

---

---

# **Beyond the Hype: A Critical Look at AI in Video Streaming**

Jan Ozer  
Streaming Learning Center  
jan.ozero@streaminglearningcenter.com

---

---

# Online Course for New Streaming Professionals

A promotional banner for the 'Streaming Media 101' course. It features a circular inset image on the left showing three men in a professional setting; one is seated at a desk with a laptop, while two others stand behind him, looking at the screen. To the right of the image, the text 'Streaming Media 101' is displayed in a large, white, sans-serif font. Below this, the subtitle 'Technical Onboarding for Streaming Media Professionals.' is written in a smaller white font. At the bottom right of the banner, there is a blue button with the text 'Click to view 12 free lessons' in white.

**Streaming Media 101**

Technical Onboarding for Streaming Media Professionals.

Click to view 12 free lessons

## Streaming Media 101: Technical Onboarding for Streaming Media Professionals

Learn the technologies, techniques, and skills to succeed in a streaming media-related role, whether it's producing and distributing streaming video or creating the tools and services necessary to do so.

<https://bit.ly/StreamingMedia101>

## What You Will Learn

In about 11 hours, this online course will teach you the terms, technologies, best practices, and skills needed to excel in a technical role in the streaming media industry. You will learn:

- ✓ How to encode and deploy streaming video using the H.264, HEVC, VP9, and AV1 codecs
- ✓ How to encode for single file and adaptive bitrate encoding and packaging for HLS, DASH, and CMAF
- ✓ About digital rights management (DRM) and distribution issues like choosing a CDN and how to measure and ensure Quality of Service and Quality of Experience
- ✓ Critical production-level decisions, like whether to encode on-premise or in the cloud, how to choose a per-title encoding technology and cloud encoder, and how to compute the breakeven on deploying an advanced codec like HEVC or AV1

## You will learn to:

- ✓ Analyze files with MedialInfo, Bitrate Viewer, Apple's AVQT, and the Moscow State University Video Quality Measurement Tool
- ✓ Encode in FFmpeg and Handbrake
- ✓ Produce mezzanine files for upload to a streaming service
- ✓ Connect to YouTube Live and Facebook Live
- ✓ Embed a live or on-demand video into a web page

# Agenda

- How to consider AI in your buying decisions
- Survey of AI applications
- Summary and conclusion

**The Best Way to Consider AI in  
Streaming-Related Product  
and Services is to ...**

**The Best Way to Consider AI in  
Streaming-Related Product  
and Services is to ...**

**Ignore it completely**

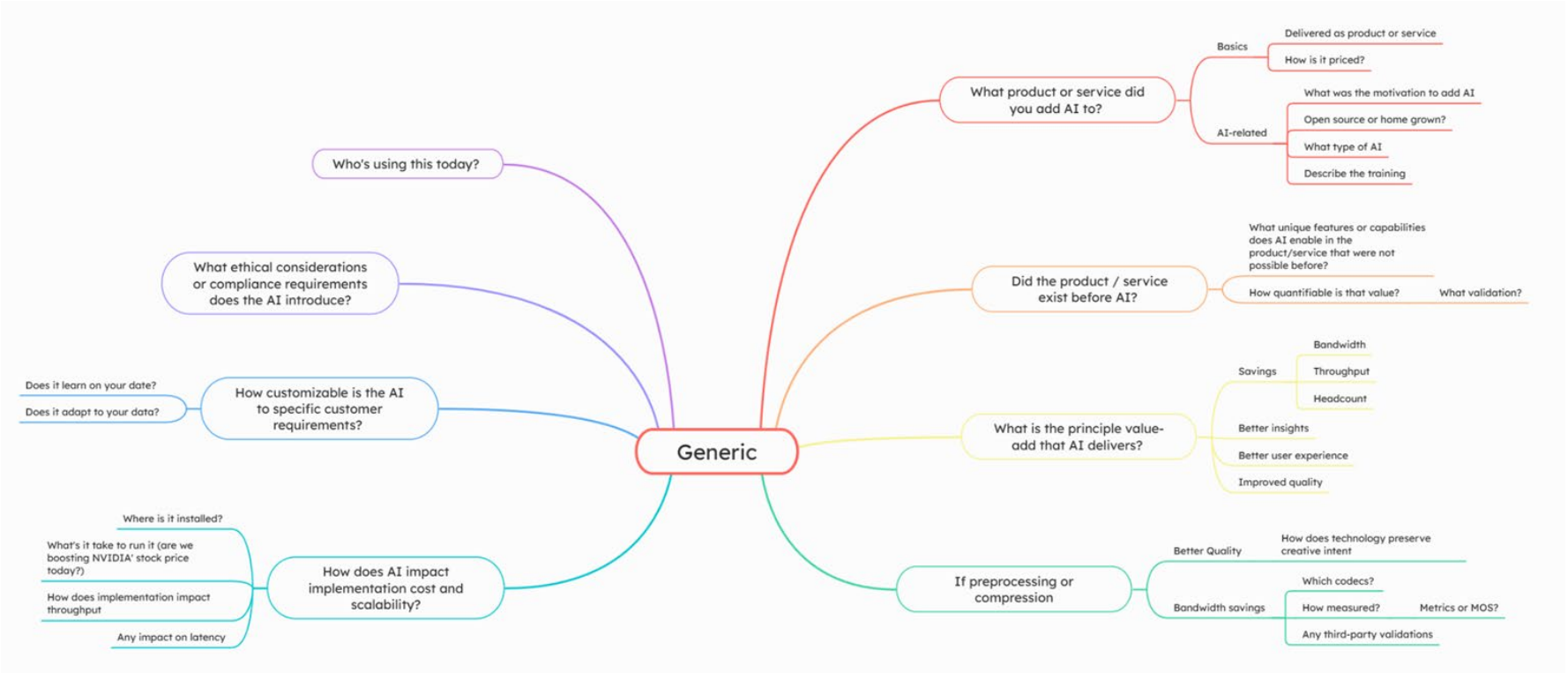
# Why?

- Not to say that AI hasn't already made significant contributions, which will only grow in the future
- But, you buy a whole product/service
- Doesn't matter if it's powered by AI or a gasoline engine;
  - If it delivers the required ROI or equivalent measure you use to make purchasing decisions, buy it
  - Not because it says AI
- Evaluate the whole product like you always did
  - If interested in AI, here are some suggested questions

# I Was Interested to See How AI Was Being Used

- Announced the session on LinkedIn
- Interviewed any interested company with deployed AI-based-technology from preprocessing through distribution
  - Spoke to 12, a tiny fraction of companies implementing AI in the space
  - Tried to understand what they were doing with AI
  - Recorded interviews – available on YouTube
- Developed questionnaire that you might find useful

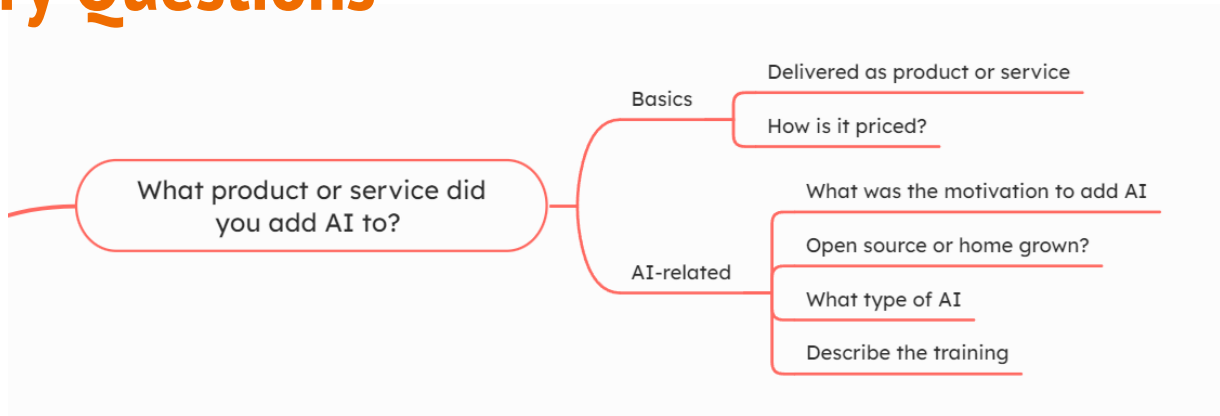
# My Evolved Question List (after speaking with 11 companies)



[https://bit.ly/AI\\_mindmap](https://bit.ly/AI_mindmap)



# Introductory Questions



- Basics

- Product or service - how delivered?
- Pricing

- Why add AI?

- **Open source or home grown**

- If open source, what's the value add that differentiates from other services (Ampere/NETINT/Whisper- throughput)
- Open source = minimum control
- Home grown requires more engineering but should deliver competitive advantage

- **What type of AI**

- ML - Machine learning analyzes data to find patterns and make accurate predictions
- Generative AI creates new data that resembles the training data
  - Interactive and free-form

# Introductory Questions

Did the product / service exist before AI?

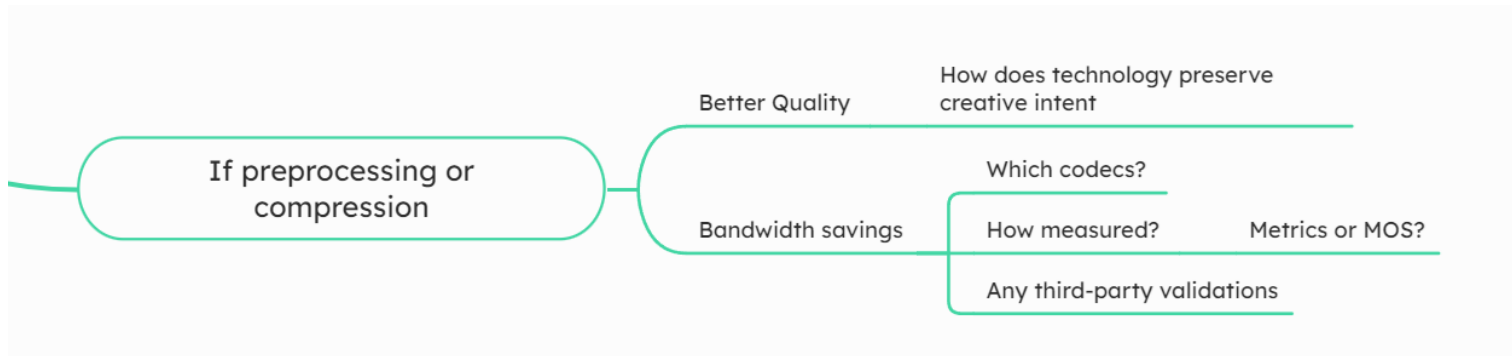
What unique features or capabilities does AI enable in the product/service that were not possible before?

How quantifiable is that value?

What validation?

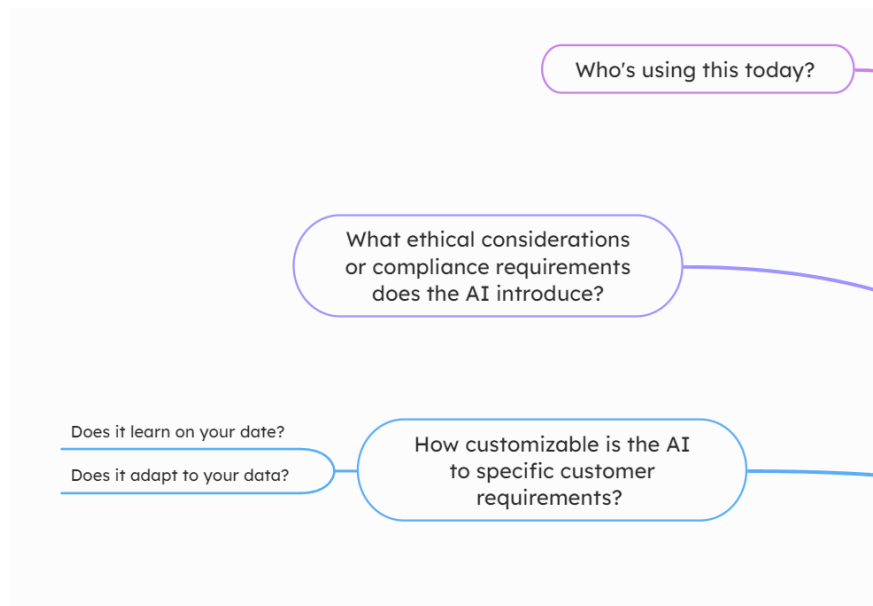
- Today, most products existed without AI, so it was incorporated to add features, performance, or both
- Try to identify key benefit:
  - Is it quantified or general?
  - Is it validated by a third-party
    - Usually a customer

# If Compression



- More questions about added value
- Compression claims are typically vague and unsupported
  - That's why you run your own tests
- **If preprocessing, "improve quality" claims**
  - Typically optional - but verify that you can disable?
  - **How to preserve creative intent?**
  - Hacking? Check VMAF/PSNR/SSIM
- **Bandwidth savings**
  - Which codecs? Impact of H.264 likely different than AV1 or VVC
  - How measured?
    - Objective - VMAF/PSNR
    - Subjective MOS (the gold standard, but expensive)
    - Typically not validated by third-parties other than customers

# Misc Questions



- Does the AI adjust to your data (or file inputs)?
  - Seldom does
- Any privacy issues?
- Who is using - get third party validations

# Bottom Line with Compression:

- Questions must be asked, but aren't terribly useful
  - Very few academic peer reviewed studies
  - Even if they did exist, you care about:
    - Your footage
    - Your codecs
    - Your encoding parameters
    - Never an apples-to-apples mix
  - When it comes to compression-related technology, you're going to have to test
- And again,
    - AI is only a component
    - Almost impossible to discern from other techniques
    - You care about whole product performance

# AI Companies (and Others) That I spoke With

## Preprocessing



- Digital Harmonics
- VideoOn

## Encoding



- Harmonic
- Visionular
- Bitmovin (Super-Rez)
- Codec Market
- Deep Render
- InterDigital

## Captions



- Interra
- See also:**
- Nanocosmos
- NETINT
- Ampere

## Quality



- IMAX

## Customer Experience



- Media Distiller
- See also:**
- Amagai

## Analytics



- Bitmovin
- Operative

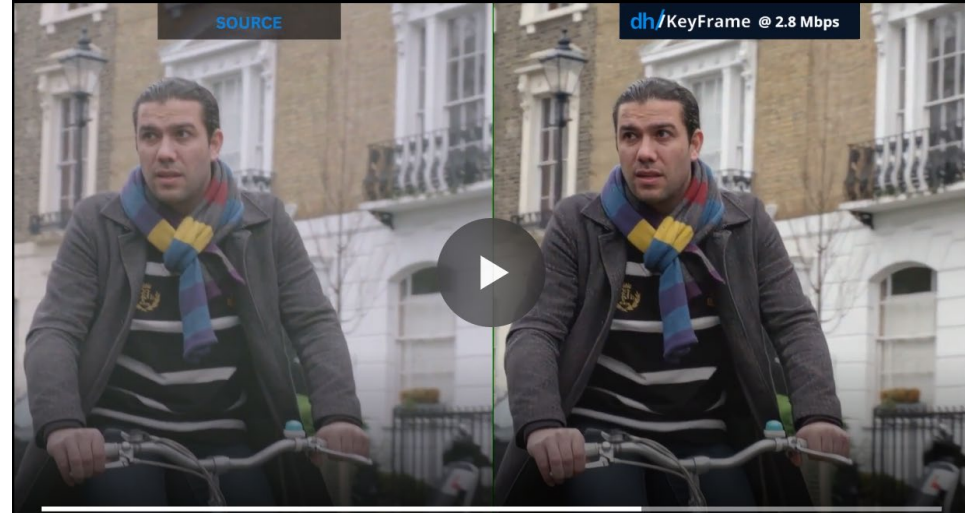
## Ad Ops



- Operative

# Digital Harmonic - Keyframe

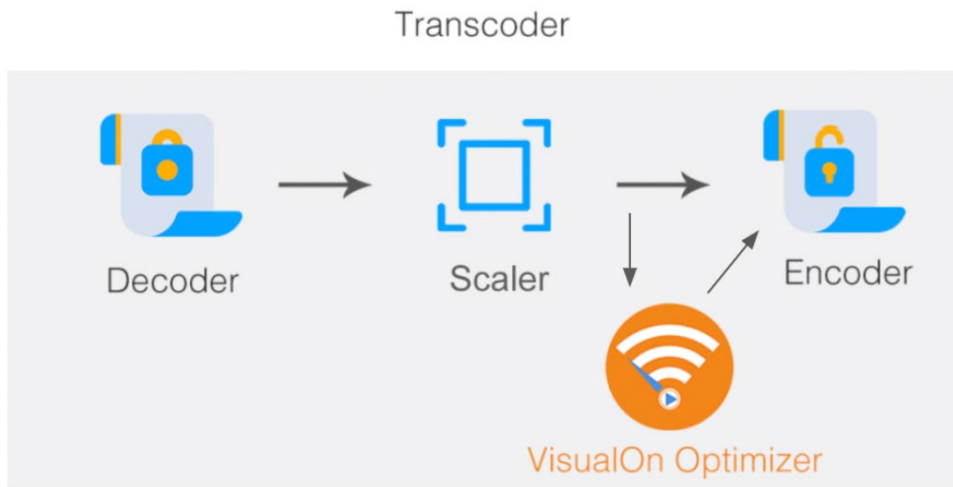
- **What:** Preprocessor with both quality improvement and bandwidth reduction
- **Product or service:** Product
- **Exist before AI:** Yes
- **Benefit of AI:** Improved performance in both modes
  - Adaptive entropy minimization
  - De-artifacting, De-noising, and De-blocking
  - Anti-aliasing
- **Bold claim:** Up to 80% bitrate reduction with no loss in quality (PSNR/MOS)
  - Quality improvement to better than source
- **Which codecs:** Agnostic



- **Implementation:**
  - Installed pre-encoder (bump in the wire)
  - Needs GPU-based system to run
  - Throughput varies with processing (quality/bandwidth reduction)
- **For more information:**
  - <https://www.dhkeyframe.com/>
  - Interview - [https://bit.ly/LI\\_DH](https://bit.ly/LI_DH)

# VisualOn Optimizer

- **What:** Preprocessor with both quality improvement and bandwidth reduction (CAE)
- **Product or service:** Product
- **Exist before AI:** Yes
- **Benefit of AI:** Improved performance in both modes
- **Bold claim:** Up to 70% bitrate reduction with no loss in quality (VMAF)
  - Quality improvement to better than source
- **Which codecs:** Agnostic



- **Implementation:**
  - Integrates via API
  - Runs on the same system as the encoder, and doesn't negatively affect throughput
- **For more information:**
  - [www.visualon.com](http://www.visualon.com)
  - Interview: [https://youtu.be/B9BI\\_YdAj78](https://youtu.be/B9BI_YdAj78)



# Harmonic: PURE/EyeQ

- **What:** AI-based compression technology used by ~ 100 customers
- **Product or service:** both
- **Exist before AI:** PURE yes/EyeQ no
- **Benefit of AI:** Improved compression efficiency
- **Bold claim:** 50% greater efficiency than open source
- **Implementation:**
  - Appliance
  - Cloud service



- **For more information:**
  - <https://www.harmonicinc.com/>  
Interview video:  
<https://youtu.be/tPDPcN21CNg>

# Harmonic Claims 50% OTT Performance Improvements with EyeQ

WAG 1-0 LAV 44:53

EyeQ

What Lattie terms *"in-loop artificial intelligence"* guides the codec to focus bits where and when they matter. "This leverages the mechanics of the human visual system to continuously assess video quality in real time and focus bits where and when they matter most for the viewer," he says. "Unnecessary bits can be eliminated, allowing content providers to consistently meet video quality targets at the lowest possible bandwidth."

[https://bit.ly/harm\\_eyeq](https://bit.ly/harm_eyeq) - 9/8/2016

- **What:** H.264, H.265, and AV1 codecs
- **Product or service:** Both
- **Exist before AI:** No - integrated from the start
- **Benefit of AI:** Improved compression efficiency
- **Bold claim:** Up to 50% bitrate reduction compared to open source
- **Implementation:**
  - License codecs
  - Use cloud service

# Smaller Videos. Great Quality. A.I. Inside.

- **For more information:**
  - [www.visionular.com](http://www.visionular.com)
  - Interview - <https://youtu.be/k44QdGMjN5U>

September 2021

# Advances in Video Compression Systems Deep Neural Networks Review and Case Studies

*This article targets video coding improvements for both in-loop and cross-loop video encoder and jointly across multiple blocks including reviews various modules in video coding that could benefit from neural networks and provides case studies for each of these modules*

By DANDAN DING<sup>id</sup>, Member IEEE, ZHAN MA<sup>id</sup>, Senior Member IEEE, DI CHEN<sup>id</sup>, Member IEEE, QINGSHUANG CHEN, Member IEEE, ZOE LIU, AND FENGQING ZHU<sup>id</sup>, Senior Member IEEE

**Zoe Liu** received the B.E. degree (Honors) and the M.E./Ph.D. degree from Tsinghua University, Beijing, China, in 1995 and 2000, respectively, and the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 2004, all in electrical engineering.

She was a Software Engineer with the Google WebM Team, Mountain View, CA, USA, and has been a key contributor to the newly finalized royalty-free video codec standard AOM/AV1. She has been devoted to the design and development of innovative products in the field of video codec and real-time video communications for almost 20 years. She is the Co-Founder and the President of Visionular Inc., Palo Alto, CA, USA, a startup delivering cutting-edge video solutions to enterprise customers worldwide. She was a 2018 Google I/O speaker. She has published approximately 50 international conference papers and journal articles. Her main research interests include video compression, image processing, and machine learning.



<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9369668>

# Codec Market

- **What:** Integrated cloud platform (encode/player/CMS/CDN)
- **Product or service:** Service
- **Exist before AI:** No - integrated from the start
- **Benefit of AI:** Improved compression efficiency
- **AI Component:** Open-source advanced implementation of VMAF
- **Bold claim:** 30% more efficient than open source
- **Implémentation**
  - Use cloud service

## Machine Learning

Codec Market leverages machine learning techniques that optimize secure encoding at scale for video-on-demand libraries and, soon, live-streamers

- **For more information:**
  - <https://www.codecmarket.com/>
  - Interview - <https://youtu.be/Dme1Q2fYEuE>

# Bitmovin Super-Resolution

- **What:** Scaling at ultra-high quality
- **Why:** Scale low resolution source content to HD and higher for superior quality distribution
- **Product or service:** Service
- **Exist before AI:** Yes
- **Benefit of AI:** Improved quality
- **Implementation**
  - Component of transcoding service



VIDTECH

## Video Tech Deep Dive: Super-Resolution with Machine Learning P1



Jameson Steiner

20 May 2020 · 5 min read

- **For more information**

- <https://bitmovin.com/encoding-service>
- Interview - <https://youtu.be/rTm26JRIkBQ>

# Bitmovin



|Bicubic|



|Super-resolution|

# Deep Render

DEEP RENDER

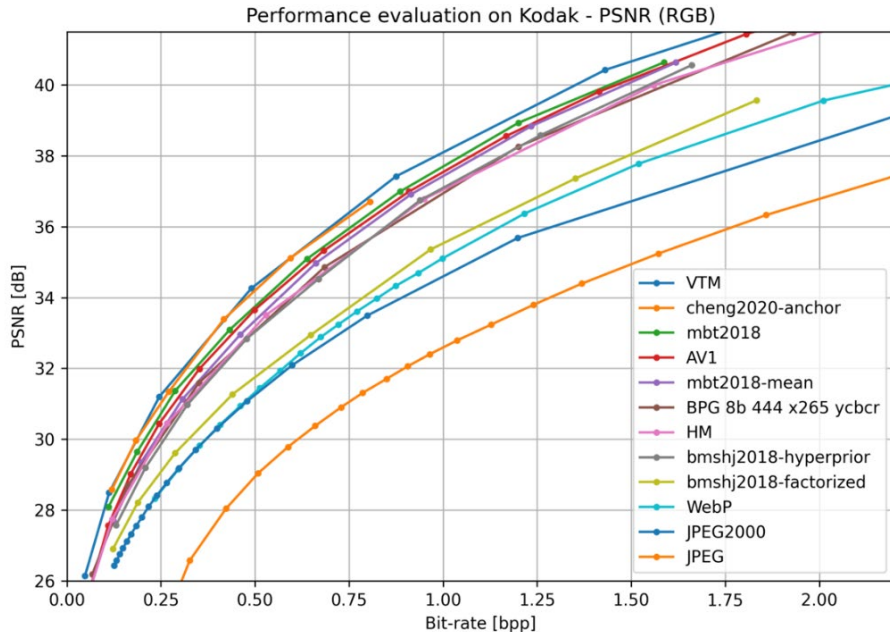


- **What:** “DeepRender is the world's only AI -based codec company”
- **Product or service:** product
- **Exist before AI:** No
- **Bold claim:** 45% greater efficiency than VVC
- **Implementation:**
  - License
- **For more**
  - [www.deeprender.ai](http://www.deeprender.ai)
  - Interview:  
<https://youtu.be/L6cy6MyGkR4>

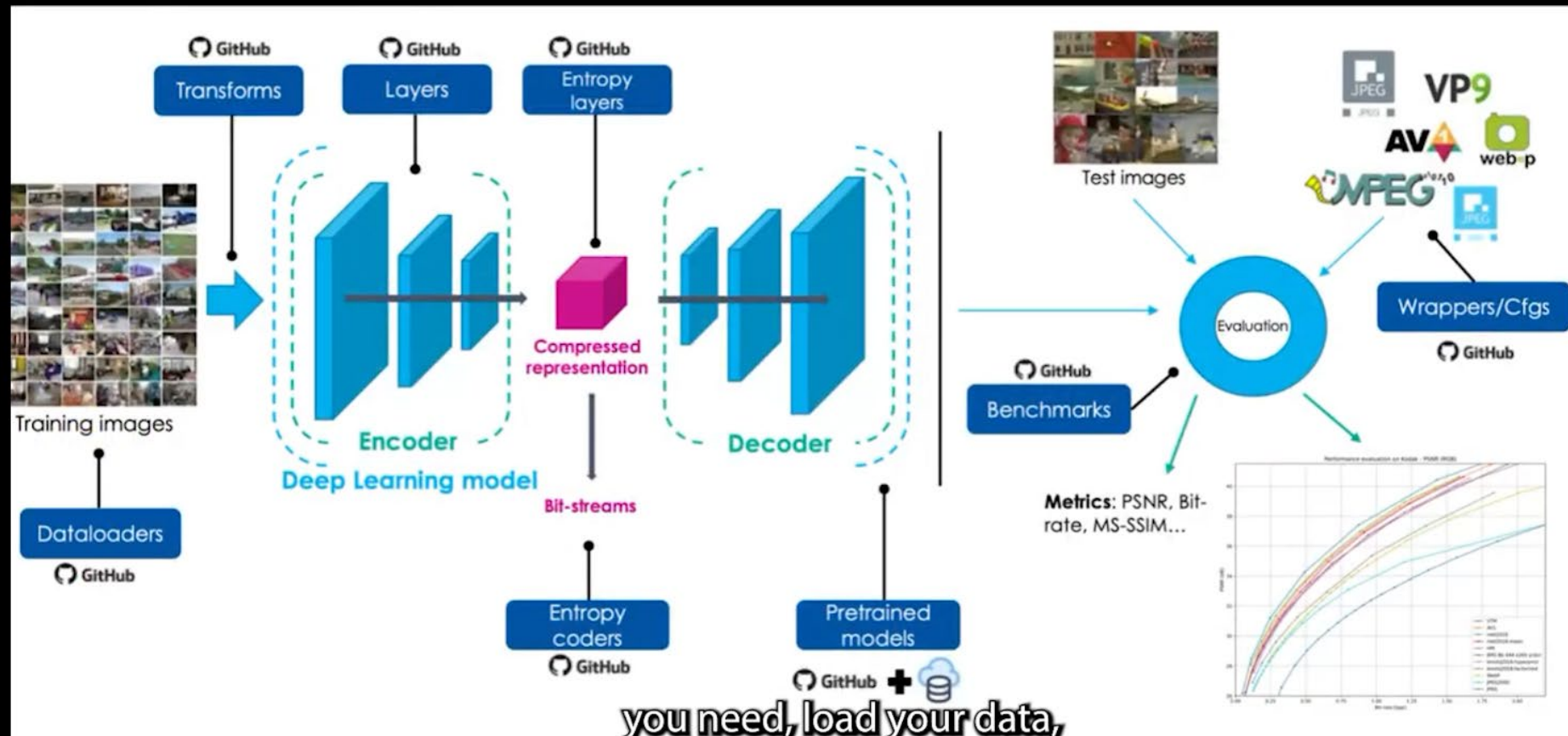


# InterDigital Compress AI

- **What:** A PyTorch library and evaluation platform for end-to-end compression research
- **Product or service:** Open source toolkit
- **Exist before AI:** No
- **Which codecs:** Agnostic
- CompressAI currently provides:
  - Custom operations, layers and models for deeplearning-based data compression
  - A partial port of the official [TensorFlow compression](#) library
  - Pre-trained end-to-end compression models for learned image compression
  - Evaluation scripts to compare learned models against classical image/video compression codecs



- **For more information:**
- [Download from Github](#)
- For more
  - Interview video: [https://youtu.be/nn9TGr\\_T8kY](https://youtu.be/nn9TGr_T8kY)



**you need, load your data,**



compressAI enables researchers to quickly design, train, test and evaluate AI-based codecs.

<https://github.com/InterDigitalInc/CompressAI>



## compressAI provides:

- A Pytorch library (compression modules: entropy bottlenecks, entropy coders) to simplify and accelerate AI-based video compression research
- An evaluation platform for apples-to-apples comparison of your compression model with state-of-the-art traditional and AI-based codecs.

**All right.**

# Interra - Baton Captions



- **What:** Content-Aware Captioning & Subtitling
  - **Product or service:** Product
  - **AI Source:** mix of open-source and home grown
  - **Exist before AI:** Yes
  - **Benefit of AI:** Faster, better, more features
  - **Bold claim:** Deploys AI multiple components: speech to text technology, natural language processing, speaker identification, scene change detection, burnt-in text detection, face detection, and machine translation.
- **For more information:**
    - Website: [www.interrasystems.com/BATON-Captions.php](http://www.interrasystems.com/BATON-Captions.php)
    - Interview video: <https://youtu.be/dp0Oah19ONo>

# Interra Captions

## Natural Language Processing



Others  
(Character count only)

They call Errol Spence "The Truth," and  
over the course of a steady rise the  
undefeated "28-year-old" has earned his  
nickname and a world title...



BATON Captions  
(Intelligent Segmentation)

They call Errol Spence "The Truth,"  
and over the course of a steady rise  
the undefeated "28 year old"  
has earned his nickname  
and a world title...

NLP delivers captions that  
read more naturally and are  
easier to understand

**Understands writer intent and sentiment**

# IMAX Viewer Score

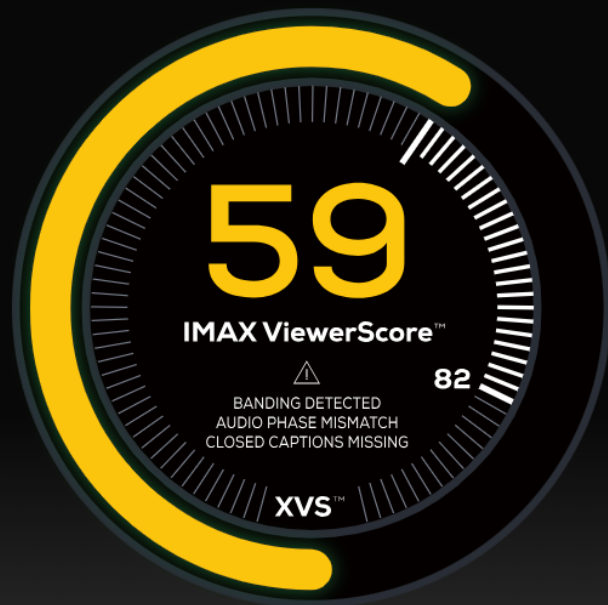
- **What:** No reference real time video metric (formerly SSIMPLUS)
- **Product or service:** Product
  - StreamAware – real-time quality monitoring and reporting
  - StreamSmart – dynamically adjusts stream quality
- **Exist before AI:** Yes
- **Benefit of AI:** More accurately matches human perception (up to 94% from 90%)



- **Bold claim:** Optimize encoding to reduce bitrate by 15% or more?
- **Which codecs:** Agnostic
- For more information
  - <https://www.imax.com/sct>
  - [interview - https://youtu.be/q6VdxaL651E](https://youtu.be/q6VdxaL651E)

Sees like a  
subscriber and a  
studio expert.

Knows exactly  
when quality and  
compliance is lost,  
and why.



### 'ANNOYING'

FAIR Picture Quality  
IMAX ViewerScore: 59



### QUALITY CHECKS:

Quality Assurance  
Quality Control  
Regulatory Compliance

- 0-100 scoring (same as VMAF/SSPLUS)
- HDR support
- Device support
- Mixed frame rate

# Media Distillery

- **What:** AI to deliver better video experiences and extended viewing time
- **Product or service:** Service
- **Exist before AI:** Yes
- **Benefit of AI:** Improved performance
- **Bold claim:** Our AI technology allows us distill topic, image, EPG correction and context from video content - all automatically and in real-time.

## Topic Distillery™

Topic Distillery™ enables topic-based discovery for any live broadcast content, allowing viewers to find content impossible to discover before and access it faster through a chapter-based playback.

Read more →

- **For more information**
  - <https://mediadistillery.com/solutions>
  - Interview - <https://youtu.be/iHqauz0rD7s>



# Bitmovin Analytics

- **What:** Supercharging Data Insights with AI for Video Analytics
- **Product or service:** Service
- **Exist before AI:** Yes
- **Benefit of AI:** Improved usability and utility
- **Implementation**
  - Added to analytics service

## AI Session Interpreter

quickly and a long startup time can lead to frustration and abandonment.

### Quality of Experience:

The user experience was relatively good. There were no buffering events or quality changes during the playback, which means that the video played smoothly without any interruptions. The lack of pauses also indicates that the user was engaged and did not need to stop the video.

### Quality of Service:

The quality of service was good. The video played smoothly without any interruptions and the average bitrate was 3000000 kbps, which is a high bitrate and indicates that the video quality was high.

## Recommendations

- **Reduce the video startup time.** This can be done by optimizing the video player for faster loading times or by providing users with more information about their internet connection speed.
- **Monitor the network conditions and adjust the video quality accordingly.** This will help to avoid buffering events and ensure that the video plays smoothly.

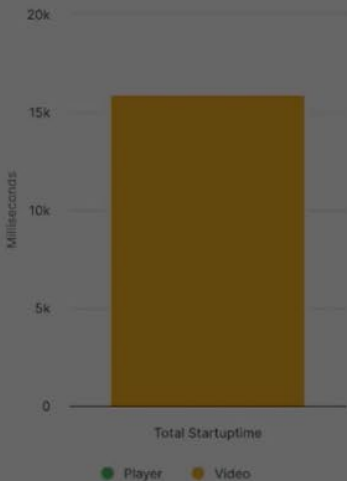
- **For more information**
  - <https://bitmovin.com/video-analytics>
  - Interview - <https://youtu.be/rTm26JRlkBQ>

- Home
- Streams
- VOD Encoding
- LIVE Encoding
- Player
- Analytics
- Getting Started
- Audience
- Quality of Experience
- Top Errors
- Monitoring new
- Advertising
- Alerts new

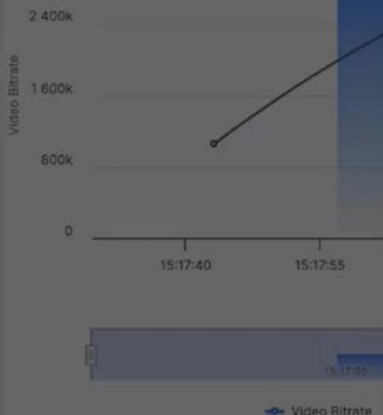
ANALYTICS / SESSIONS / DDDDDDD-F27A-4D0E-81DD-F120D9C714F4

Analytics Version: v2.17.2 Player Version: bitmovin-8.51.0 Operating System: Android 10

### Startup Time



### Session Visualization



User Information Sessions

## AI Session Interpreter

quickly and a long startup time can lead to frustration and abandonment.

### Quality of Experience:

The user experience was relatively good. There were no buffering events or quality changes during the playback, which means that the video played smoothly without any interruptions. The lack of pauses also indicates that the user was engaged and did not need to stop the video.

### Quality of Service:

The quality of service was good. The video played smoothly without any interruptions and the average bitrate was 3000000 kbps, which is a high bitrate and indicates that the video quality was high.

## Recommendations

- **Reduce the video startup time.** This can be done by optimizing the video player for faster loading times or by providing users with more information about their internet connection speed.
- **Monitor the network conditions and adjust the video quality accordingly.** This will help to avoid buffering events and ensure that the video plays smoothly.

 Good Response Bad Response Regenerate

### Feedback makes us better!

Our AI powers your insights, but it's not perfect - mistakes can happen. If you spot something off, please let us know. Your feedback is invaluable in helping us enhance your experience.

Every individual session is different has

# Operative

- **What:** Generative AI to automate proposal generation from texts, emails, and voice prompts
- **Product or service:** Service
- **Exist before AI:** Service yes, features no
- **Benefit of AI:** Intelligently speed sales tasks to optimize yield across linear and digital inventory
- **Implementation**
  - Added to Operative services

 ADELINE AI  
**YOUR NEW MEDIA COMPANION**

Meet Adeline, the assistive AI platform coming soon to Operative media solutions.

- To learn more:
  - <https://www.operative.com/adeline/>
  - Interview - <https://youtu.be/9JaappNDF2k>

# King.com Digital Proposal

Deal Status: MS - Proposal Draft | Deal ID: 228611 | Viewing Version: V1- Version 1 🏆 | Revision No: 0



Header Accounts Conditions Goals Workspace

Open Revision

Save Cancel View Approvals ▾ Copy ▾ Change Version ▾

Deal Name * King.com Digital Proposal	Platform Linear Off	Plan Class Select	Order Type Digital	Opportunity ID 0	Version Name* V1- Version 1
--	------------------------	----------------------	-----------------------	---------------------	--------------------------------

## General Details

Advertiser * KING.COM	Buying Agency * OMD - NY	Primary Account Executive * Jawed A		
Calendar * Broadcast Calendar	Start Period * 2024-Q1	End Period * 2024-Q1	Deal Year * 2024/25	Marketplace * Broadcast Upfront
Date Range * 1/1/24 - 31/3/24	Contract Close Date	Time Zone		
Digital Ratecards General Market Scatter	Deal Notes Optimized Deal	Deal Link ID 0		
Account Manager	Media Strategist	Pitch ID		
Opportunity ID	Deal Type	Budget		
	Discount Profile			

It gave it a name  
I asked for digital inventory and AOS.

# AOS Landing Page

Home 🕒 ★ ?

	01-28	CONTACTS	AGENCY LLC	05				
	2024-03-31	1-800 CONTACTS	RAIN AGENCY LLC	2023-11-28	36893	PravinR	Approval	Open
Revenue	2025-05-01	YAHOO DSP	YAHOO DSP	2023-06-14	36783	Tubi - Trafficker Approval	Approval	Open
Entertainment/WWE/UF2324	2024-09-21	DISNEY - WALT DISNEY STUDIOS	OMG23	2023-09-25	36779	Sports Division - General Plan Approval	Approval	Open

approve task 36893

# Looking Ahead

- ML integration into encoding technologies is old news
  - Harmonic, Visionular, Bitmovin have integrated AI for years
  - Most short-term benefit from pre-filtering, content adaptive, and AI integrated into codecs compatible with existing players
  - Mid term affect on specialized codec markets (video for non-humans) could be very profound
  - New AI codecs for broadcast 6+ years off
- Generative AI, in particular, could significantly enhance utility and usability of many products in the next 2-3 years
  - AI for dummies