Envisioning the Future of Communication

Τόμ. 2, Αρ. 1 (2025)

Envisioning the Future of Communication - Conference Proceedings vol. 2

Intelligent processing and data management of sports content in media and communications

Nikolaos Vryzas, Lazaros Vrysis, Stamatis Poulakidakos

doi: 10.12681/efoc.7896

Intelligent processing and data management of sports content in media and communications

Nikolaos Vryzas, Lazaros Vrysis, Charalampos Dimoulas*

Abstract

This work presents a framework for intelligent processing and management automation of sports content utilizing algorithmic techniques and Artificial Intelligence (AI) methods, such as Machine/ Deep Learning (ML/DL). In the modern digital media landscape, sports data is among the most popular news/informing categories, favoring mediated communication and audience interaction. The main subject of interest may concern any sports event (e.g., a sports match) and all accompanying data related to player statements, athlete profiles, historical records of similar sports competitions, broader events, audience reactions, and more. The proposed approach introduces a series of techniques for semantic processing, annotating, and data linking mechanisms, providing a broader framework for indexing, and retrieving interconnected information. These data may include audiovisual material from event recordings, textual streams of unstructured or standardized descriptions, contributed content, comments, and reactions from ordinary users, before, during and after the main event. A modular ontology for organizing and describing events allows for the structured management of these informatory streams, making them useful for coaches, sports analysts, editors, journalists, and the broader audience. The basic functional capabilities and pilot results of the initial techniques applied in basketball are presented.

Keywords: sports semantics, intelligent processing, content management, machine learning, artificial intelligence, quality of experience.

Introduction

In today's world, where media is critical in shaping information dissemination and audience engagement, this work presents a cutting-edge framework that leverages algorithmic techniques and Artificial Intelligence (AI) methods to process and manage sports content. From the excitement of live events to the rich history of archives, sports content is a vital component of modern communication, connecting communities and captivating audiences worldwide (Bellamy, 2009; Dimoulas, 2020; Fuss et al., 2013; Owens, 2015; Raney, & Bryant, 2009; Romney, & Johnson, 2020). Multiple content entities (i.e., text, images/graphs, photographs, audio, video) are employed for this purpose, aiming at engaging audiences and enhancing the offered Quality of Experience (QoE), ensembled in multimodal presentations, non-linear

* Nikolaos Vryzas, Post-Doc researcher, School of Journalism and Mass Communications, Aristotle University of Thessaloniki, Greece, nvryzas@auth.gr
Lazaros Vrysis, Post-Doc researcher, School of Journalism and Mass Communications, Aristotle University of Thessaloniki, Greece, lvrysis@auth.gr
Charalampos Dimoulas, Professor, School of Journalism and Mass Communications, Aristotle University of Thessaloniki, Greece, babis@eng.auth.gr

storytelling, multichannel/spatial audio, and 3D video arrangements, or even eXtended Reality (XR) media applications (Dimoulas, 2016; Kalliris et al., 2014; Kotsakis et al., 2014a; Matsiola et al., 2015; Vegiris et al., 2008). The enhanced experience purposes to entertain the public and also to offer cultural and pedagogical interventions for society, given that sports have also been part of cultural spectacles and demonstrations. In this context, what content we capture, document, and broadcast today will be the sports audiovisual heritage for future generations, thus special care needs to be taken during pre-production, production, and post-production phases (Dimoulas et al., 2014; Dimoulas, 2022; Podara et al., 2021). In today's digital era, social media and User Generated Content (UGC) play a very significant role that cannot be left out of the discussion, with the television experience being wholly altered with the use of multiple cross-/trans-media channels, second-screen television applications, mobile footage contributions, and more, which are dominant in the consumption of television-broadcasted sports content (Dimoulas & Symeonidis, 2015; Dimoulas & Veglis, 2023; Katsaounidou et al., 2018; Katsaounidou et al., 2019; Palioura & Dimoulas, 2022; Podara et al., 2019; Saridou & Veglis, 2021; Sidiropoulos et al., 2019).

Elaborating on the above remarks, the advent of social media thoroughly enhanced audience participation through content playing, sharing, commenting, and interacting, therefore augmented the overall user experience (Boulianne, 2020; Jenkins & Ito, 2015; Matsiola et al., 2015; Vázquez-Herrero, & López-García, 2019). Among others, the "new media" era (Web 2.0 and beyond) multiplied the times and ways that content is produced, distributed, getting reactions, searched, and retrieved, thus creating new needs for more sophisticated media assets description, documentation, and management automations (Dimoulas et al., 2018; Kalliris & Dimoulas; 2009; Katsaounidou et al., 2019; Matsiola et al., 2015). Hence, apart from textual metadata, content-extracted descriptors, data-handling ontologies, and intelligent systems are now employed for providing deeper insights into the presented information, i.e., thematic classification, emotions, semantic conceptualization, relevant events detection/identification with aligning and linking mechanisms and more (Dimoulas & Veglis, 2023; Filippidis et al., 2018a; 2018b; 2019; Kapela et al., 2015; Kotsakis et al., 2014b; Tsolakis, 2023; Tsolakis et al., 2023; Vryzas et al., 2018; 2021a). Data-driven approaches relying on smart systems trained through Machine/Depp Learning (ML/DL) and hybrid decision making algorithms are now part of both the production and post-production processes, intended to facilitate the capturing/framing and streaming processes (Vegiris et al., 2008; Vryzas et al., 2021b), serve content indexing during editing (Dimoulas et al., 2007a; Vegiris et al., 2009), face emotional/semantic classification and personalization needs (Dimoulas & Symeonidis, 2015; Kotsakis et al., 2014b; Vryzas et al., 2018; 2021a; 2021b). Apart from documentation and management purposes, AI-based approaches are currently employed to enhance content quality by suppressing noise and removing other degradation artifacts, increasing image and audio resolution and, overall, augmenting the offered experience (Braun et al., 2021; Dimoulas, 2020; Fuss et al., 2013; Jo et al., 2018; Rota et al., 2023; Zhang et al., 2017). Such services are considered critical when it comes to the utilization of archived material in heritage applications

and reuse scenarios, with imperative usefulness in the sports media domain (Chatzara et al., 2019; Dimoulas et al. 2014; 2018; Dimoulas, 2022; Dimoulas & Veglis, 2023; Romney, & Johnson, 2020).

The current work elaborates on the need for intelligent documentation, indexing, and warehouse of sports data, including the incorporation and proper enhancement of archived material. The research hypothesis is that data-driven AI approaches utilizing sophisticated ML and DL solutions are technologically mature and efficient in delivering the stated requests of quality enhancement and media assets management automations by detecting and utilizing multiple/multimodal streams describing the same sports event. The paper envisions a modular architecture of different sub-systems that can be adaptively combined to face the needs of each explicit case, indicating a framework of data-driven solutions with the associated dataset repositories that can be configured to deliver the corresponding needs. This framework offers a structured approach to handling the diverse array of sports-related data and unlocks new opportunities for semantic analysis, annotation, and data linkage. By harnessing the power of ML/DL algorithms, it provides a sophisticated platform for indexing, retrieving, and interpreting interconnected content, ranging from player insights to fan reactions. Furthermore, this approach transforms raw data streams into actionable insights through a meticulously crafted ontology, benefiting a broad range of stakeholders. Coaches, sports analysts, journalists, and enthusiasts can utilize its comprehensive organizational structure, enabling informed decision-making and enhancing the overall sporting experience.

The structure of the paper is organized as follows. The next section states/defines the problem in discussion, its multidisciplinary scientific background, and the associated literature review. Subsequently, the introduced *mineCourt* framework (*Mining Enhanced Content semantics thru sports data)* is presented and discussed. Finally, pilot results and discussion summarize the proposed modular solution with its innovating and contributing research aspects, demonstrating use cases for future implementation and further progress anticipation.

Problem Definition and Background

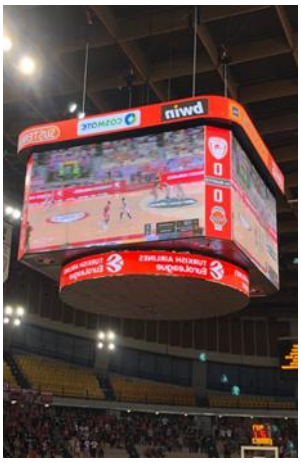
Based on the preceding analysis, when it comes to sports broadcasting and archiving, there are two significant areas of focus. The first is the Quality of Experience (QoE) offered to viewers, which has seen significant improvements in recent years. These include enhancements in visual resolution, more vivid colors, increased frame rates (and motion dynamics), multichannel sound-tracks with spatial/3D audio reproduction capabilities, content profiling and device adaptation functionalities and more (Dimoulas, 2020; Fuss et al., 2013; Owens, 2015; Raney, & Bryant, 2009). Additionally, the broadcasting process has been elevated, with various augmentation layers now included in the streamed data and metadata, thus turning the broadcasting process into a whole new experience. No doubt, there is no comparison between the grade of archived content and new sports footage, where quality degradations are evident in the old material (image noise, poor resolution, audio noise contamination, sound artifacts and

more), mainly due to the initial capture and preservation in analog form (Figures 1, 2). Nevertheless, sophisticated Machine/Deep Learning (ML/DL) approaches can be employed nowadays for archived material restoration through deep noise suppression and super-resolution enhancement, utilizing Generative AI models to synthetically recreate missing/poor information in audio and visual/image content.

The second pillar emphasizes the importance of efficient content archiving and management automations through content annotation and indexing mechanisms (Figure 3), which is especially critical when dealing with vast amounts of prolonged audio and video recordings. It is no coincidence that such semantic web services have been systematically pursued since the advent of the Web 2.0 era, as already commented, making media assets management a practical necessity in most related everyday applications (Boulianne, 2020; Dimoulas & Veglis, 2023; Jenkins, 2015; Kalliris & Dimoulas, 2009; Katsaounidou et al., 2019; Matsiola et al., 2015; Saridou & Veglis, 2021). With the abundance of content and considering the progress and maturity of media compression algorithms and encoding protocols, the research attention was turned to innovative, intelligent data management solutions. Therefore, content indexing, documentation, and retrieval services are crucial for making information usable in post-/ batch-processing modes, such as reviewing footage for reexperiencing an event, making it usable in content reuse scenarios or implicating further correlation and analysis processes (Dimoulas, 2020; Nikolopoulou & Papagianni, 2015; Papachristou, 2016). Appropriate documentation and archiving are equally essential as a minimum heritage mediation for future generations, not just in Journalism and Sports Businesses (Dimoulas et al., 2014; 2018). In this context, metadata augmentation and sophisticated content description are employed to archive data, along with quality enhancement tasks, to facilitate broader and easier usage.

a.

b.



c.

Figure 1: Snapshots of basketball events: (a). archived material of an old event with quality degradations and poor visual aids; (b). contemporary views with visual analysis aids that alter the offered experience (both in the physical venue and in broadcasted media); (c). use of action cameras for a different viewing experience (Dimoulas, 2020; https://youtu.be/iXFJMNiCzJo?si=yLRcDlWcVSnqxh0a; https://www.youtube.com/live/V_AQYFLLMqA?si=sEWKw0bcmCczrf0Z)

a.

b.

c.

d.

Figure 2: Snapshots of football broadcasting footage: (a). archived material of an old event with obvious quality degradations; (b, c, d). contemporary views with visual aids that alter the offered experience (Dimoulas, 2020; https://youtu.be/rrcsLwUJdQk?si=ZJh1G_gPZ2jF8lsP).

Figure 3: Indexing and annotation visualization of football and basketball matches: (a). indicating event points within a football match; (b). providing visual aids and statistics concerning spatiotemporal football possession/movements; (c) providing visual aids and statistics concerning spatiotemporal basketball possession/movements (Dimoulas, 2020; Nikolopoulou & Papagianni, 2015; Papachristou, 2016; Filippidis et al. 2018a; 2018b; 2019).

Elaborating on the above remarks, the problem definition can be projected in a block diagram by formulating a mathematical function to be resolved (Figure 4). The implicated solution (indicated by the red dotted frame) is fed by the massive volumes of unstructured data covering multiple sports events (system input), with the desired output being the enhanced and semantically augmented media assets (bluish dotted frame). Indexing and data linking cues are

involved for specific records to be matched (and internally aligned to each other) with specific sports events timelines. More specifically, following the data-flow pipeline of Figure 4, unimodal content entities are extracted from the available media records and are further subjected to detection, matching, and alignment of the targeted athletic events/matches, with content restoration and enchantment needs to be detected and deployed. Then, the implicated modalities (and their advantageous processing/ analysis insights) are combined in a multimodal fusion, thus leading to semantic analysis indexing and metadata augmentation.

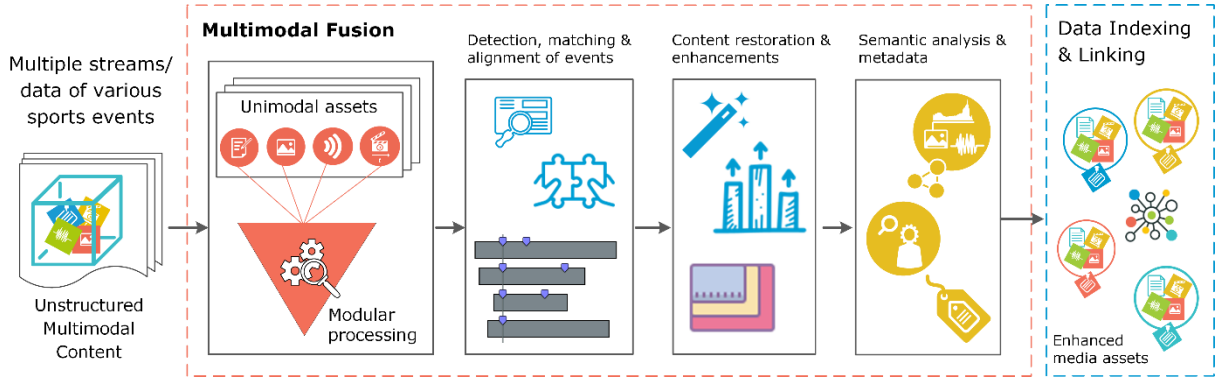


Figure 4: Block diagram presenting the given problem and the involved processes.

Content preservation is one of the oldest and most critical tasks regarding archives of audiovisual material heritage. Given that technology has been tremendously advanced, old footage and formats used to support lower quality concerning audio and visual resolution, bit depth, noise presence, etc. (Dimoulas, 2022; Dimoulas et al., 2014; Jack, 2011; Kalliris & Dimoulas, 2009; Mourjopoulos, 2005). Hence, traditional audio and image/video enhancement techniques have been employed for many decades to get rid of unwanted contamination components (broadband acoustic background noise, reverberation, salt-and-paper visual artifacts, etc.) utilizing variations of spectral subtraction and spatial filtering, which could also be combined with signal detection and segmentation techniques (Dimoulas et al., 2007a; 2007b; Kalliris et al., 2001; 2016; Tsardoulias et al., 2016; Tsilfidis & Mourjopoulos, 2011; Vegiris et al., 2009). Super-resolution imaging (and audio) techniques were also developed to increase the content sharpness by augmenting the available pixels/sample with new predictions, which, in the case of video enhancements, were also combined with accurate sub-pixel motion estimation (Dimoulas., 2020; Dimoulas et al. 2007b; 2014; Jack, 2011; Konstantoudakis et al., 2018a; 2018b). Nowadays, most of the above restoration and enhancement techniques can be deployed with machine learning algorithms, which can settle both restoration and spatiotemporal resolution enhancement needs by detecting specific motifs/patterns and synthetically generating missing patches (Braun et al., 2021; Jo et al., 2018; Lucas et al., 2019; Rota et al., 2023; Ulyanov et al., 2018; Wang et al., 2018; Zhang et al., 2023).

Concerning media indexing and documentation, data-driven approaches are utilized for content-based audiovisual description and management automations, which have also been employed in recognition of sports actions (Dimoulas, 2020; Filippidis et al., 2019; Kapela et

al., 2015; Pan & Li; 2020; Liu et al., 2021; Wu et al., 2020; Liu et al., 2021). During the last decades, AI-assisted solutions have been dominant, with a broad multidisciplinary presence in multiple fields, from art and (cultural) heritage to media applications, audiovisual production automations, human-machine interaction and more (Chatzara et al., 2019; Dimoulas, 2022; Dimoulas & Veglis, 2023; Dimoulas et al., 2014; Vryzas et al., 2018; 2021a; 2021b). Hence, audio novelty and motion activity tracking techniques have been initially implemented for event detection, segmentation, and summarization purposes, utilizing relevant content-extracted features (Dimoulas et al., 2007a; 2007b; Filippidis et al., 2019; Dimoulas & Kalliris, 2013; Kalliris et al., 2016; Müller, 2007). Sentiment analysis and quality of experience estimation/profiling methods have also been employed in that direction (Kotsakis et al., 2014a; 2014b; Vryzas et al., 2018; 2021a). Multimodal approaches are usually encountered, with audio entailing a significantly lighter computational burden compared to handling visual/motion assets and also proving advantageous in event detection/ matching tasks due to the higher likelihood of shared audio information across different content versions, irrespective of recording device placement or direction—a feature not inherently present in video (Dimoulas & Kalliris, 2013; Dimoulas & Symeonidis 2015; Kalliris et al., 2016). At the same time, the text modality can be proved significantly useful in sports events detection and social media reactions monitoring, baring the most lightweight processing nature (Dimoulas & Veglis, 2023; Fillipidis et al., 2018a; 2018b; Tsolakis, 2023; Tsolakis et al., 2023).

Shifting from machine learning to the era of deep learning architectures, with the availability of large datasets and more sophisticated solutions (i.e., convolutional neural networks, generative adversarial networks, transfer learning techniques and more), the potentials of such data-driven solutions seem unparallel (Liu, 2021; Liu et al., 2021; Lucas et al., 2019; Pan & Li, 2020; Vryzas et al., 2021a; Wang et al., 2018). Still, careful attention is needed while deploying such systems, especially regarding archiving, documentation, and heritage applications. For instance, by applying visual enhancement through generative AI models, there is no safeguard to check the authenticity and preciseness of the synthetically produced imaging patches. While the outcome may seem realistic, improving the offered experience in perceptual terms, a dataset bias may cause unwanted style and pattern errors, altering the character and the whole impression provided by the enhanced content, which might misalign with the heritage aims. Likewise, AI enhancement of heavily noisy speech has been shown to create voicing and pronunciation artifacts, which new generations might not even notice. Finally, with the vast amount of data, available datasets, and the plethora of AI solutions, careful treatment is needed within a standardized framework, like the one proposed by the current research. Otherwise, no matter how smart the deployed solutions are, increased complexity, computational overheads, and unpredictable algorithmic behaviour might cause more problems than the ones we are trying to solve.

The Proposed Framework: Mining Enhanced Content semantics thru sports data (mineCourt)

The proposed mineCourt approach (Mining Enhanced Content semantics thru sports data) suggests a modular integration framework, targeting enhanced semantics extraction of sports content towards sophisticated indexing and analysis of specific events. With the main aim being to automate the media assets documentation and management utilities, quality enhancements are considered part of the process (even if they have not been extensively applied so far) and can be delivered to facilitate the event detection and segmentation tasks (Dimoulas et al., 2007a; 2007b) or as a sole/explicit procedure, in case old/archived footage or UGC material containing noising artifacts and quality degradations are involved. The first related attempt focused on the implementation of an annotation approach for indexing, summarizing, and visualizing the implicated events in football matches. Specifically, the idea was that radio producers, lively commenting/describing a game, could spatially annotate the actions on a top-view court field view through a web service listing all the encountered football activities (pass, dribble, shot, goal, foul/penalty, etc., Figure 3a) (Nikolopoulou & Papagianni, 2015). The main challenge was to test whether a journalists' pair could commit to the parallel task of commenting and annotating the game without interruptions or flaws in their radio broadcasting descriptions (which proved feasible). Hence, it came up that content could be appropriately annotated with the event in progress, producing rough segmentation and indexing ques while also offering the generated visualization to radio audiences (the ones with no access to video broadcasting) (Dimoulas, 2020; Nikolopoulou & Papagianni, 2015).

The crafted football annotator (Figure 3a) allowed further elaboration, encompassing more specific details on the succession of the activities (e.g., goals will always follow kick actions, i.e., shot, foul, penalty, etc., penalties can be annotated only within the penalty area and so on). This time, the effort was focused on basketball matches, with a tighter state transition diagram providing a roadmap of all possible endings/successions during each event. For instance, a possession change can occur within the following conditions: after a turnover, an offensive foul, a defensive rebound, or a successful basket. Hence, the developed basketball annotator (Figure 4c, left) exhibited enhanced standardization on the available annotation choices, thus minimizing the sets of choices on each decision node and making it easier for the indexing process to be lively conducted (provided that necessary automations are incorporated in the basketball annotator) (Papachristou, 2016; Filippidis et al., 2018a; 2018b). Hence, the fact that basketball matches contain much more internal events that succeed in short times and smaller court areas, using less space (e.g., compared to football), was one of the reasons that pointed our attention to this direction (apart from the personal interest, preferences, and background of the involved researchers). Another challenge was the difference between the broadcasting time and the match clock (with the first being continuous and the second interrupted during the inactive play moments and pauses), making it more difficult to establish a ground time reference for all associated footage, e.g., main coverage steams, highlights, UGC contributions, etc.

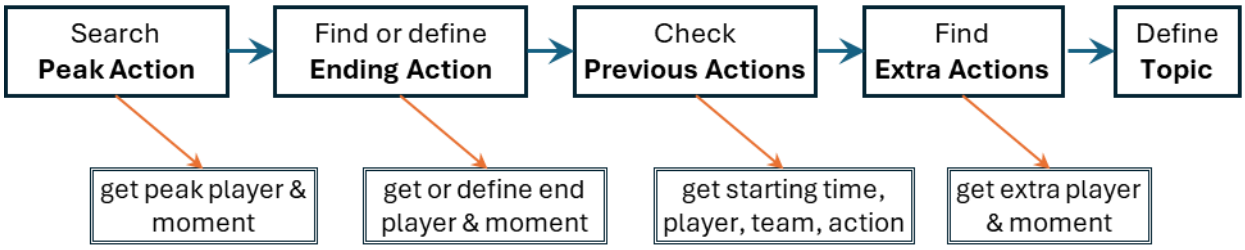
(again, compared to football or other sports). A final reason for working on basketball matches/events is that related contemporary research is conducted with similar aims (Liu, 2021; Liu et al., 2021; Pan & Li, 2020; Wu et al., 2020), facilitating, inspiring, or even motivating the current effort. Nonetheless, the selected area is considered a case study in our integration framework, meaning that individual modalities could easily be updated to meet the needs of each different sports match (with less complicated plays).

Returning to the helpful insights and controls provided by the state transition diagram and the annotator updates, further modalities were pursued to facilitate the process by elaborating on the User Interface (UI) of the annotation web service and/or incorporating additional content entities and tools (i.e., broadcasted audio with the match descriptions, textual streaming and standardized play-by-play (PBP) descriptions, voice commanding to state the indexing labels, etc.). The most convenient and impactful data turned out to be the textual information, especially the PBP text provided by each league (e.g., NBA, Euroleague, Greek BasketLeague, etc.). Utilizing Natural Language Processing (NLP) algorithms, these messages can be combined, interpreted, and projected into meaningful events components, dataflows, and summaries (Figure 5), as successfully tested with the minor differentiations in terms and rules used of different leagues (Filippidis et al., 2018a; 2018b). Similar processing can be established with live textual streaming provided by sports websites during a match, utilizing the offered enhanced descriptions to detect buzzwords of possession actions and define the associated events (in combination with the PBPs or in cases they are not available). Apart from its computationally lightweight nature, the important aspect of this textual process lies in its ability to fully align with the states transition diagram and the annotator utility. More interestingly, further audio and video assets can be controlled through a multimodal ontology framework, allowing for multiple broadcasting media and UGC streams to be ingested, matched, and coupled (Anagnostou, 2022; Filippidis et al., 2019; Tsolakis, 2023; Tsolakis et., 2023).

a.

Min		Score	
09:27		**60 - 52**	TIMMA, JANIS Three Pointer (1/1 - 3 pt)
09:25		60 - 52	SHENGELIA, TORNIKE Assist (3)
09:08	MCLEAN, JAMEL Turnover (1)	60 - 52	
09:07		60 - 52	TIMMA, JANIS Steal (3)
08:54	MANTZARIS, VANGELIS Foul (3)	60 - 52	
08:54		60 - 52	HUERTAS, MARCELINHO Foul Drawn (5)
08:43		**60 - 54**	SHENGELIA, TORNIKE Two Pointer (5/8 - 13 pt)

b.



c.

08:54		43 - 35
08:46		43 - 35
08:44		43 - 35
08:42		43 - 35
08:40	PRINTEZIS, GEORGIOS Def Rebound (1)	43 - 35
08:32	MILUTINOV, NIKOLA Two Pointer (4/6 - 9 pt)	45 - 35
08:30	PRINTEZIS, GEORGIOS Assist (3)	45 - 35
08:07		45 - 35
08:06	Def Rebound (2)	45 - 35
07:42	PRINTEZIS, GEORGIOS Two Pointer (5/8 - 10 pt)	**47 - 35**
07:22		**47 - 38**
07:20		47 - 38

Possession Event

start action: Def Rebound
start player: Printezis, Georgios
start time: 08:40
peak action: Two Pointer
peak player: Milutinov, Nikola
peak time: 08:32
end action: score change (45-35)
end player: Home Team
end time: 08:32
extra action: Assist
extra player: Printezis, Georgios
extra time: 08:30
topic: home successful offence

d.

Printezis took over after the break with a couple of jumpers, a great assist to Milutinov and a show-stopping dunk that restored a 49-38 Olympiacos lead. Diop followed a slam with a three-pointer and Granger added a driving layup that caused the hosts stop the game at 51-45. Agravanis hit a triple, Thompson added a driving basket and a power layup by Jamel McNeal made it 60-49 through three quarters. A three-pointer by Janis Timma and a layup off a great spin move by Shengelia gave Baskonia hope at 60-54 early in the fourth quarter. Agravanis rescued the hosts with a jump hook and Papanikolaou struck twice from downtown to restore some breathing room. Huertas singlehandedly brought Baskonia within 71-64 with under two minutes left, but it proved to be too little, too late for the guests.

00:41	MCLEAN, JAMEL Two Pointer (1/3 - 2 pt)	60 - 49
00:41	PAPANIKOLAOU, KOSTAS Assist (3)	60 - 49

PLAY-BY-PLAY

Two Pointer

skos:narrower: lay up
skos:narrower: dunk
skos:narrower: jump-shoot
skos:narrower: tip-in

Figure 5: Textual processing of the play-by-play (PBP) descriptions for basketball event identification and indexing: (a). example of an indicative PBP segment; (b). PBP events identification flow; (c) events detection and indexing (d). summaries and different vocabulary terms (Filippidis et al. 2018a).

Specifically, voice commanding was considered an alternative input to speed up the annotation process, indicating the class of the associated action (dribble, shoot, pass, etc.) by voice. A multilingual limited-vocabulary speech recognition system (in Greek and English languages) was developed to detect around fifty buzzwords associated with basketball actions. One thousand words per language and user were recorded for a total of five different (5) users to serve system training and testing purposes. Despite the small dataset size, the system achieved excellent performance with a 0.4% word error rate in Greek and 3.18% in English, probably due to pronunciation deficits (Gougousis, 2017). Similar audio recognition was tested in the noise conditions of real basketball games, with poorer accuracy. While buzzword detection performance on the commentary speech was decreased, audio-driven event detection was significantly helpful (detecting audience reaction, referee whistles, court air horns, etc.) (Filippidis et al., 2019; Sfyrakis, 2023).

It is essential to mention that the deployed audio processing utilities have yet to be fully exhausted since recent advantages that significantly extend the offered capabilities have been marginally tested and/or incorporated. For instance, blind vocal separation methods have become more powerful combining AI-driven with traditional approaches (Braun et al., 2021; Kalliris et al., 2016; Tsilfidis & Mourjopoulos, 2011), also being available in state-of-the-art applications that average users can efficiently operate with impressive results (e.g., https://www.lalal.ai/, https://ultimatevocalremover.com/, https://fadr.com/stems). Thus, the commenting voices within a TV or radio broadcasting stream can now be much more easily separated from stadium-crowded sounds, thus resulting in precise speech-to-text conversion (even in the real-world conditions of a basketball game) and/or accurate buzzword recognition, which could be further combined with the textual indexing. Moreover, since the official video broadcasting streams can be accurately timed with the textual events by applying machine vision to the superimposed scoreboard (Figure 6), audio synching cues can be instrumental in aligning extra/UGC content and applying metadata augmentation to all the synchronized segments. Specifically, machine learning solutions facilitate scoreboard detection and tracking, with subsequent match clock extraction supported by Optical Character Recognition (OCR) techniques (Anagnostou, 2022). Therefore, a link between the media/play timeline and the match reference clock is established, with textual event metadata to be easily incorporated and propagated to all the coupled footage (Figure 6).

Figure 6: Detection, recognition, and extraction of the match clock to align the textual events with the timelines of the video segments (Anagnostou, 2022).

Figure 7 illustrates the conceptual mineCourt approach with its envisioned functionalities. The modular architecture and integration aspects are of primary focus, in such a fashion that will enable knowledge transfer to similar cases of media coverage and management practices of other (sports) events. The diagram encompasses input sources, proposed unimodal processes, and multimodal fusion, all aimed at delivering content enhancement, documentation, analysis, and management automations. In today’s news ecosystem, apart from mainstream media, news websites, blogs, and UGC content shared through social media platforms have a significant impact on shaping the news reporting landscape, especially for sports events that captivate users’ attention and receive great audience engagement (Dimoulas, 2022; Katsaounidou et al., 2019; Touri & Kostarella, 2017). Hence, mineCourt embodies all the available sources of information contributing to the coverage of an athletic event, namely official broadcasting channels (radio, TV, web streaming), UGC streams captured at the events venues or in associated occasions, add-on material (reports, analyses, interviews, etc.) provided by the sport organizers, media professionals/ journalists, and plenary users/citizen journalists, social media posts that circulate all this information through shares, reactions, and comments. The semiotics behind the co-placed mis-organized/non-aligned circles in the entrance of Figure 7 (left rectangle with cyan dotted line) is that multiple non-alike streams are formed within each content production–distribution cell, requiring integrated/uniform treatment. Therefore, massive unstructured multimodal content items represent the mineCourt inputs, which must be handled in unimodal/semantic processing and multimodal fusion to enhance media assets and management automations. After that, an ontology framework can serve adaptive (sports) events

detection, alignment, and indexing, encompassing data linking mechanisms, semantic meta-data processing, and data analysis tools.

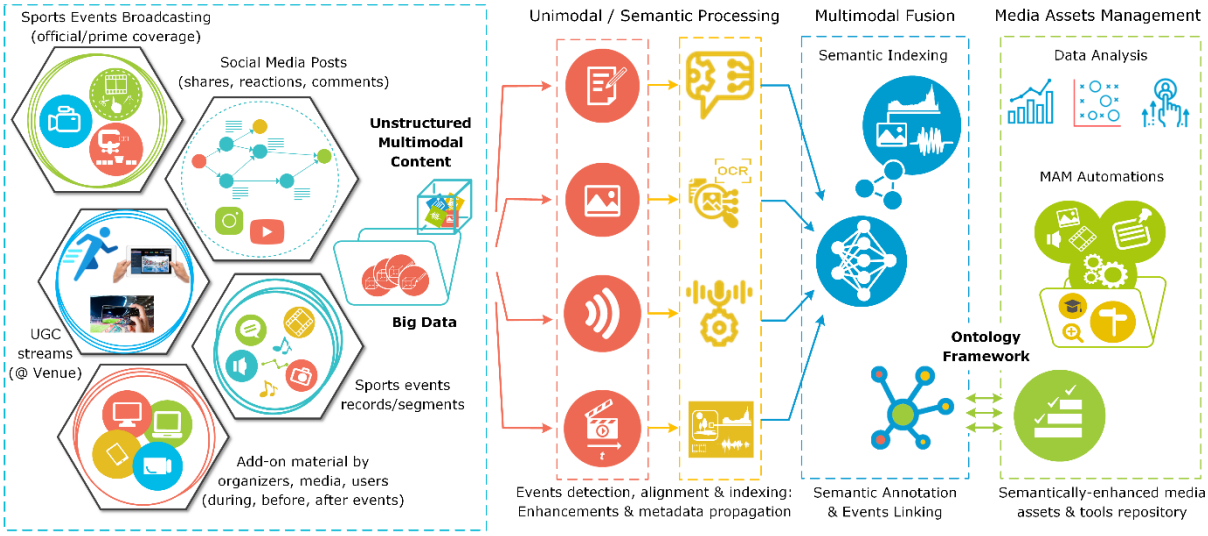


Figure 7: Conceptual diagram and functionalities of the proposed mineCourt framework.

Pilot Results and Discussion

The present study delves into the importance of intelligent sports data documentation, indexing, and management automations, integrating content enhancement and metadata augmentation mechanisms. As a result, a modular architecture is proposed, comprising various sub-systems that can be flexibly combined to address specific requirements. As a result, an integration framework is proposed, containing data-driven solutions with associated dataset repositories that are customizable to meet corresponding needs, coupled with a corresponding data description and linking ontology. This framework provides a systematic approach to managing the diverse range of sports-related data, unlocking opportunities for semantic analysis, annotation, and data linkage. Also, it establishes a sophisticated platform for organizing, retrieving, and interpreting interconnected content, spanning from player insights to fan reactions. This outcome also positively answers the stated research hypothesis: Data-driven/AI-assisted solutions are technologically mature and effective in meeting the desired objectives of improving quality and automating media assets management in the sports/basketball domains. By identifying and utilizing multiple/multimodal streams that describe the same sports event, raw data streams are converted into actionable insights, benefiting various stakeholders such as coaches, sports analysts, journalists, and enthusiasts, enhancing the overall sports experience.

Elaborating on the above remarks in relation to the inputs and data flow of Figure 7, apart from the main media coverage of a sports event, multiple linked assets are equally important or more, even if they have entirely different spatiotemporal characteristics. For instance, with the occasion of a basketball game, similar content of previous seasons, including players, teams,

and analysis perspectives, might be as valuable/enjoyable for media professionals, sports analysts, the involved (and other) basketball teams, and the public audience. Even when someone is interested only in the main event, the availability of manifold streams can cover multiple aspects. Hence, radio broadcasting can stimulate emotions different from those caused by the TV coverage experience, with more vivid/thorough speech narrations. Likewise, UGC posts provide subjective views from multiple angles, encompassing versatile semantics, conceptional, and sentimental insights (i.e., regarding cheering sounds, public reactions, events-related marginal details, small-scale/local incidents, etc.).

Furthermore, textual descriptions of PBP and web-streaming services usually contain stricter/more accurate (and less verbose) points to facilitate the time-indexing process. The superimposed video scoreboard can be used for extracting match clock as a synching reference to texts and other modalities (which can be easily deployed with reduced computational load), thus serving alignment and semantic metadata propagation (i.e., spatiotemporal, contextual, and emotional labels). Content syncing is imperative for subsequent audio/video enhancements by selecting the best quality footage and/or deploying time delay compensation and spatial/imaging filtering techniques. The sound modality seems preferable in this direction due to its lightweight nature and omnidirectional presence, and it can also serve speech recognition, buzzword detection, audio-driven segmentation, and indexing, which can be further utilized in multimodal fusion. The available material can also drive the generative synthesis of missing viewpoints through AI. Most importantly, all these processes are to be controlled under an ontology framework to deliver all the coupling between data, metadata, events linking/ aligning, and algorithmic tools.

Extending the above, Table 1 showcases individual components developed and tested within the mineCourt framework, many stemming from Ph.D. research and diploma dissertations conducted at the Laboratory of Electronic Media of Aristotle University (https://e-media.jour.auth.gr/). Their presentation aims to demonstrate the maturity and validity of the proposed integration rather than evaluate individual modalities. Even so, most results are comparable to the current state of research. Hence, the crafted football and basketball annotators, to start with, have been put in related (subjective) evaluation procedures and proved that they do not lack in the offered usability, visualization and indexing functionalities, compared even to market solutions of media/sports domains (Dimoulas, 2020; Nikolopoulou & Papagianni, 2015; Papachristou; 2016). A manifold of audio classification and speech recognition tools have been implemented with remarkable accuracy (≥90%) in segmenting basketball events (commenting speech, referee whistles, court air horns, audience reactions like boos, applauses, and cheering) (Filippidis et al. 2019; Sfyrakis, 2023), detecting buzzwords associated with basketball actions (Gougousis, 2017), extracting bimodal segmentation footprints for content matching and alignment purposes (Dimoulas & Symeonidis, 2015). Textual/PBP processing enabled successful detection and indexing of possession events (Filippidis et al., 2018a; 2018b), synchronized with video timelines using visual scoreboard

detection/tracking and game clock extraction (with remarkable accuracy of over 90-95%) (Anagnostou, 2022). Finally, a prototype sports indexing and retrieval front-end has been developed and coupled with a unified data ontology for serving data description, linking, and metadata propagation/enhancement facilities (Anagnostou, 2022; Tsolakis, 2023; Tsolakis et al., 2023). This ontology has been established and validated with basketball games of three different leagues (NBA, Greek Basket League, Euroleague) (Tsolakis, 2023; Tsolakis et al., 2023). Most of all, the proposed integration framework drives best practices adoption and knowledge transfer to (sports) media examples with similar management/analysis needs.

Table 1: Presentation of deployed mineCourt components with the associated maturity, integration, progress, and future expectations.

Related work	Description	Outcomes: Maturity – Integration
Nikolopoulou & Papagianni (2015)	Radio commentary of a football match with (spatial) annotation of the possession events	- football annotator (web service) - football possessions visualization and indexing
Papachristou (2016)	Basketball actions vocabulary and state transition diagram	- basketball annotator (web service) - basketball possessions visualization and indexing
Dimoulas & Symeonidis (2015)	Syncing shared media through audio-driven bimodal segmentation	- audio-driven segmentation and footprint extraction method - synching of shared media with emphasis on sports events
Gougousis (2017)	Limited-vocabulary automatic speech recognition of basketball buzzwords/actions	- accurate recognition of basketball words (in clean recordings) - speech recognition dataset/code - voice dictation/input to basketball annotator
Filippidis et al. (2018a; 2018b)	Detection and semantic indexing of basketball possession events using textual streams (PBPs)	- basketball annotator (web service) with enhanced control automations - natural language processing of textual PBP descriptions to define/index basketball possession events
Filippidis et al. (2019), Sfyrakis (2023)	Audio-driven basketball events segmentation and indexing	- audio segmentation and classification of basketball events (code/datasets) - speech and buzzword recognition - voice separation & future expectations (enhancement, re-mixing, etc.)
Anagnostou (2022)	Audiovisual basketball events detection and spatiotemporal localization/indexing	- scoreboard detection/tracking and match clock extraction in official videos - textual/PBP events processing and synching with video timelines - content indexing/retrieval interface and algorithmic backend (code/dataset)

Tsolakis (2023), Tsolakis et al. (2023)	Ontology-based Framework for Sports Media Management	- data handling and managing framework - semantic data linking and metadata enhancement - basketball vocabulary adaptation with events indexing mechanisms

References

Anagnostou A. (2022). *Audiovisual event detection and localization techniques in sports content*. Unpublished Diploma Thesis (https://github.com/Nasos-Anagnostou/Audiovisual-event-detection-and-localization-techniques-in-sport-content/tree/main/The%20final%20stack). School of Electrical and Computer engineering, Aristotle University of Thessaloniki, Greece.

Bellamy, R. V. (2009). Sports media: A modern institution. In *Handbook of sports and media* (pp. 66-79). Routledge.

Boulianne, S. (2020). Twenty years of digital media effects on civic and political participation. *Communication research*, 47(7), 947-966.

Braun, S., Gamper, H., Reddy, C. K., & Tashev, I. (2021, June). Towards efficient models for real-time deep noise suppression. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 656-660). IEEE.

Chatzara, E., Kotsakis, R., Tsipas, N., Vrysis, L., & Dimoulas, C. (2019). Machine-assisted learning in highly-interdisciplinary media fields: A multimedia guide on modern art. *Education Sciences*, *9*(3), 198.

Dimoulas, C. A., & Veglis, A. (2023). Theory and Applications of Web 3.0 in the Media Sector. *Future Internet*, *15*(5), 165.

Dimoulas, C. A. (2022). Cultural heritage storytelling, engagement and management in the era of big data and the semantic web. *Sustainability*, *14*(2), 812.

Dimoulas, C. (2020). Media Technologies & Sports: A timeless relationship of strong interaction. In Twigil's N. & Mastrogiannakis D. (Eds.) *Sport and Media: Rhetoric, identities and representations,* pp. 211-242. Zygos, Thessaloniki.

Dimoulas, C., Veglis, A. A., & Kalliris, G. (2018). Semantically enhanced authoring of shared media. In *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 6476-6487). IGI Global.

Dimoulas, C. A. (2016). Audiovisual spatial-audio analysis by means of sound localization and imaging: A multimedia healthcare framework in abdominal sound mapping. *IEEE Transactions on Multimedia*, *18*(10), 1969-1976.

Dimoulas, C. A., & Symeonidis, A. L. (2015). Syncing shared multimedia through audiovisual bimodal segmentation. *IEEE MultiMedia*, *22*(3), 26-42.

Dimoulas, C. A., Kalliris, G. M., Chatzara, E. G., Tsipas, N. K., & Papanikolaou, G. V. (2014). Audiovisual production, restoration-archiving and content management methods to preserve local tradition and folkloric heritage. *Journal of Cultural Heritage*, *15*(3), 234-241.

Dimoulas, C. A., & Kalliris, G. M. (2013, May). Investigation of wavelet approaches for joint temporal, spectral and cepstral features in audio semantics. In *Audio Engineering Society Convention 134*. Audio Engineering Society.

Dimoulas, C. A., Avdelidis, K. A., Kalliris, G. M., & Papanikolaou, G. V. (2007). Joint wavelet video denoising and motion activity detection in multimodal human activity analysis: application to video-assisted bioacoustic/psychophysiological monitoring. *EURASIP Journal on Advances in Signal Processing*, *2008*, 1-19.

Dimoulas, C. A., Avdelidis, K. A., Kalliris, G. M., Papanikolaou, G. V., & Vegiris, C. (2007, May). Automated audio detection, segmentation and indexing, with application to post-production editing. In *Audio Engineering Society Convention 122*. Audio Engineering Society.

Filippidis, P. M., Dimoulas, C., Bratsas, C., & Veglis, A. (2018). A multimodal semantic model for event identification on sports media content. *Journal of Media Critiques [JMC]*, *4*(14).

Filippidis, P. M., Dimoulas, C., Bratsas, C., & Veglis, A. (2018). A unified semantic sports concepts classification as a key device for multidimensional sports analysis. In *2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)* (pp. 107-112). IEEE.

Filippidis, P. M., Vryzas, N., Kotsakis, R., Thoidis, I., Dimoulas, C. A., & Bratsas, C. (2019). Audio Event Identification in Sports Media Content: The Case of Basketball. In *Proceedings of the 146th Audio Engineering Society Convention,* Dublin, 20-23 March 2019.

Fuss, F. K., Subic, A., Strangwood, M., & Mehta, R. (2013). *Routledge Handbook of Sports Technology and Engineering*. New York, USA: Routledge.

Gougousis H. (2017). *Speech recognition and voice command assisted human-machine interaction* (in Greek). Diploma Thesis (http://ikee.lib.auth.gr/record/294611). School of Electrical & Computer engineering, Aristotle University of Thessaloniki, Greece.

Jack, K. (2011). *Video* demystified*: a handbook for the digital engineer*. Elsevier.

Jenkins, H., & Ito, M. (2015). *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. John Wiley & Sons.

Jo, Y., Oh, S. W., Kang, J., & Kim, S. J. (2018). Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3224-3232).

Kalliris, G., Dimoulas, C., & Papanikolaou, G. (2001). Broad-band acoustic noise reduction using a novel frequency depended parametric Wiener filter: implementations using filter-bank, STFT and wavelet analysis/synthesis techniques. *Audio Engineering Society Preprint, Proceedings of the 110th AES Convetion, Amsterdam*.

Kalliris, G., & Dimoulas, C. (2009, May). Audiovisual content management issues for the new media environment. In *Proceedings of the International Conference on New Media and Information: Convergences and Divergences, Athens, Greece* (pp. 3-6).

Kalliris, G., Matsiola, M., Dimoulas, C. A., & Veglis, A. (2014, July). Emotional aspects in Quality of Experience and Learning (QoE & QoL) of audiovisual content in mediated learning. In *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications* (pp. 198-203). IEEE.

Kalliris, G. M., Dimoulas, C. A., & Uhle, C. (2016). Special Issue on Intelligent Audio Processing, Semantics, and Interaction. *Journal of the Audio Engineering Society*, *64*(7), 464-465.

Kapela, R., Świetlicka, A., Rybarczyk, A., & Kolanowski, K. (2015). Real-time event classification in field sport videos. *Signal Processing: Image Communication*, *35*, 35-45.

Katsaounidou, A., Dimoulas, C., & Veglis, A. (2019). *Cross-media authentication and verification: emerging research and opportunities: emerging research and opportunities*. IGI Global. https://doi.org/10.4018/978-1-5225-5592-6.

Katsaounidou A., Matsiola M., Dimoulas C. (2018). Controversial content broadcasting and the role of television. In *Proceedings of the International Conference 50 Years of Greek Television,* Epikentro, Thessaloniki 2018.

Konstantoudakis, K., Vrysis, L., Tsipas, N., & Dimoulas, C. (2018). Block unshifting high-accuracy motion estimation: a new method adapted to super-resolution enhancement. *Signal Processing: Image Communication*, *65*, 81-93.

Konstantoudakis, K., Vrysis, L., Papanikolaou, G., & Dimoulas, C. (2018). High accuracy block-matching sub-pixel motion estimation through detection of error surface minima. *Multimedia Tools and Applications*, *77*, 5837-5856.

Kotsakis, R., Dimoulas, C. A., Kalliris, G., & Veglis, A. (2014, July). Emotional descriptors and quality of experience (QoE) metrics in evaluating mediated learning. In *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications* (pp. 232-237). IEEE.

Kotsakis, R., Dimoulas, C., Kalliris, G., & Veglis, A. (2014). Emotional prediction and content profile estimation in evaluating audiovisual mediated communication. *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, *2*(4), 62-80.

Liu, R., Liu, Z., & Liu, S. (2021). Recognition of basketball player's shooting action based on the convolutional neural network. *Scientific Programming*, *2021*, 1-8.

Liu, L. (2021). Objects detection toward complicated high remote basketball sports by leveraging deep CNN architecture. *Future Generation Computer Systems*, *119*, 31-36.

Lucas, A., Lopez-Tapia, S., Molina, R., & Katsaggelos, A. K. (2019). Generative adversarial networks and perceptual losses for video super-resolution. *IEEE Transactions on Image Processing*, *28*(7), 3312-3327.

Matsiola, M., Dimoulas, C.A., Kalliris, G., & Veglis, A., (2015). Augmenting User Interaction Experience through Embedded Multimodal Media Agents in Social Networks. In: J. Sahlin, (Ed). *Social Media and the Transformation of Interaction in Society*. (pp.188–209) Hershey PA: IGI Global.

Mourjopoulos, J. N. (2005). The evolution of digital audio technology. In *Communication acoustics* (pp. 299-319). Berlin, Heidelberg: Springer Berlin Heidelberg.

Müller, M. (2007). *Information retrieval for music and motion* (Vol. 2, p. 59). Heidelberg: Springer

Nikolopoulou V. & Papagianni C. (2015, Oct.). *Broadcasting and intelligent analysis of audiovisual sport events: the case of football* (in Greek). Unpublished Thesis. School of Journalism and Mass Communications, Aristotle University of Thessaloniki, Greece.

Owens, J. (2015). *Television sports production*. New York & London: Focal Press (5th Ed.).

Palioura, M., & Dimoulas, C. (2022). Digital storytelling in education: A transmedia integration approach for the non-developers. *Education Sciences*, *12*(8), 559.

Pan, Z., & Li, C. (2020). Robust basketball sports recognition by leveraging motion block estimation. *Signal Processing: Image Communication*, *83*, 115784.

Papachristou T. (2016, September). *Intelligent analysis of sports events – the case of basketball*. Unpublished Thesis (in Greek). Unpublished Thesis. School of Journalism and Mass Communications, Aristotle University of Thessaloniki, Greece.

Podara, A., Giomelakis, D., Nicolaou, C., Matsiola, M., & Kotsakis, R. (2021). Digital storytelling in cultural heritage: Audience engagement in the interactive documentary new life. *Sustainability*, *13*(3), 1193.

Podara, A., Matsiola, M., Maniou, T. A., & Kalliris, G. (2019). Transformations of television consumption practices: An analysis on documentary viewing among post millennials. *Particip. J. Audience Recept. Stud*, *16*(2), 68-87.

Raney, A. A., & Bryant, J. (2009). *Handbook of sports and media*. Routledge.

Romney, M., & Johnson, R. G. (2020). Show me a story: narrative, image, and audience engagement on sports network Instagram accounts. *Information, communication & society*, *23*(1), 94-109.

Rota, C., Buzzelli, M., Bianco, S., & Schettini, R. (2023). Video restoration based on deep learning: a comprehensive survey. *Artificial Intelligence Review*, *56*(6), 5317-5364.

Saridou, T. A., & Veglis, A. (2021). *Exploring the Integration of User-Generated Content in Media Organizations Through Participatory Journalism*. In Encyclopedia of Information Science and Technology, Fifth Edition (pp. 1152-1163). IGI Global.

Sfyrakis E. (2023). *Detection and recognition of audio events in basketball games* (in Greek). Diploma Thesis (http://ikee.lib.auth.gr/record/351509). School of Electrical and Computer engineering, Aristotle University of Thessaloniki, Greece.

Sidiropoulos, E., Vryzas, N., Vrysis, L., Avraam, E., & Dimoulas, C. (2019). Growing media skills and know-how in situ: Technology-enhanced practices and collaborative support in mobile news-reporting. *Education Sciences*, *9*(3), 173.

Touri, M., & Kostarella, I. (2017). News blogs versus mainstream media: Measuring the gap through a frame analysis of Greek blogs. *Journalism*, *18*(9), 1206-1224.

Tsardoulias, E., Thallas, A. G., Symeonidis, A. L., & Mitkas, P. A. (2016). Improving multilingual interaction for consumer robots through signal enhancement in multichannel speech. *Journal of the Audio Engineering Society*, *64*(7/8), 514-524.

Tsilfidis, A., & Mourjopoulos, J. (2011). Blind single-channel suppression of late reverberation based on perceptual reverberation modeling. *The Journal of the Acoustical Society of America*, *129*(3), 1439-1451.

Tsolakis, N., Vryzas, N., Dimoulas, C., Maga-Nteve, C., Meditskos, G., & Vrochidis, S. (2023, July). An Ontology-Based Framework for Sports Media Data Interpretation. In *2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA)* (pp. 1-8). IEEE.

Tsolakis N. (2023). *Ontology-based framework for semantic content in sports: application in* basketball *games retrieval*. Diploma Dissertation (http://ikee.lib.auth.gr/record/349551). Post-Graduate Program on Advanced Computing and Communication Systems, School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece.

Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2018). Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 9446-9454).

Vázquez-Herrero, J., & López-García, X. (2019). When media allow the user to interact, play and share: Recent perspectives on interactive documentary. *New Review of Hypermedia and Multimedia*, 25(4), 245-267.

Vegiris, C. E., Avdelidis, K. A., Dimoulas, C. A., & Papanikolaou, G. V. (2008). Live broadcasting of high definition audiovisual content using HDTV over broadband IP networks. *International Journal of Digital Multimedia Broadcasting*, *2008*.

Vegiris, C., Dimoulas, C., & Papanikolaou, G. (2009, May). Audio Content Annotation, Description, and Management Using Joint Audio Detection, Segmentation, and Classification Techniques. In *Audio Engineering Society Convention 126*. Audio Engineering Society.

Vryzas, N., Vrysis, L., Kotsakis, R., & Dimoulas, C. (2018, September). Speech emotion recognition adapted to multimodal semantic repositories. In *2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)* (pp. 31-35). IEEE.

Vryzas, N., Vrysis, L., Kotsakis, R., & Dimoulas, C. (2021). A web crowdsourcing framework for transfer learning and personalized speech emotion recognition. *Machine Learning with Applications*, *6*, 100132.

Vryzas, N., Vrysis, L., & Dimoulas, C. (2021). Audiovisual speaker indexing for Web-TV automations. *Expert Systems with Applications*, *186*, 115833.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., ... & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops* (pp. 0-0).

Wu, L., Yang, Z., Wang, Q., Jian, M., Zhao, B., Yan, J., & Chen, C. W. (2020). Fusing motion patterns and key visual information for semantic event recognition in basketball videos. *Neurocomputing*, *413*, 217-229.

Zhang, K., Zuo, W., Gu, S., & Zhang, L. (2017). Learning deep CNN denoiser prior for image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3929-3938).