Fraunhofer

Digital Media

In the Business Area Digital Media five Fraunhofer Institutes are cooperating to provide technologies and solutions for the media industry.
Fraunhofer IIS, HHI, FOKUS, IAIS, IDMT

www.digitalmedia.fraunhofer.de

Follow us on LinkedIN

Trends and Technologies in Digital Media

Innovation for our Future

issue 16

Topics

Preface 3

Speed-up Your Production Workflow with JPEG XS. 4

InsightPersona – The Tailored Media Intelligence Solution to Optimize Your Communication Strategy 8

Generative AI and Agentic Systems in Broadcasting 10

MPEG-H DIALOG+ Delivering Next Generation Accessibility 12

Shaping the Future of High-Resolution AoIP Soundfield Capturing 14

Fraunhofer ContentServer R8 Technology 16

BOOTH 18

Fraunhofer Business Area DIGITAL MEDIA 22

Imprint 23

Preface

The media industry is facing ever-increasing challenges, manifested in rapid technological innovations and significantly changing user behavior. Personalized services, media library offerings, and content from social media and streaming platforms now dominate the market.

Although the hype surrounding podcast content continues unabated, demand for video content with appropriate subtitles is also increasing. Authenticity and multilingual support are more in demand than ever.

In particular, topics such as AI-supported tools and applications are driving our next-generation media solutions. Whether it's content verification, fake news detection, convenient, customized localization tools for broadcasting in digital cinema formats, or reliable deepfake detection – Fraunhofer experts are leading the way when the industry is looking for new, powerful solutions to improve its workflows and portfolios.

With our Dialog+ technology, we are presenting for the first time a solution that is capable of distinguishing between spoken and signed voices in order to facilitate access to audio content and, for example, to ensure easy access to interview formats in a festival environment and to ensure barrier-free access to this content.

With our award-winning JPEG XS coding technology, we and our licensees are paving the way for IP-based software-defined contribution workflows and are experiencing increasing market penetration. The advantage of such solutions lies in high transmission quality while simultaneously reducing costs and increasing flexibility.

On the production side – also new at IBC 2025 is a microphone array for easy and quick recording of 360° sound.

Are you curious about all our innovations? Then we cordially invite you to discover the full spectrum of technologies and market-ready solutions – it's worth experiencing!

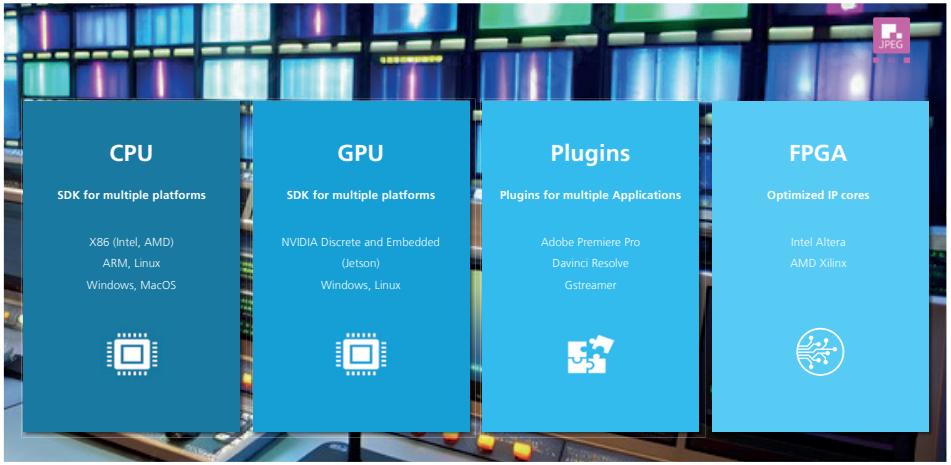
Enjoy reading!

Prof. Dr. Siegfried Foessel

Spokesman Fraunhofer Business Area Digital Media

Speed-up your Production Workflow with JPEG XS

The award-winnig, and ISO-standardized JPEG XS coding technology provides the ideal foundation for software-defined workflows and transmission of high-quality video. The Fraunhofer JPEG XS SDK enables manufacturers and broadcasters to easily set up IP-based workflows like ST-2110, IMPX and MPEG-TS in contribution environments or for live broadcast production.



JPEG XS Portfolio of Fraunhofer IIS

A JPEG XS – A game changer for media contribution workflows!

Especially designed for low-latency streaming of high-quality video over IP, JPEG XS provides a lot of advantages that count when re-thinking studio and contribution workflows to overcome bottlenecks and cost-intensive progress steps and to make your production environment resilient for future challenges.

- **High-Quality Transmission** for visually lossless video transfer over IP
- **Ultra-Low-Latency** making JPEG XS a premium solution for live broadcast transmission
- **Resource Efficiency** saving CPU and GPU load as well as memory usage and enables integration into all kinds of devices
- **Flexible Profiling and Scalability** balance quality, latency, and complexity, allowing adoption across a variety of use cases and support camera-to-cloud, camera-to-studio or studio-to studio. Suitable for HDR content and various color spaces through flexible coding options.
- **Robustness for Pro-Video and Broadcast Environments** by error resilience and streaming compatibility.

Fraunhofer JPEG XS SDK makes the difference

The JPEG XS SDK, implemented in software en- and decoders, enables a lot of cutting-edge devices, production and post-production workflows. The SDK allows for higher flexibility compared e.g. hardware-based monitoring solutions. Regarding cost, time efficiency and sustainable aspects JPEG XS really makes the difference.

The Fraunhofer IIS JPEG XS experts concentrated also on seamless system integration. Plugins for the NVIDIA Holoscan for Media framework allow direct data transfer from network interface cards (NICs) to graphic cards in parallel with JPEG XS encoding or decoding. JPEG XS is therefore a convincing live studio transmission codec.

Together with ST2110 audio-video-data synchronization using PTP (Precision Time Protocol) clocks JPEG XS ensures the accurate synchronization of all streams regardless of how the packets were routed.

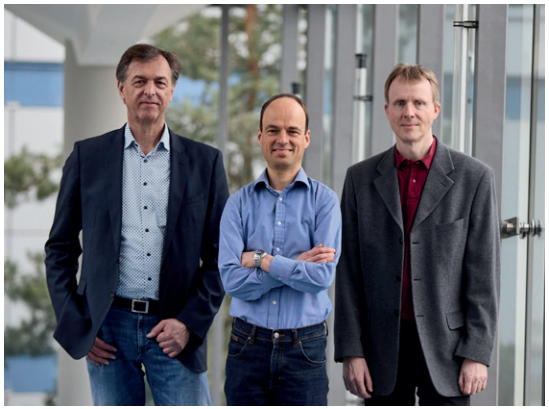
The Fraunhofer JPEG XS SDK is also compatible with IMPX (Internet Protocol Media Experience) to enable the transmission of compressed video, audio and data over IP networks especially for the prosumer market.

Fraunhofer JPEG XS SDK optimized now for MPEG-TS

JPEG XS can be used as codec for image data within an MPEG-TS stream. The images compressed by JPEG XS would then be transported in an elementary stream together with audio and additional meta-data streams. The standardized MPEG-TS format is designed to transfer in challen-ging streaming environments for example for IPTV, OTT or DVB providing synchro-nization, error tolerance and streaming compatibility. For live streaming or studio pipelines JPEG XS can extremely reduce the amount of uncompressed image data before the video stream is packed and transmitted. You profit from low band-width and low latency thanks to JPEG XS, while MPEG-TS ensures robustness and interoperability for live broadcasting.

Increasing Partner Network

Together with a permanently increasing partner network, Fraunhofer IIS can prove the performance of their JPEG XS software-defined coding as software-only solution or as IP-Core in hardware. From broadcasters to studio equipment manufacturers, from post-production software to production hardware – JPEG XS revolutionizes the way of high-quality contribution workflows.



Prof. Dr. Siegfried Foessel, Dr. Joachim Keinert, Dr. Thomas Richter

The renowned scientific Joseph-von-Fraunhofer Prize selected by a jury of scientists and industry representatives

And the winner is... JPEG XS

Existing video codecs involve accepting longer lag times or quality losses during transmission of image data in production environments. They also require a relatively large amount of computing power to decode the data, which means expensive hardware must be used and more energy is consumed. A team of researchers from the Fraunhofer IIS has developed JPEG XS, a state-of-the-art and forward-looking image compression format that transfers high-quality images with minimal latency, low resource use and almost lossless image quality. For this innovative project, the researchers were presented with the Joseph von Fraunhofer Prize for 2025.

JPEG XS

Contact

Maseeh Habibzay
Business Development and Marketing
Fraunhofer IIS
jpeg-xs-business@iis.fraunhofer.de

InsightPersona – The Tailored Media Intelligence Solution to Optimize Your Communication Strategy

Figure it out! Your customer asks for a customized strategy for their communication campaign or wants to optimize the next steps for moving picture communication or recorded presentations on stage. Imagine having a quick and reliable tool at hand that simplifies the process of crawling all video/audio archives and channels to get hold of important information about the protagonist of your campaign.

With InsightPersona the media analysis specialists from the Fraunhofer Institute for Digital Media Technology IDMT in Ilmenau and Oldenburg provide a technology that can support communication agencies, broadcasters and archives to optimize their clients' content strategy.

An important method for effective communication is a data-based strategy. "Therefore, data search in all available text, sound and video sources is the essential touchpoint", explains Christian Rollwage, Head of Audio Signal Enhancement at the Fraunhofer IDMT´s Oldenburg Branch. Based on audio and video sources that are freely accessible or available in your databases you get an in-depth look at content, context and frequency of specific quotes, topics or individuals, who speak or appear.

AI-based instant search

InsightPersona is based on dedicated AI-based audio and video analysis components optimized for huge media archives or live content. The speaker recognition engine was trained on more than 20,000 male and female speakers of every age and different pronunciations. Compared to common engines it uses a very small number of parameters to do its job – very fast and with low energy consumption.

Combined with the instant face search it enables a cross-modal analysis of visual and audio content to achieve better and more accurate search results for any reference person. The recognition and analysis of both audio and video data is crucial. "Especially when analyzing talk shows or group discussions it's important to get a comprehensive picture of how the protagonist of your communication strategy reacted, and how to improve his/her presence for the next stage appearance," explains Uwe Kühhirt, Head of Audiovisual Systems at Fraunhofer IDMT in Ilmenau.

InsightPersona supports your content strategy

InsightPersona provides detailed statistics to support the targeting of your communication and marketing strategy and informs you about whether your content complies with guidelines regarding gender, age classification or barrier-free access e.g. listening effort. It makes it easier for agencies to check if important key messages have been placed. In addition, InsightPersona can reflect people's vocal excitement by interpreting the corresponding audio stream.

The software of InsightPersona runs completely cloud-free if required. This ensures that sensitive data does not have to be transferred via external networks. Using a chatbot interface, you can ask your question in natural language.

Important facts to know

- Enables data-based strategy planning
- Combines face and speaker recognition for high-quality search results
- Instant access to video/audio sources for live reporting
- Easy handling via a chatbot interface
- Can be customized to your applications
- Analysis of vocal excitement levels and word-cloud functionality

As a result, you instantly get an overview of all visible appearances and speech/text passages with data source, broadcast and streaming date, information about the channel or TV station and time code. The findings are directly accessible through the links provided in the chatbot answer. The intuitive dashboard and user interface can be customized for your special needs, allowing you to work with InsightPersona as easily as possible.

Contact

Fraunhofer IDMT
Christian Rollwage
christian.rollwage@idmt.fraunhofer.de
Dr. Uwe Kühhirt
uwe.kuehhirt@idmt.fraunhofer.de

Generative AI and Agentic Systems in Broadcasting

Generative and agentic AI are transforming broadcasting – from adapting content for younger audiences to automating video editing and semantic archival search. Multi-agent systems could soon handle entire workflows in newsrooms and production. As adoption grows, so do concerns about bias and editorial control – raising the need for thoughtful integration of human–AI collaboration and human quality checks to keep audience trust.

Since the release of ChatGPT in late 2022, generative AI has been rapidly adopted in the broadcasting industry. Starting with text-based workflows in newsrooms, image and video models have since become available to analyze and produce visual content, and agentic systems now help automate entire workflows.

Current Adoption of Generative AI

Broadcasters such as the BBC and RTL are testing GenAI tools in their newsroom workflows to generate headlines and teasers or to summarize articles. The Norwegian public broadcaster NRK, for instance, automatically adapts summaries for younger audiences to boost engagement. Warner Bros. Discovery has created a "Cycling Central Intelligence Platform" to provide easy access to comprehensive information about cyclists and venues for sports commentators.

Agentic AI: from Assistants to Autonomous Agents

Multi-agent systems take this a step further by automating entire workflows while incorporating human-in-the-loop mechanisms for adjustments and quality checks. This may include optimizing subtitles for readability, shortening transcripts to meet time constraints, or rewriting articles for specific audiences. These approaches are also applicable to other modalities.

Multimodal AI Agents

Automated video editing can produce news segments from raw video footage and text, selecting and cutting suitable material. Multimodal semantic search can retrieve relevant archival content based on high-level semantic descriptions by applying reasoning models to find relevant material.

OdishaTV introduced a multilingual news anchor avatar 'Lisa' to produce regional news content. Similarly, automatic dubbing and lip-syncing can scale distribution of content across different languages. Fraunhofer IAIS is collaborating with broadcasters such as the German WDR and RTL to develop tailored GenAI solutions.

Opportunities and Risks

The rapid adoption of generative and agentic AI raises concerns about editorial accountability, content bias, and the erosion of human journalistic judgment. Fraunhofer IAIS developed human quality checks and structured human–AI workflows to help mitigate these risks by enabling correction and guiding AI output. These practices can also strengthen audience trust.

Contact

Dr. Christoph Andreas Schmidt
Fraunhofer IAIS
christoph.andreas.schmidt@iais.fraunhofer.de

MPEG-H DIALOG+
Delivering Next Generation Accessibility

Fraunhofer IIS is the first to deliver a solution to an industry-wide challenge in the field of dialogue enhancement: The latest MPEG-H Dialog+ feature recognizes singing voices and excludes them from dialogue enhancement.

MPEH-H Dialog+ is based on artificial intelligence and uses a deep neural network to automatically separate speech from background (music, effects, ambiences) of a final audio mix. The background is attenuated whenever speech is present, and it automatically remixes the content to a new, dialogue-enhanced version.

MPEH-H Dialog+ offers an alternative to the original audio mix that viewers can select if they want to enhance the dialogue, giving people with hearing disabilities, for example, easier to understand options. It is the perfect solution for content providers to enhance audio material when only the final audio mix is available.

Until now, conventional codecs have struggled to detect singing in a film's music and to avoid enhancing it. Fraunhofer IIS has solved this issue with the new feature of MPEG-H Dialog+.

The technology prevents automatic dialogue enhancement from being applied to singing that occurs in the background. This allows musical passages to retain their full sound quality while still enabling personalization of dialogue sequences.

Recently, French-German public broadcaster ARTE joined forces with Fraunhofer IIS to use the dialogue enhancement technology Dialog+ for its streaming service arte.tv. Recipients will be able to switch between the original production and a version with "Klare Sprache" in German or "Confort Audio" in French content.

By adopting the new feature, ARTE is now taking another important step toward accessible streaming.

»Fraunhofer IIS debuts new MPEG-H Dialog+ feature: The world's first dialogue enhancement to intelligently distinguish speech from singing.«

Contact

Daniela Rieger
Fraunhofer IIS
amm-info@iis.fraunhofer.de

Shaping the Future of High-Resolution AoIP Soundfield Capturing

The evolution of microphone array technology towards high-channel-count, real-time capable, IP-based systems opens new horizons for spatial audio acquisition and analysis. The MicA platform demonstrates how AES67-based networking, nanosecond-precise synchronization, and flexible array geometries can enable applications from higher-order Ambisonics to drone detection using beamforming techniques. This article outlines the future potential of such systems in broadcasting, conferencing, and security domains.

Microphone arrays are essential tools for spatial audio acquisition, beamforming, and source localization in broadcasting, conferencing, and surveillance. Conventional systems often embed signal processing within dedicated hardware, which can restrict scalability, computational flexibility, and access to raw channel data. The MicA system represents a shift towards IP-native, network-distributed architectures. By enabling real-time streaming of all microphone signals, it supports flexible post-processing, integration of advanced algorithms, and coordinated operation of multiple arrays.

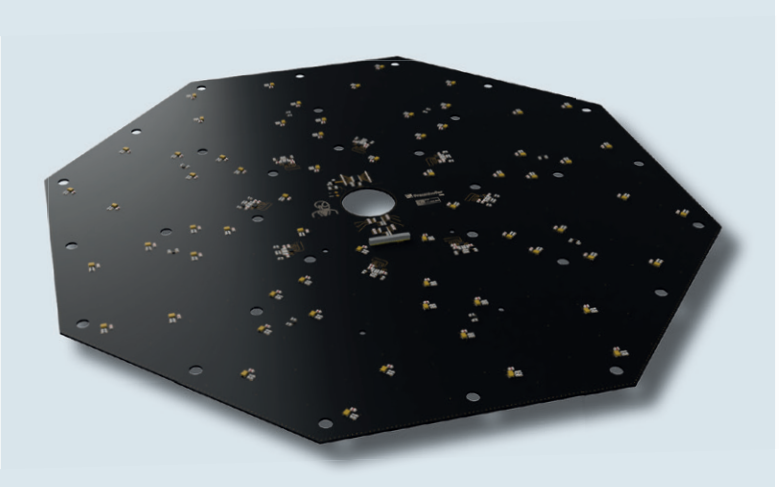
Technological Foundations

MicA employs AES67-compliant Audio-over-IP transmission for up to 256 MEMS microphones per module. Precision Time Protocol (PTP) ensures nanosecond-level synchronization, enabling coherent multi-array operation. The printed circuit board design supports custom 1D, 2D, and 3D array configurations, tailored to application-specific spatial and frequency resolution requirements. Multiple modules can be deployed in a synchronized network to increase aperture size and angular coverage. Using Power-over-Ethernet (PoE) minimizes infrastructure requirements, facilitating deployment in distributed sensing scenarios.

Application Domains

High-resolution Sound Field Capturing

Large aperture, high-density arrays support advanced beamforming, acoustic holography, and higher-order Ambisonics with minimal spatial aliasing.
Multi-array synchronization enables the reconstruction of complex sound fields for immersive broadcasting.



Microphone array

Conferencing

Enhanced direction-of-arrival estimation and adaptive beamforming improve speech intelligibility and spatial separation, enabling natural telepresence experiences even in reverberant environments.

Sound Source Acquisition for Auralizations

Array Measurements can be used to isolate partial sound sources within complex acoustic environments, even in scenarios with multiple simultaneously active emitters. The ability to track and separate individual sources, including moving objects, allows the derivation of highly realistic digital models, which can form the basis for perceptually accurate auralizations, preserving directivity patterns and spatial relationships.

Surveillance and Object Detection

Large scale acoustic sensing allows for detection, classification, and trajectory prediction of security sensitive objects such as unmanned aerial vehicles based on their acoustic signatures.

Distributed networks enable wide-area coverage and 3D-tracking, offering valuable capabilities for monitoring and early-warning systems.

Outlook

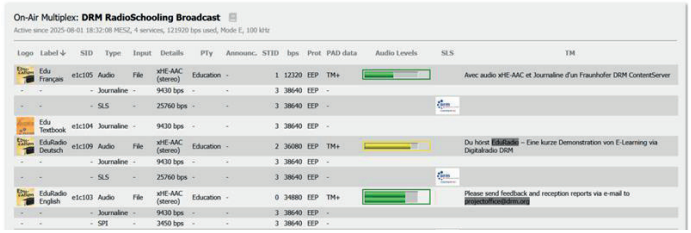
The transition from closed, hardware-centric systems to scalable, synchronized, IP-native microphone array platforms enables new approaches in industrial noise monitoring, environmental sensing, and AI-driven spatial analysis. Open-standard compliance ensures future-proof integration and the possibility of building adaptive, large-scale acoustic sensor networks for a wide variety of applications.

Contact

Thorben Kron
Fraunhofer HHI
thorben.kron@hhi.fraunhofer.de

Fraunhofer ContentServer R8 Technology

At IBC, Fraunhofer IIS launches the latest generation of its trusted Content-Server™ technology. The new R8 version delivers powerful enhancements for broadcasters worldwide. It supports both, DAB and DRM standards with unmatched flexibility, reliability, and ease of integration.
Discover the future of digital radio at IBC 2025.



Broadcasting made smarter: ContentServer R8 for DAB and DRM

The Fraunhofer ContentServer R8 is a professional head-end technology that combines internal audio encoding, data service management, and multiplex generation into one robust platform. Designed for broadcasters of all sizes, from local stations to international networks, it offers seamless integration into existing workflows and full remote control via a user-friendly web interface.

Built for Innovation and Reliability

The ContentServer R8 supports advanced receiver testing with a built-in DRM Modulator and enables innovative network structures like xSTI-C for DAB regionalization.

Its powerful redundancy features – including cross-site failover – ensure uninterrupted service, even in critical environments.

New in R8: More Power, More Possibilities

The latest R8 upgrade introduces:

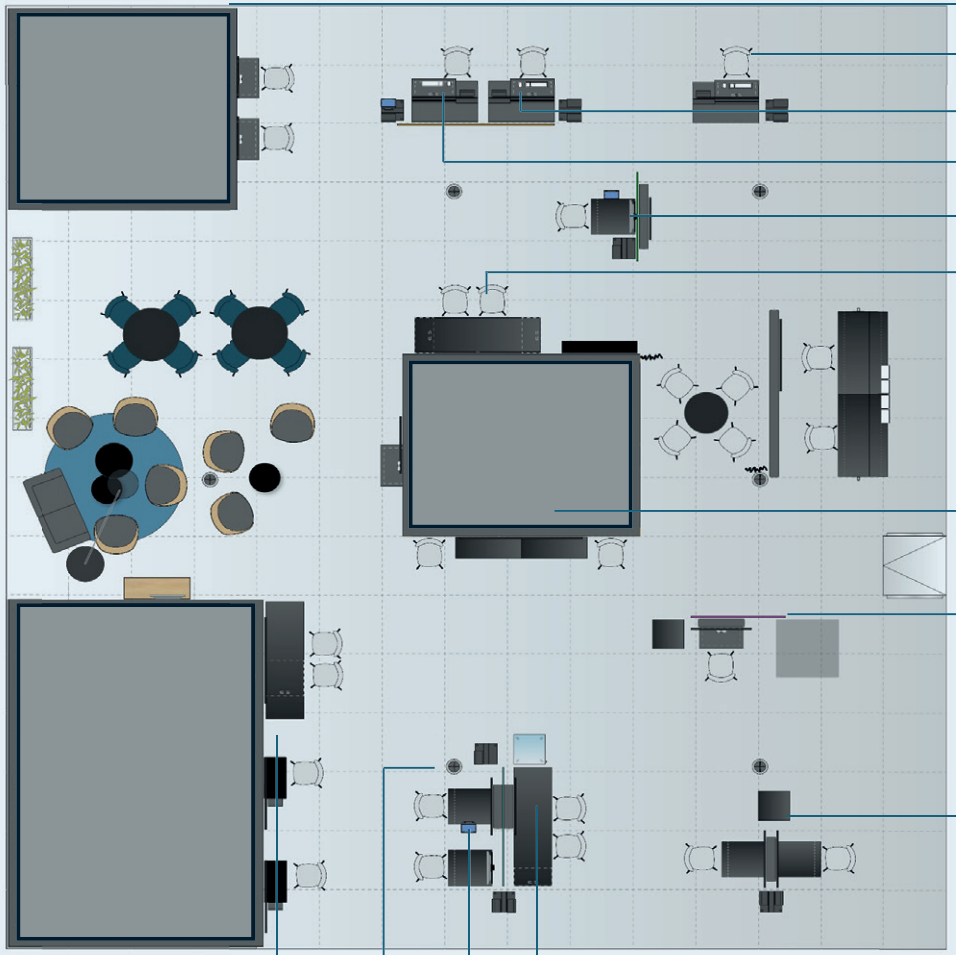
- On-Air Multiplex Overview with live content previews and source status
- Automatic loudness normalization and limiting
- Secure DCP streaming via SRT
- A new Journaline encoder with enhanced XML and image handling
- Support for all regional text profiles, including Arabic and bidirectional scripts
- Improved virtual machine compatibility and secure SFTP interfaces

These features ensure that broadcasters can deliver high-quality, multilingual, and interactive digital radio content with minimal effort.

Contact

Alexander Zink
Fraunhofer IIS
cs-support@iis.fraunhofer.de

Booth



Multi-language Subtitling / easyDCP

AI4Media

Content Verification Toolbox

InsightPersona: Media Intelligence for Personas

Digital Radio Solutions for DAB+ and DRM

Allinga TTS

Trustworthy Authentication for Compressed Video Data

Shaping the Future of High-Resolution AoIP Soundfield Capturing

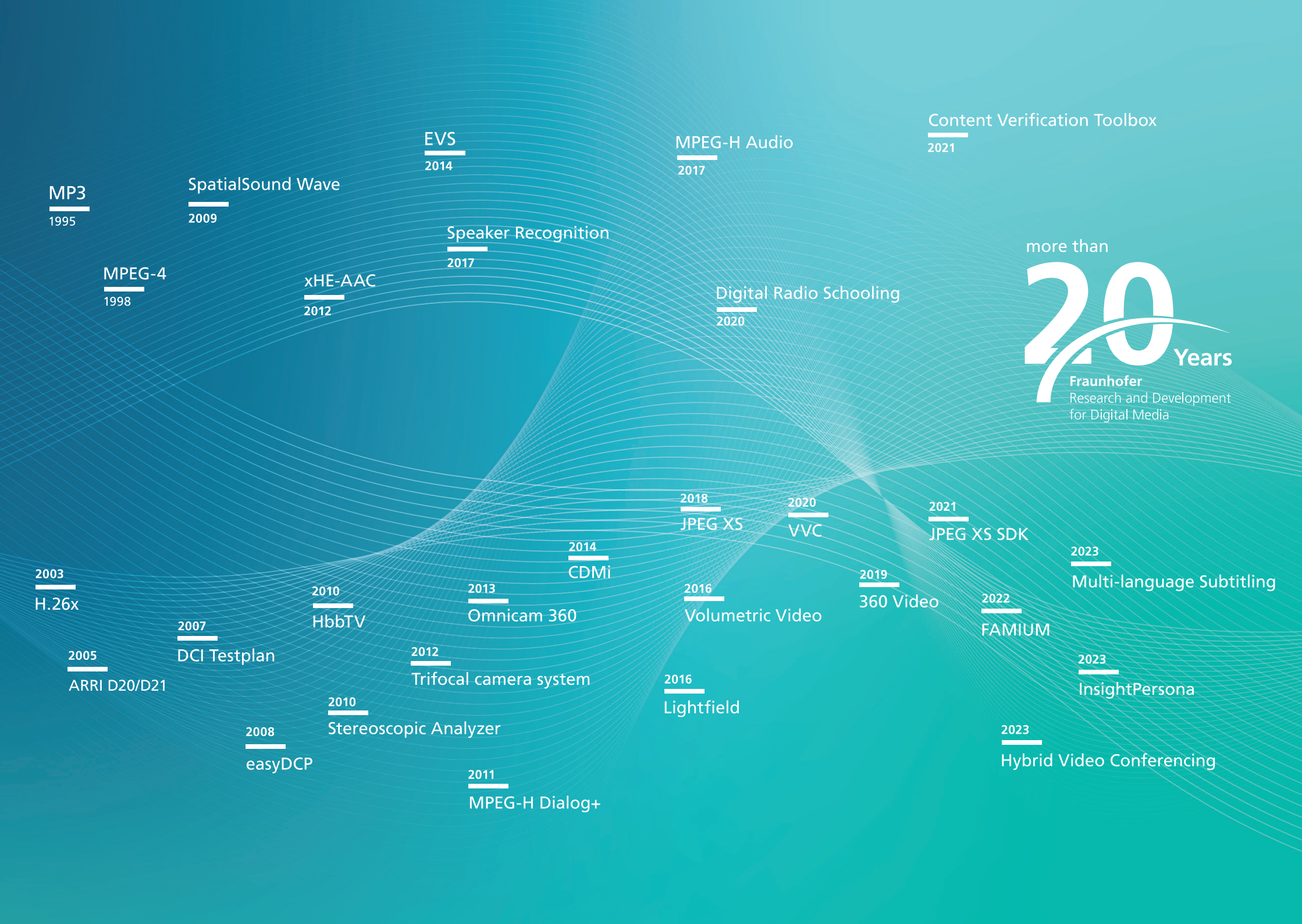
FAMIUM
LLM Solution; Streaming Media Test Suite, Green Streaming, Provenance and Authenticity

MPEG-H Audio – Next-generation Broadcast and Streaming

Speed-up your Production Workflow with JPEG XS

MPEG-H Dialog+: Delivering Next Generation Accessibility

Efficient Audio Streaming with xHE-AAC


MP3
1995
SpatialSound Wave
2009
EVS
2014
MPEG-H Audio
2017
Content Verification Toolbox
2021
MPEG-4
1998
xHE-AAC
2012
Speaker Recognition
2017
Digital Radio Schooling
2020
more than
20 Years
Fraunhofer
Research and Development
for Digital Media
2003
H.26x
2005
ARRI D20/D21
2007
DCI Testplan
2008
easyDCP
2010
HbbTV
2010
Stereoscopic Analyzer
2011
MPEG-H Dialog+
2012
Trifocal camera system
2013
Omnicam 360
2014
CDMi
2016
Lightfield
2016
Volumetric Video
2018
JPEG XS
2019
360 Video
2020
VVC
2021
JPEG XS SDK
2022
FAMIUM
2023
Multi-language Subtitling
2023
InsightPersona
2023
Hybrid Video Conferencing

Fraunhofer Business Area DIGITAL MEDIA

The cooperation of Fraunhofer Institutes within the business area Digital Media provides innovative solutions and products for the digital age of motion picture.
We provide technological innovations for digital media workflows and for immersive viewing and sound experiences. Benefit from our expertise in research and development as well as in standardization. The institutes offer research and development in the areas of production, audio systems, data compression, post processing, transmission, projection, distribution and digital archiving.

As an one-stop competence center for digital media we provide for our customers scientific know-how and the development of solutions that can be integrated in workflows and optimize process steps.

The members of the Digital Media Business Area are actively working in renowned organizations and bodies like International Standardization Organization ISO, SMPTE (Society for Motion Picture and Television Engineers, FKTG (German Society for Broadcast and Motion) and VESA (Video Electronics Standards Association).
We are also a partner of the 3IT, the Innovation Center for Immersive Imaging Technologies and the Fraunhofer Digital Media Technologies, Fraunhofer USA, Inc. These contributions enable research and development activities based on international standards.

The members of the Fraunhofer Business Area DIGITAL MEDIA are

- Fraunhofer Institute for Integrated Circuits IIS, Erlangen
- Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut HHI, Berlin
- Fraunhofer Institute for Open Communication Systems FOKUS, Berlin
- Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS, St. Augustin
- Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau

Imprint

Publication Information

Fraunhofer
Business Area Digital Media
c/o Fraunhofer Institute for
Integrated Circuits IIS
Am Wolfsmantel 33
91058 Erlangen, Germany

Concept and Publisher
Angela Raguse
Fraunhofer Business Area Digital Media

Editors
Angela Raguse
Jennifer Karbach
Hannah Pohl
Joran Poortstra
Thorben Kron

Layout and production
Ariane Ritter

Photo acknowledgements
Cover picture: © Fraunhofer-Gesellschaft

Page 3: © HFF / Fraunhofer IIS
Page 5: © Adobe Stock / Fraunhofer IIS
Page 6-7: © Fraunhofer IIS
Page 8: © Fraunhofer IDMT mit Canva erstellt
Page 11: © Adobe Stock / Fraunhofer IAIS
Page 13: © Fraunhofer IIS
Page 15: © Fraunhofer HHI
Page 16: © Fraunhofer IIS
Page 17: © Miljan Živković – stock.adobe.com
Page 19: © Fraunhofer IIS / etGlobal

Contact

Fraunhofer
Business Area DIGITAL MEDIA
Angela Raguse M.A.
Phone +49 9131 776-5105
contact-digitalmedia@iis.fraunhofer.de
www.digitalmedia.fraunhofer.de