churn averages around 35% in the USA market. As a result, content creators/producers need to provide more personalized ways to appeal to consumers and stay ahead of the competition. Media houses must aim to speak to audiences by analyzing and acting on data based on viewership patterns.

For the M&E industry, saving time is of paramount importance. For teams to deliver high-quality content and focus on the creative aspects, it is imperative to automate processes and reduce manual efforts. Custom-built AI-ML solutions such as video deduplication assist in enhancing the productivity of content storage teams by automatically detecting duplicate videos and efficiently managing content archives.

Content hyper-tagging is vital for effective management as it makes searching content easier by using meta tags. In addition to that, it can help monetize archived content and improve its reusability. Custom-built AI can further help generate relevant metadata by combining technology with business needs.
Why the M&E Industry Needs to Explore Video Intelligence?

Let’s look at how video intelligence helps key stakeholders from across the media value chain:

- A brand needs to do contextual advertising.
- A TV/Web series creator wants to understand which characters/story arcs are responsible for maximum engagement in a show.
- A sponsor needs to measure the screen time and screen share of its advertisements.
- Editors need to search for specific content at an instant and retrieve it.
- Broadcasters want to automatically identify and rate videos with explicit content.
Using Video Intelligence models over content allows content owners to do all of the above and much more. M&E organizations have the following broad requirements around their content archives, and **Customized Video Intelligence** finds applications in all of them.
Video Intelligence involves combining computer vision with business acumen to solve niche media industry challenges. Depending on the content - TV Shows, News or Web series - AI offers a contextual description of content through content entities and uses this understanding to solve the content objectives.

What is a content entity?

Content entities refer to people, objects, location, dialogues, keywords, or any useful concept that meaningfully describes the constituents of content. This applies to various forms of content. With the recent developments in AI and cloud technology, businesses can now process content at scale. Media houses can generate highly granular time-coded streams of content entities (metadata) that explain the content structure with the desired details. This metadata is further used to create a meaningful taxonomy of content entities.
Metadata Matters. However, It Doesn’t Solve Everything

It is necessary to understand that video intelligence isn’t only about generating volumes of metadata. Its purpose is to create meaningful metadata that describes your content per the desired use case. The effort of content entity creation must justify its business utility.

Metadata generated by AI may not always be relevant to the use case and may have pockets of contextual inaccuracy. As a result, post-processing techniques are used to improve the overall relevance of the metadata once content entities are created. Data quality parameters are defined to keep track of the overall quality of the metadata generated. Some post-processing techniques used are as follows:

1. In identifying scene boundaries, additional scene cuts can get introduced by unexpected visual artifacts. Triangulating this information with sharp dips in the audio profile can be an excellent indicator of a potential scene boundary.

2. When character cluster tracks are created, there are instances where a character appears on the screen bearing a couple of camera cuts in between. Considering empirical thresholds to normalize the data even if the character does not appear on the screen often gives better results.
Content entities like faces and objects detected in a non-central position with a tiny region of interest are neither relevant for analysis nor for editorial use, and can be filtered out.

Some content entities like camera angles can be self-curated using the meta context around other entities. Camera angles, closeups, and medium shots can be curated by looking at the screen share and screen placement of the detected faces at specific camera angles.
Relevant metadata that uniquely defines video content can be broadly classified into three categories:

1. **Time-Specific Metadata to Logically Define a Scene**

   This involves defining a scene as a temporal entity. Scene identification involves determining the contiguous camera cuts bearing a similar narrative. AI identifies them by picking sharp changes in the color histogram profile of two adjacent frames. A camera cut doesn’t always mean a change in the scene. For example, even when the camera changes between two people during a conversation, the scene remains the same. This temporal metadata helps identify contiguous video cuts that represent the same narrative.

   **How are the scenes identified?**

   The similarity of two adjoining camera cuts in terms of the narrative (audio, visual and color profile) provides the context to identify homogeneous groups of camera cuts and that defines a logical boundary of the narrative - the scene boundary.
This class of metadata concerns itself with people. It includes identifying character-based information that drives the narrative in a particular scene. This includes tagging people, emotions, celebrities, and character pairs. These classes of metadata may also represent age and gender information.

These content entities are the active foreground elements—the characters or the people performing in the content. The detection of these entities involves identifying key characters that drive the narrative in a particular scene. In a movie, it could be the protagonist, in a game of soccer, it would be the players, and in a news debate, it’s the anchor and the panelists.

However, defining the content entities using meaningful relationships and parameters is more important than merely identifying the people present in the scene. AI creates this by picking frequent overlaps of screen presence and layering in contextual information.

Here are some examples of meaningful tracks of characters.

- Relationship Tracks - e.g., father-son, protagonist-antagonist, lead couple
- Age and Gender Tracks - e.g., family with kids, kids with grandparents, teenagers
- Gameplay Tracks - e.g., arch-rivals, team-mates, team celebration
- Studio Panel Tracks - e.g., Presidential nominees, political rivals, a celebrity couple
Characters: Jon Snow & Ygritte
Camera Angles: Low-Level Shot
Set Locale: Dubrovnik, Croatia
3. Elements of Interests and Background Information

These meta-tags contain background information about the video content. Here the content entities constitute the backdrop of the story. These are passive elements in the overall viewership experience. This is the canvas on which the foreground elements operate.

<table>
<thead>
<tr>
<th>MetaTags</th>
<th>Content Entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera angles</td>
<td>Close-up, long shot, wide-angle</td>
</tr>
<tr>
<td>Set locale</td>
<td>In-studio, indoor, outdoor</td>
</tr>
<tr>
<td>Objects of interest</td>
<td>Car, ball, trees, doors, windows, red-yellow cards, brand logos, trophies</td>
</tr>
<tr>
<td>Screen graphics</td>
<td>Split screen, scoreboard, text elements, credits, transitions</td>
</tr>
<tr>
<td>Overall ambiance</td>
<td>Music, silence, high decibel drama, color palette and lighting</td>
</tr>
</tbody>
</table>

© 2022, ContentAI: Video Intelligence For Media & Entertainment
Quantiphi has combined computer vision, deep learning, and AI techniques to develop a bleeding-edge metadata tagging solution to make content more discoverable.

**Training: Custom Model Development**
- Video tagging and generation of custom tags and data
- Training models on a custom-built AI platform
- Customization of pipelines based off use cases and business requirements

**Inferencing: AI-ML Inferencing for Media Applications with Contextual Metadata**
- Extraction of valuable metadata about the video and its content
- Processing the metadata to conform to the predefined business rules for a use case/application
- Integration with Media Asset Management (MAM) systems to write metadata and make archives smart
- Active learning feedback loops to customize the pipeline for ever-changing requirements/rules and newly added tags/data

**Management: One-Stop-Shop Platform for Media Workflow Operations**
- Application of meta-tags on the content
- Enabling deep metadata-based search
- Easy content editing, repurposing, and publishing
- Integration into mission-critical content analytics workflows
Enriching Content Archives with Searchable Video Metadata

Problem Context
Sportscast, a subsidiary of Deutsche Fußball Liga (DFL), needed AI enriched metadata to archive decades of football content.

Quantiphi Solution
Quantiphi developed a bespoke video intelligence platform to generate editorial-worthy moments in the DFLs content library.

Impact
- 400+ hrs of historic content processed via automatic archiving
- 1500+ players and DFL members added to the platform
- 1.5M data points encompassing league players, camera shots, logos and custom entities

Video Deduplication and AI-powered Content Management

Problem Context
A media house was facing high content archiving costs across cloud and on-prem servers due to duplicate videos with minute differences being created at every stage of the content lifecycle.

Quantiphi Solution
Quantiphi leveraged audio and computer vision-based feature extractors to build an automated solution that identifies unique logical segments in the content and deletes duplicate content stored in the clients’ archive.
AI-Powered Closed Captioning for Live News Broadcasts

Problem Context
The client, a US broadcasting major with eleven news broadcasting studios, wanted real-time transcriptions on their live broadcasted news feeds on the web and pre-recorded video news segments.

Quantiphi Solution
Quantiphi developed a scalable web-based solution that consisted of closed captions obtained using speech-to-text API synced with the video feed. The solution was trained on news content to maximize the accuracy of created captions.

Impact
- Completed transcriptions for recorded and live stream videos on clients’ portal
- Over 85% transcription accuracy with zero glitches

Content Monetization and Audience Intelligence

Problem Context
Alphonso, a TV data company, wanted to determine the switching pattern of viewers and predict users that would lapse for a show. They needed to assign propensity scores to users with a greater tendency to watch a new show for user group segmentation.

Quantiphi Solution
Leveraged statistical analysis and ML-driven clustering to generate viewership patterns and understand user switching behavior to hyper-personalize ad campaigns.

Impact
- Achieved real-time viewership preferences leading to effective campaign optimization
Management and Archiving

Every M&E enterprise strives to save time, cost, and effort in managing content. The existing manual tagging process is effort-intensive, expensive, and time-consuming. Enabling a deep search of useful micro-moments within content unlocks new ways to manage content. Here are a few use cases:

- Characters/emotions detection
- Camera angles/shot locale
- Custom entities and events
- Specialized meta-data architecture for news, media, and entertainment
- Video deduplication to save content storage costs
- AI-powered workflows to repurpose/reuse desired content moments

Distribution and Compliance

Video intelligence solutions accelerate the process of content publishing and distribution by automating processes in accordance with constantly changing OTT rules and regulations. These include the following use cases:

- OTT compliance
- Subtitle/audio conformance
- AI-assisted content moderation
- Automated closed captioning and localization
Audience Intelligence and Monetization

Contextual Advertising for Ad-Effectiveness

Video intelligence plays a vital role in understanding the audience’s behavior and creating a targeted approach. Key metrics and insights can be leveraged to enhance the audience data at hand.

In today’s economy, ads are considered an unpleasant interruption in the viewing experience. However, that is not true for contextual advertising. Video intelligence enables OTT platforms to target specific moments in line with the content and prevent tone-deaf ads. These include the following use-cases:

- Contextual ad-placements
- Determining ad-compliance
- Brand visibility for media and television
- Hyper Commercialization: Blending OTT with e-Commerce

Deep Content Insights

While audience metrics provide details about content performance, they do not give the reasons for it. A better understanding of deep content parameters bridges this gap by helping creative teams make data-driven decisions in their storytelling. Mapping audience behavior with latent content parameters helps with:

- Reach summary
- Describing trend and genre analysis
- Generating character resonance
- Depicting audience demographics and behavior
- Defining prominent content themes/story plots
- Social engagement

Content Personalization

- Create and recommend the content of choice
- Content-based recommendation engines
Quantiphi’s cutting-edge AI capabilities across the media value chain empower content owners with ample tools to innovate and differentiate. Our media intelligence offerings help media houses seamlessly adopt AI into their workflows and improve productivity, produce higher quality content, streamline operations, uncover deep content insights and unlock scalable revenue-generating business models.
Launch Partner For AWS Media Intelligence Services 2021

Google Cloud Partner of the Year Awards - Data Analytics, 2021

Americas Service Delivery Partner of the Year, 2021

AI-powered Video Deduplication -Winner ‘Store’ Category, 2021

Leader in the Forrester New Wave™: Computer Vision Consultancies Q4, 2020

Leader in IDC MarketScape: Worldwide AI IT Services 2021

Solutions Portfolio

Conversational AI

Speech Interfaces

Document AI

Visual QA

Intelligent Voice Automation

Recommendation Engine

Semantic QnA

Time Series Forecasting

Predictive Analysis

Experience

Sight

Sound

Language

Pattern

Platform Solution

Security & Access Management

Containerization & Deployment

Infrastructure Modernization

Data Lake & EDW Modernization

Cloud Migration

Data Integration

Data Consolidation & Reporting

Data Visualization

Business Intelligence

Big Data

Data Solutions

Insights
Quantiphi is an award-winning AI-first digital engineering company driven by the desire to reimagine and realize transformational opportunities at the heart of business.

Visit: www.quantiphi.com

Amsterdam | Boston | Bangalore | London | Mumbai | Princeton | San Jose | Toronto | Trivandrum