DanXe: an Extended Artificial Intelligence Framework to Analyze and Promote Dance Heritage

Article *in* Digital Applications in Archaeology and Cultural Heritage · May 2024

DOI: 10.1016/j.daach.2024.e00343

6 authors, including:

Lorenzo Stacchio
University of Macerata
38 PUBLICATIONS 176 CITATIONS

SEE PROFILE

Silvia Garzarella
University of Bologna
5 PUBLICATIONS 2 CITATIONS

SEE PROFILE

Pasquale Cascarano
University of Bologna
31 PUBLICATIONS 172 CITATIONS

SEE PROFILE

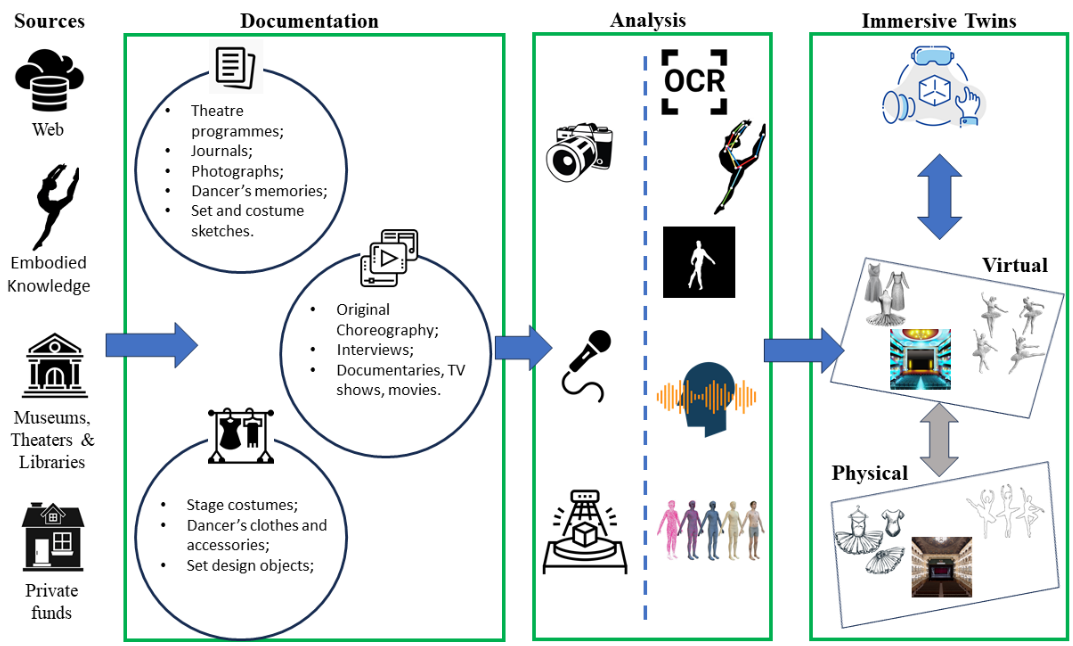
Allegra De Filippo
University of Bologna
28 PUBLICATIONS 125 CITATIONS

SEE PROFILE

Graphical Abstract

DanXe: an Extended Artificial Intelligence Framework to Analyze and Promote Dance Heritage

Lorenzo Stacchio, Silvia Garzarella, Pasquale Cascarano, Allegra De Filippo, Elena Cervellati, Gustavo Marfia



Highlights

DanXe: an Extended Artificial Intelligence Framework to Analyze and Promote Dance Heritage

Lorenzo Stacchio, Silvia Garzarella, Pasquale Cascarano, Allegra De Filippo, Elena Cervellati, Gustavo Marfia

- Introduction of DanXe, a novel framework combining Artificial Intelligence strategies for digitization and automatic analysis with Extended Reality solutions for immersive visualization.

- DanXe allows to put the basis for the creation of Digital Twins of the dance world, to preserve, and promote Dance Heritage, for both intangible and tangible materials. This also facilitates simulations, idea dissemination, and multi-modal analysis.

- DanXe follows a multidisciplinary approach, that aims at enriching the storytelling of dance heritage, and supports researchers through automated data analysis while extending its implications to various creative industries and cultural heritage preservation efforts.

- DanXe was applied and validated through a case study focusing on the artistic legacy of Rudolf Nureyev, emphasizing comprehensive organization and analysis of materials related to his impact, offering demonstrations of the strategies of Danxe for its analysis and exploration.

DanXe: an Extended Artificial Intelligence Framework to Analyze and Promote Dance Heritage

Lorenzo Stacchio[a,*], Silvia Garzarella[b], Pasquale Cascarano[b], Allegra De Filippo[c], Elena Cervellati[b], Gustavo Marfia[b]

[a]*Department of Political Sciences, Communication and International Relations, University of Macerata,*
[b]*Department of the Arts, University of Bologna,*
[c]*Department of Computer Science and Engineering, University of Bologna,*

Abstract

Motivated by the need to leverage technologies to enhance the preservation, accessibility, quantitative data analysis, and valorization of Dance Heritage, this work introduces DanXe, a framework based on Artificial Intelligence and Extended Reality for the digitization, automatic analysis, and immersive manipulations of tangible and intangible cultural heritage material. Our contribution offers a framework to define the Digital Twin of documentary assets associated with dance heritage, enriching the storytelling of its traditions, performers, and specific languages. The implications of the proposed framework extend beyond the realm of dance, impacting various creative industries and cultural heritage preservation efforts. The approach is validated through a specific case study, linked to the artistic legacy of the dancer and choreographer Rudolf Nureyev intending to narrate his artistic impact. Moreover, we contribute with the description of an effective strategy for the comprehensive organization and analysis of all materials related to the legacy of the dancer.

Keywords: Artificial Intelligence, Extended Reality, Cultural Heritage, Computational Dance

---

*Corresponding Author

Email addresses: `lorenzo.stacchio2@unibo.it` (Lorenzo Stacchio), `silvia.garzarella@unibo.it` (Silvia Garzarella), `pasquale.cascarano2@unibo.it` (Pasquale Cascarano), `allegra.defilippo@unibo.it` (Allegra De Filippo), `e.cervellati@unibo.it` (Elena Cervellati), `gustavo.marfia@unibo.it` (Gustavo Marfia)

1. Introduction

Between September 29 and October 17, 2003, the General Conference of UNESCO established a pivotal document in Cultural Heritage Studies: the *Convention for the Safeguarding of Intangible Cultural Heritage*. Within this document, dance plays a significant role but studying it demands a particularly careful approach. This *Convention* builds upon and expands the principles established by the *World Heritage Convention* of 1972, which was ratified in Paris on November 16, 1972, to identify, valorize, and protect cultural and natural heritage, including monuments, groups habitats, and sites shaped by both human activity and natural processes (art. 1-2) (1).

Later, in 2003, considering "*the profound interdependence between Intangible Cultural Heritage, Tangible Cultural Heritage, and Natural Heritage*" (2) UNESCO endorsed a new agreement that supplements the original by acknowledging the significance of practices and traditions transmitted from one generation to the next, namely: "*the practices, representations, expressions, knowledge, skills – as well as the instruments, objects, artifacts and cultural spaces associated therewith – that communities, groups and, in some cases, individuals recognize as part of their cultural heritage*" (2).

This heritage, referred to as Intangible Cultural Heritage (ICH) by UNESCO, is *"transmitted from generation to generation, is constantly recreated by communities and groups in response to their environment, their interaction with nature and their history, and provides them with a sense of identity and continuity, thus promoting respect for cultural diversity and human creativity*" (2). Following the provisions of the 1972 *Convention*, the new document aims to safeguard cultural diversity and promote respect for the cultural practices of communities worldwide. It does so by offering Member States a range of tools to identify, document, preserve, and promote their ICH.

To keep track of the results, two lists were established: the *Representative List of the Intangible Cultural Heritage of Humanity*, which highlights the cultural diversity of the assets and underscores their importance, and the *List of Intangible Cultural Heritage in Need of Urgent Safeguarding*, which has a more operational focus aimed at protecting at-risk assets with the involvement of the international community. In addition to these two instruments, a third one is added, the *Register of Good Safeguarding Practices*, dedicated to sharing particularly effective safeguarding experiences (3; 4).

The criteria for inscribing a heritage on either of the two lists or for contributing to the *Register* are established by precise *Operational Directives* (5), which today recognize 730 assets (referred to as “elements”), distributed across 145 countries. According to the *Convention*, these elements have characteristics attributable to one or more of the five domains identified in Article 2.2, indicated by alphabetical letters from A to E as listed below:

(A) oral traditions and expressions, including language as a vehicle of the intangible cultural heritage;

(B) performing arts;

(C) social practices, rituals and festive events;

(D) knowledge and practices concerning nature and the universe;

(E) traditional craftsmanship.

Within the 2003 *Convention*, dance occupies a prominent position. A search for the word “dance” in both *Lists* reveals the presence of 97 elements distributed across 67 nations, predominantly referencing traditional and ritualistic forms of dance. The situation remains practically unchanged when shifting the search parameter from “dance” to “choreography.” The heritages of the 45 registered countries are predominantly traditional and confirm, by their distribution, a strong presence in Asia and the Pacific (6), as depicted in Figure 1. Apart from the practice of modern dance in Germany, officially recognized as UNESCO Heritage in 2022, western ‘theatrical’ dance lacks explicit representation, though it could be fully classified within the sectors identified by the document and warrants equal attention in terms of safeguarding.

The importance of this form of dance - rooted in medieval treatises, first codified at the court of Louis IX, and widely performed on global stages from the 18th century to today’s contemporary forms (7) - arises from the specificity of the associated heritage, combining a substantial amount of written documentation with a legacy of knowledge and techniques continually passed down by masters and performers. It’s worth noting how such diverse material is not only diversified in its typology (both tangible and intangible) but also in its spatiotemporal definition and the use of different technical languages, with repercussions in terms of accessibility. Furthermore, due to the specificity of performing arts, a comprehensive analysis that considers both intangible and tangible elements is essential. Only through such an integrated approach can the plural identity of dance heritage be effectively restored.

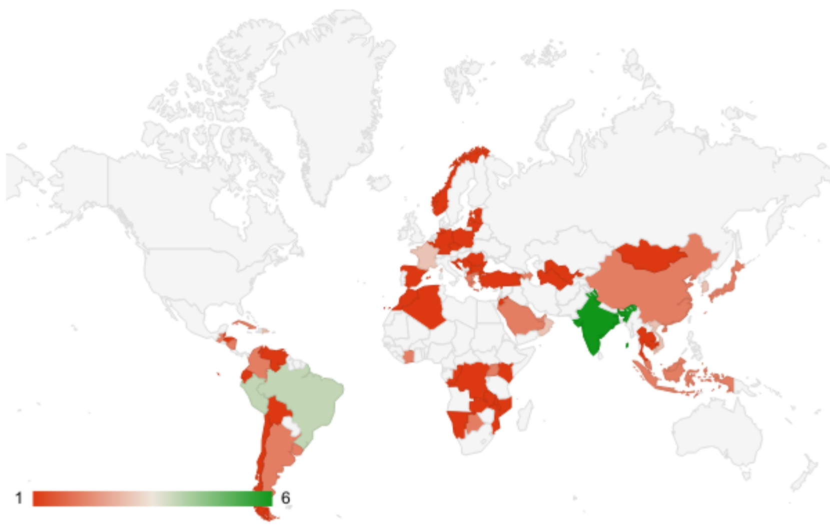


Figure 1: Intangible Cultural Heritage Distribution (2008-2023). The heatmap shows the distribution of the 97 “dance” elements recognized by UNESCO across the world.

In 2003, as stated in the *Convention*, there was “no instrument for the safeguarding of intangible cultural heritage” and the safeguarding measures identified both at national (art. 11-14) and international (art. 16-18) levels were primarily based on their recognition, inventorying, and promotion. This phenomenon also touches those CH items, which expose both intangible and tangible nature, such as dance. Such a kind of material increases the difficulty of its preservation and promotion. Moreover, without any kind of automatic support, dance researchers and analysts would consume an important amount of time, to both reach, correlate, and perform analysis on the sources and the materials under examination. This often forced experts, to both concentrate on a single modality (or a limited multi-modal one) and choose qualitative analysis, instead of quantitative ones (8).

Twenty years after the *Convention*, novel technologies, such as Artificial Intelligence (AI) and Extended Reality (XR) had an exponential growth of application across a spectrum of disciplines towards the humanities, in particular to CH (9). It is worth noticing that, the combination of AI and XR paradigms was recently defined as a novel branch of research, named Extended Artificial Intelligence (10; 11). In this landscape, the involvement of such digital technologies, in the dance domain, was applied considering diverse data, tasks, and use cases (12; 13). Despite these vast applications, to the best of our knowledge, the following aspects of the integration of such technologies were overlooked.

First, those applications focused on analyzing specific computational dance aspects while not concentrating on a holistic approach that could cover the majority of its tangible and intangible material. Secondly, those digital analyses and outcomes remain static, without providing dynamic triggers to evolve and connect them according to how the world changes (12; 14).

To cope with both of these aspects, we could resort to AI and XR, which are fundamental pillars to (i) create general digital libraries of complex multimedia and 3D objects to preserve the cultural assets under examination; (ii) create interactive experiences to let them be accessed by the majority of the population (i.e., storytelling) (13). In such a context, Digital Twins (DT) paradigms amount to a unique opportunity to include the temporal dimension for such assets. DTs are computer-based models, either physical or virtual, that simulate, emulate, or mirror real-life entities such as objects, processes, humans, or human-related features (15). Considering dance, DTs could help professionals and researchers to update information related to dance material in the digital realm, but also produce simulations, facilitate the dissemination of creative ideas, and provide an easier way to analyze its multi-modal nature, particularly employing XR paradigms (16). Following such a line of thought, we here introduce a novel Extended Artificial Intelligence framework, DanXe, which is a careful composition of (i) AI algorithms designed for the digitization and automatic analysis of tangible and intangible materials (aiming at creating the DT of the dance cultural heritage); (ii) XR solutions to enable immersive visualization of the derived insights to demonstrate how the digital space is suitable to address the complexity of performative immateriality. This would, in principle, provide an alternative space for the simultaneous analysis of all the aspects that constitute the essence of dance.

To design such a framework, we followed a multidisciplinary approach, combining computer science and dance research backgrounds, to implement functionalities for exploring and preserving the entire range of documentary assets associated with dance heritage, enriching the storytelling of its traditions, its performers, and its specific languages. Our framework can support researchers by enhancing accessibility and providing valuable insights through automated data analysis processes, which are typically confined to qualitative analyses and narrow topics (17; 8). To demonstrate its effectiveness and flexibility, DanXe was validated through a specific case study: the artistic legacy of the dancer and choreographer Rudolf Nureyev (1938-1993), intending to narrate his artistic impact. However, such an application would only be possible with a preliminary comprehensive organization and analysis of all the assets related to Rudolf Nureyev. This process represents another significant contribution to our work, underscoring our commitment to enhancing understanding and appreciation of his immense artistic contributions. Finally, it is worth highlighting that the implications of the proposed framework extend beyond the realm of dance, impacting various creative industries and cultural heritage preservation efforts.

2. Dance Research Background

In this Section, we want to elucidate the dance research background that motivated the definition of our system and its exemplary application to theatrical dance and the case of a famous dancer: Rudolf Nureyev. The decision to focus on a single type of dance form arises from the specificity of the associated heritage, which creates a unique multi-modal documentation, defining an embodied knowledge continually passed down by professionals (18). However, this source is characterized by widespread distribution on a global scale, resulting in dispersion and multiplicity of sources, a significant typological variance (texts, audiovisual, memories, and embodiments), and limited accessibility due to the use of technical languages.

Long before the *Convention*, the reflection on the ephemeral nature of performing arts has extensively occupied twentieth-century debate in the field of Cultural and Performance Studies, especially its later decades. As a result of a widespread 'performative turn' in practices, during the 1960s sector studies shifted away from the prevailing text-centric approach to prioritize the analysis of processes and, consequently, the actual 'documentability' of the performative act (19). Hence, the need to approach the study of performing arts memory with two different perspectives: the more traditional investigation of traces left behind the performative act (texts, scores, photographs, drawings, costumes, and other archival documents) and the "incorporated knowledge" (20) that the performance leaves in those who have seen it (the audience), practiced it (the dancers), and transmitted it (the masters).

Only an integrated analysis of these aspects can help restore the plural identity of heritage such as dance, whose complexity—unlike what often occurs with spoken theater—is further augmented by semantics entrusted to movement. The instance of ballet, in this regard, stands as particularly emblematic. Consider nineteenth-century masterpieces like *Swan Lake* (1877) and *The Sleeping Beauty* (1889), still staged with the signature of the choreographer who conceived them but with results decidedly distant from the nineteenth-century editions and with modalities that combine direct teacher-student transmission with more historicist operations (21). Except for initiatives with explicitly reconstructive purposes, eg.

the choreographer Alexei Ratmansky's works, indeed, the restaging of the classical repertoire is generally curated by *maîtres répétiteurs*, experienced dancers who follow the professionals accompanying them throughout the learning and perfecting process of the choreography (22; 23). The mastery of the teachers is mainly given by the authority of their experience, a manifesto of embodied but strictly specialist knowledge and not necessarily reliable from a philological point of view. To this, a rich documentary heritage must be added, consisting of choreographic notations, artists' memoirs, stage notes, sketches, etc. Hence the urgency of an approach suitable for accommodating both dimensions.

2.1. Rudolf Nureyev (1938-1993)

We here depict the historical context surrounding Rudolf Nureyev's career, examining his impact on the ballet repertoire and the challenges of defining his legacy within dance historiography. Furthermore, we provide an overview of the diverse array of archival materials and sources available for studying Rudolf Nureyev's legacy, ranging from historical archives and theater documentation to personal collections and testimonies from collaborators and fans. Finally, we highlight the challenges and opportunities inherent in accessing and interpreting these sources focusing on the complexities of preserving and reconstructing the artistic heritage of a figure as multifaceted as Nureyev.

Rudolf Nureyev was a creative interpreter of the ballet repertoire, capable of revitalizing male roles through an innate stage presence and a charismatic approach to technical virtuosity. Constantly moving from one stage to another, the Russian dancer was able to assimilate the styles of the great national ballet schools, combining the exuberance of the Soviet tradition with the airiness of European technique, without disregarding hybridization with the free forms of modern dance, inclusively encompassed within his repertoire (24).

Despite international fame, interpretive talent, and authorial abilities, however, Nureyev still lacks a defined place within dance historiography. The mythologization of the character, which has deeply influenced the communication of his image, seems to be reflected in the specific literature, abundant in details about his personal life but lacking in studies dedicated to a precise framing of his biographical and authorial profile.

Placing the 'Nureyev myth' within the broader panorama of the so-called "cultural Cold War" (25) can enhance the bibliographic corpus and provide the basis for a better interpretation of existing documentary sources, yet the specificity of his style and its actual impact on dance historiography remains largely unexplored. In-depth knowledge of the artist's repertoire remains tied to the memories of direct witnesses of the creative process, particularly the dancers of the Ballet de l'Opéra national de Paris, which Nureyev directed from 1983 to 1989 and followed as *maître de ballet* until 1992. Drawing on dancers' testimonies, through direct observation of dance practice, is essential to enable the profiling of Nureyev's specific contribution, but it is not sufficient to ensure adequate objectivity.

In the absence of a sufficient bibliography, the cross-analysis of all available documentary sources has become indispensable and, with all the specificities and problems involved, can provide a suitable case study to verify the possibilities of enhancing and valorizing the artistic legacy of a dancer.

2.2. Materials and sources

The career of Rudolf Nureyev unfolded between the 1960s and 1980s, a period marked by a significant explosion of mass media that propelled his global renown, resulting in extensive coverage across print media (both periodicals and specialized publications), television (specials, interviews, shows), and cinema (documentaries and films as director and actor).

This distribution reflects exactly the general ones previously highlighted for theatrical dance. Moreover, considering its activity, Nureyev's traces were left all over the world, creating a sparse spatial distribution, which is here reported in Figure 2.

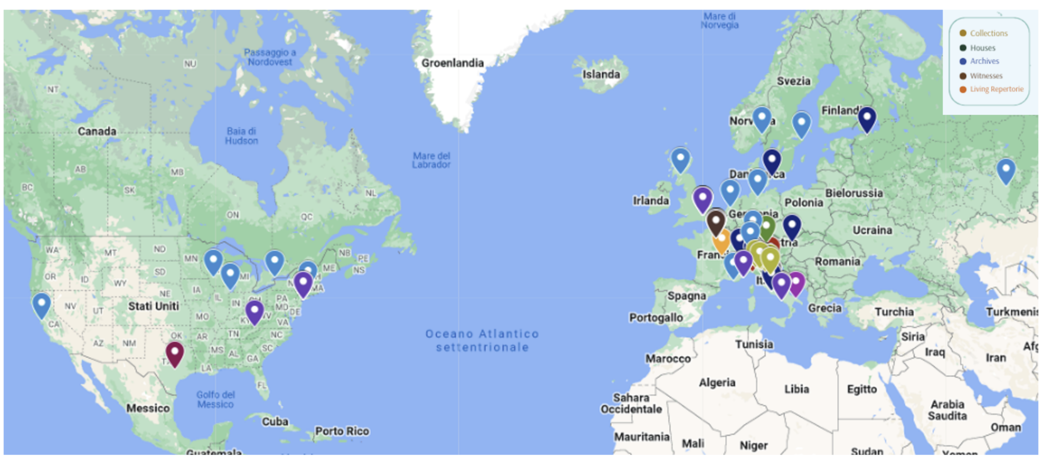


Figure 2: Distribution of part of both tangible and intangible assets related to the dancer Rudolf Nureyev (map updated until 2023).

Of notable importance is the documentation preserved within the historical archives of theaters where Nureyev performed as a guest or resident artist, particularly: the Royal Opera House (London), Opéra National de Paris (Paris), Wiener Staatsoper (Vienna), Metropolitan Opera House (New York), Teatro alla Scala (Milan), Mariinsky Theatre (St. Petersburg). This list, potentially extendable to all theaters that hosted the dancer throughout his career, corresponds to predominantly paper-based documentary assets (programs, photographs, stage sketches, costumes, press reviews) and audiovisual recordings (rehearsal footage, premieres, audio recordings of performances and interviews).

In the case of Nureyev, while manuscript sources are practically absent, the presence of heritage assets is notably significant, encompassing not only objects such as furniture, textiles, personal items, and collections but also work-related correspondence and fan letters. These materials, often held by private collectors, are owned by the Rudolf Nureyev Foundation, which has been entrusted with managing the dancer's legacy since 1993.

The Foundation oversees the Archives de la Fondation Rudolf Noureev, which are distributed among: the New York Public Library (New York), Rambert – Maude Lloyd and Wallace Potts Archive (London), Royal Opera House Collections (London), CND – Centre national de la danse (Pantin), CNCS – Centre national du costume de scène (Moulins). Within these collections, the French *Fonds Noureev* stands out as the most substantial resource for studying the dancer's career, and, like the others, comprises diverse sources that belonged directly to the artist or were donated to the Foundation by his closest collaborators.

The Rudolf Nureyev Foundation also manages the administration of Nureyev's 'intangible' legacy, namely his original choreographic works, which continue to be performed in major opera houses worldwide and are a permanent part of the Opéra national de Paris repertoire. The re-staging of these choreographies is entrusted to selected artists, *maître répetiteurs* who were direct witnesses of the creation processes and who are tasked with coordinating and supervising every production bearing the dancer's signature.

In the absence of written productions certified by Nureyev himself (such as choreographic scores or notes), the survival of the repertoire depends on the embodied memory of the performers and videos produced during the premieres. Despite the temporal proximity, both sources are vulnerable due to the fragility of their mediums (linked to personal memory in the former case, and to outdated formats like VHS in the latter). This renders the preservation of the original works particularly precarious, a fragility exacerbated by the intricacy of Nureyev's choreographic language, which is highly technically complex and thus subject to greater variation when interpreted by different dancers.

To these 'institutionalized' records, to fully reconstruct the artistic legacy of a dancer of such widespread fame, must necessarily be added the testimonies of spectators, ballet corps dancers, partners, and the many fans – who have experienced and contributed to the 'Nureyev phenomenon', a concept challenging to encapsulate in purely scientific terms yet profoundly influential in shaping the dancer's authorial identity and measuring his impact. Within this framework, such testimonies primarily manifest as oral or audiovisual sources, analyzable through audio recordings, video documentation, and comparison of current dance practices with the dancer's distinctive style.

3. Related Works

Although there is a general lack of exemplary case studies regarding the preservation of repertoires, from a multi-modal perspective, the connection between dance and computing is deeply historicized. Starting from the pioneering experiences of Merce Cunningham ("People just don't understand the future possibilities of the computer and dance", he said) (26). In 1986 he started working with *Life Forms*, to digitally manipulate avatars of human bodies, providing a platform for choreographers to experiment with movement sequences and explore new choreographic ideas in a virtual space(27). His pioneering efforts recognized the potential of leveraging digital techniques to address challenges within the field of dance.

Later, in (28) the authors started using AI strategies to identify a movement sequence that would organically transition between two specified body postures. Furthermore, the authors of (29) introduced a collaborative video-based workflow accessible through a web-based platform known as *The Choreographer's Notebook* which enables choreographers and dancers to annotate and provide feedback on various aspects of the performance, including movement, timing, and expression. In 2016, the development of *BalOnSe* aimed to streamline the exploration and retrieval of multimedia content. This was achieved by leveraging metadata, including title information and featured dancers. Notably, the platform introduced a feature enabling users to search based on movement concepts. This functionality allows users to filter videos according to specific terms from a predefined vocabulary, drawing from previously submitted annotations. Other initiatives such as *SCHEDAR* (Safeguarding the Cultural Heritage of Dance via Augmented Reality) have embraced emerging technologies to archive, analyze, and thoroughly document intangible heritage works, particularly within the realm of dance (30).

(31) introduced *WhoLoDancE*, aimed at contributing to future technologies for dance education and creation, exploring opportunities for teachers, learners, and choreographers in the dance studio. In (32), the authors proposed an inter-

active VR platform designed to train couples (in Salsa dancing). The platform includes components such as a virtual partner with interactive controls, visual, and haptic feedback, and was tested with both experienced dancers and novices. In (33), to enhance the teaching efficiency of rhythmic gymnastics courses, the paper introduces a system combining VR image recognition technology and digital twins, tailored to the teaching needs of rhythmic gymnastics. Furthermore, AI tools tailored for the task of human posture recognition have been used in (34; 35) to classify dance movements and for the control and synthesis of basic dance gestures. Lastly, in (36), XR has been identified as having the potential to enhance dance learning by providing students with more information. The study developed an XR interactive learning system to aid teachers in dance instruction and explored the impact of XR on the process and outcomes of dance skill acquisition.

The aforementioned research works, presented in chronological order, underscore the deep historical roots of the connection between dance and computing. They provide evidence that this relationship has been extensively explored over time (37; 38).

In summary, the here-mentioned works applied AI and XR paradigms for the benefit of the dance world, with three main tasks: (i) digitizing physical material; (ii) automatic analysis of such digitized/digital representation; (iii) the visualizations of both in classical/immersive environment (16; 39).

However, none of the works described designs a general-purpose toolbox able to analyze, preserve, and promote, physical or digitized cultural heritage assets related to the dance world, without focusing on a specialized hardware-software architecture for a single task.

4. DanXe Framework

DanXe is an extended artificial intelligence framework designed to (i) digitize the majority of tangible and intangible cultural heritage dance material (ii) automatically analyze digital content to provide insights that can be re-used in further human-centered analysis, and (iii) provide a modular immersive interface to visualize any of those information in orchestrated Digital Twin (DT) environments. Its architecture is visually depicted in Figure 3.

As mentioned in Section 2, many different sources nowadays preserve, generate, and maintain tangible and intangible Cultural Heritage assets for the dance world. For example, Museums and Private funds collect and preserve: documents to textually store dance practices or dancers' memories, pictures and videos of exhibitions, soundtracks, and stage costumes/accessories. These items maintain

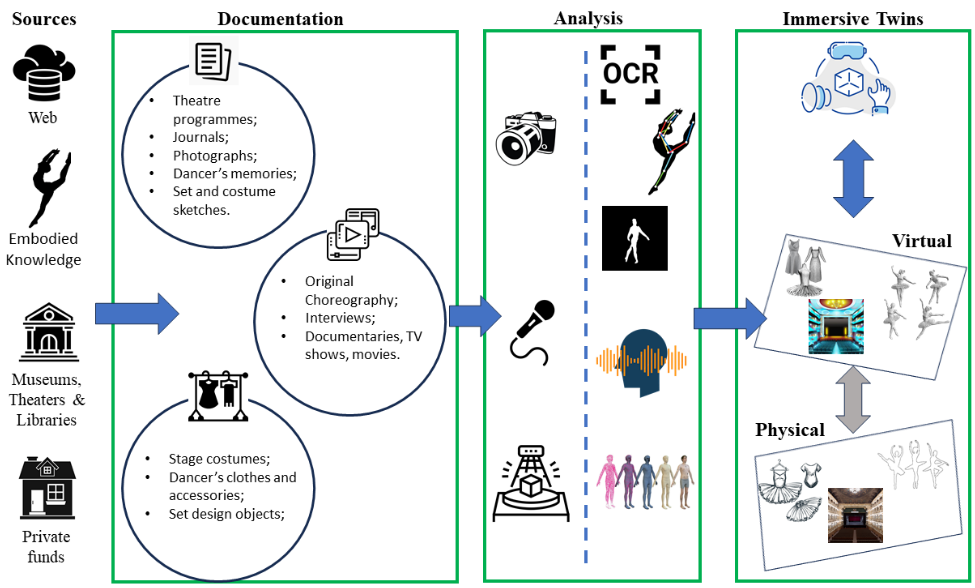


Figure 3: DanXe framework overview.

track of both tangible and intangible materials. Such assets could be digitized by adopting data-capturing devices, such as photography, microphones, or 3D scanners (included in the DanXe pipeline).

Such digitized information, along with native digital one, can be fed to the multi-modal analysis component of the DanXe framework. In this layer, different AI-based paradigms are employed to analyze such dance heritage materials. Its main functionalities are reported in the module diagram depicted in Figure 4.
The Textual Analysis Module encompasses diverse sources such as theatre programs, essays, journals, and dancer's memories. It associates this data with the extraction of references, and relevant topics while fetching relevant documents and their summaries. The Visual Analysis Module deals with image and video data generated by sources like dance shows, movies, and TV programs. Key activities involve the extraction of dance moves, identification of people, and synthesizing novel costume images from sketches. Lastly, this module is also responsible for extracting text from images depicting it, like theatre programs. The Audio Analysis Module examines music tracks and audio speeches to recognize emotion, and musical instruments, and provide transcription and synchronization of music or audio. The 3D object data module focuses on extracting 3D representations from images/video sources, in particular involving dancers and costumes. Finally, the

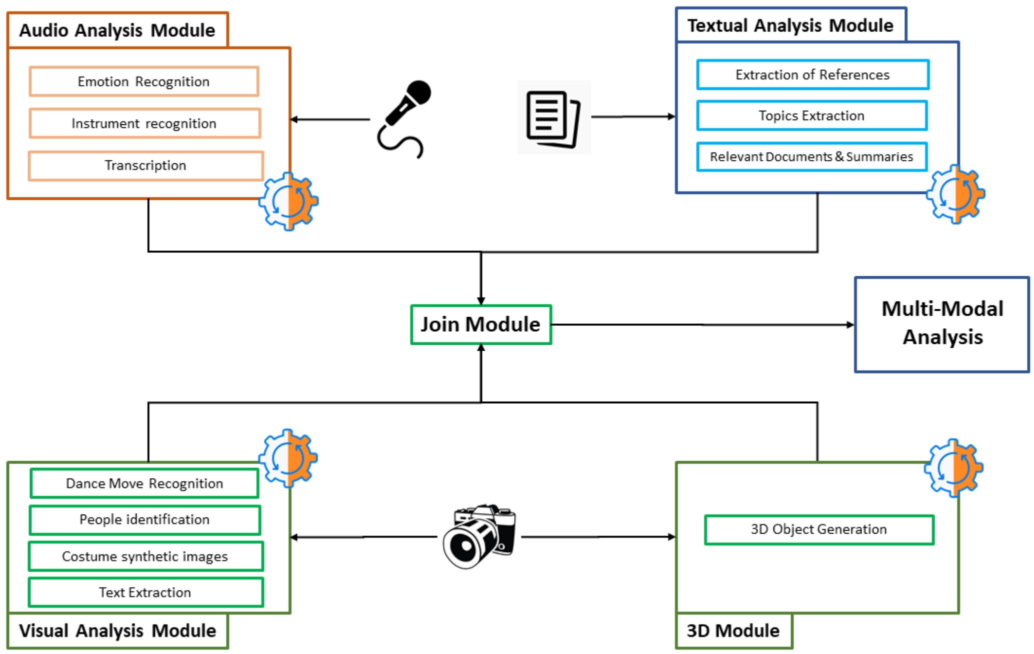


Figure 4: Module Diagram of the different DanXe analysis layer modules concerning the data from dance heritage material.

provided analyses are multimodally merged by the Join Module, which can be queried by any interface.
It is worth highlighting that, each of the analysis modules, is labeled with a continuous development symbol, indicating that the actual functionalities implemented in DanXe will be continuously updated and integrated with novel methods in a seamless fashion.

This analysis layer produces a double-edge output: (i) analytics that can be used by both researchers and professionals for several purposes (e.g., educational, choreography design, performance emotional impact evaluation); (ii) produced information that can be put to good use to define a DT of the dance world. This twin can be then empowered by XR visualization and manipulations, as proposed in (16). Considering the latter, we also provide a first design and development of the final layer of DanXe: Immersive Twins, where a user could visualize and manipulate multi-modal information related to a certain dance material in any degree of reality, through XR paradigms.

Holistically considered, DanXe provides a unique sequential step to completely map the dance world, which nowadays is mostly physically stored in the digital realm, through XAI paradigms. In the following sections, we will describe each of the logical layers proposed in the DanXe pipeline: Digitisation, Analysis, and Visualizations.

4.1. Digitisation

The world of dance is defined through physical items including costumes, sets, paper documents, audiovisuals, and other tangible objects. Those, along with human knowledge and memories, can be used to produce intangible material such as choreographies, masteries, productions, and performances in general. Such information is produced and maintained by various sources, like theaters, museums, libraries, and private funds, which, together with performers' embodied knowledge, emerge as the main ones.

These assets, however, are scattered across different sources without a defined structure for their organization. To this date, it is essential to define an arrangement for such material considering the specific domain of application and related use cases (as described in Section 2). After such data structuring, the digitization process allows projecting those assets from the physical world, into the digital realm (40; 41; 38). These could be permanently stored in databases for their preservation while making them more accessible for both fruition and analysis (41; 38). A vast audience of stakeholders, ranging from consumers to dance professionals (e.g., choreographers, dancers) and researchers, could benefit from this process, considering dimensions such as history, dance theory, practices, and creativity (38; 42; 39; 43).

To this date, a general framework for the projection and analysis of dance-related entities should also expose services for their digitization. Such interfaces should carefully combine hardware and software systems, optimized per each considered application (38). In DanXe we designed different entry points to capture physical signals and analyze them to be reconstructed in the digital realm (44; 45; 46). As depicted in Figure 3, the access to the digital realm is mediated by photography, audio capture, and 3D scanner hardware, which generate multimedia data describing what is captured in the real world: images, videos, audio, and 3D models (e.g., represented as point clouds or meshes (47)). These data structures can capture the majority of the documentation available in the field of dance (12; 48). It is worth noticing that such digitized information can be aggregated with the ones that are already stored in online platforms (e.g., Naxos [1]).

[1] https://www.naxosvideolibrary.com/

These raw and heterogeneous digital items define a set of data that can be used to extrapolate high-level knowledge, using automatic and semi-automatic paradigms (49), overcoming classical manual analysis approaches (12; 49). However, such kinds of outputs could only be obtained by carefully combining data analytics and AI paradigms, while integrating Human-In-The-Loop approaches (50).

4.2. Analysis

The analysis of textual documents, images, videos, and 3D models holds immense significance in unraveling the rich cultural heritage of dance. These assets are repositories of invaluable information, encapsulating choreographic nuances, historical performances, and artistic evolution. By subjecting these items to automatic or semi-automatic analysis, we can unlock a wealth of insights into the cultural heritage and artistic expressions of dance. Both such analyses can be done resorting to AI paradigms, which perform predictive analysis mimicking actions that would instead be performed manually by human experts, requiring less time with reasonable accuracy (12). It is so mandatory to understand which kind of reasoning dance professionals and researchers do in their everyday activities, to support them by automatizing repetitive and time-consuming ones. To this date, considering the possible space of application and the range of analysis, we initiated a process of isolation of the considered material, data, and tasks, and established methods to perform the mentioned analysis.

Considering the variety of the analysis that can be carried out for each kind of data, Deep Learning (DL) emerge as the most efficient and effective paradigms to adopt (38). This is because DL models expose state-of-the-art performances in several tasks for analyzing multi-media data in all the considered modalities (i.e., 2D/3D computer vision, audio signals, natural language processing). Moreover, recent works highlighted how DL approaches allow also for multi-modal analysis (e.g., image question answering, audio-video choreography synchronization) (51). For this reason, our framework integrates out-of-the-box DL-based techniques for the automatic analysis of such a corpus of data. Analyzing available DL approaches allowed us also to match the kind of data, and tasks, that we could examine by adopting such methods. The result of such a process was schematized and reported in Table 1. The next sections will refer to such a Table to discuss the different adopted approaches for textual, visual, and audio, detailing inputs, methods, and outputs for each of them.

Table 1: Different kinds of analysis performed with Deep Learning approaches on dance heritage material codified in different types of data.

	Data type	Output	Method	Refs
Text	Theatre Program, Essays, Journals, Dancer's memories	Extraction of References	Named Entity Recognition & Relations Extraction	(52)
	Essays, Journals, Dancer's memories	Topics extractions	Topic Modeling	(53)
	Essays, Journals, Dancer's memories	Relevant Documents & Summaries	Information Retrieval & Summarization	(54)
Image/Video	Dance shows, Movies	Dance moves recognition	Keypoint Regression and analysis	(55; 56)
	Dance shows, Movies, TV programs	People identification	Object Segmentation, Re-ID	(57; 58)
	Sketches	Costume synthetic images	Sketch2Image	(59)
	Theatre Program	Text extraction	OCR	(60)
Audio	Music track/Audio speech	Music/Audio Emotion	Music emotion recognition	(61)
	Music track	Musical instrument recognition	Instrument classification	(62)
	Music track/Audio speech	Music or audio transcr/synchro	Audio signal analysis	(63)
3D Object	Image/Video	3D Obj of Dancer/-costume	3D Reconstruction	(64)

4.2.1. Text

Despite Dance being a mostly visual-audio domain, a huge amount of data that captures knowledge regarding different aspects of its lifecycle is codified in texts (65; 66). Essays, commentaries, notations, biographies, and online platforms like websites and social media are just examples of databases of unstructured text from which we could extract knowledge related to general or specific dance topics.

Considering such documents, it is worth noticing the thoughtful reflection performed by (66) on hypertextuality exposed by dance essays and journals. The authors referred to hypertextuality as the interconnectedness of information within a digital textual document, where words and phrases are linked to other relevant information or resources. This work also highlighted how it enables non-linear reading experiences, facilitating exploration and explicit delineation of discourse shape, organization, and functions (66). Finally, the author highlighted how those links should be connected not only to other textual data but also to the multi-modal dance materials (e.g., images, video, and audio representations of tangible and intangible items).

However, we here hypothesize that the linking domain should be expanded to metadata and structural annotations, aiming at describing schemes and categorizations within it (12). Such structure provides a two-wise advantage: furnishes additional information about documents; and poses the foundations for a more efficient query mechanism and concept discovery.

In the past, hypertextuality was often performed by analytic, yet manual, methods to find correspondences and analyze them together to make new concepts emerge (concerning dance research). However, the variety (e.g., essays, commentaries, journals, memories) and quantity of textual documents and the possible metadata/labels that can be inferred to facilitate its organization, prevent this manual analysis approach (workforce and time-consuming). To this date, we could resort to DL paradigms to automatically analyze these texts.

Considering now the plethora of DL models and tasks they were taught to solve (67), we here consider a subset of them, to extract the following knowledge from the text, which covers the aforementioned use cases: *topic modeling*, *information retrieval*, and *named entity recognition*. How those techniques are composed and applied to existing documents is depicted in Figure 5, which also describes the core component of the middleware, which aims at extracting raw text from heterogeneous data sources.

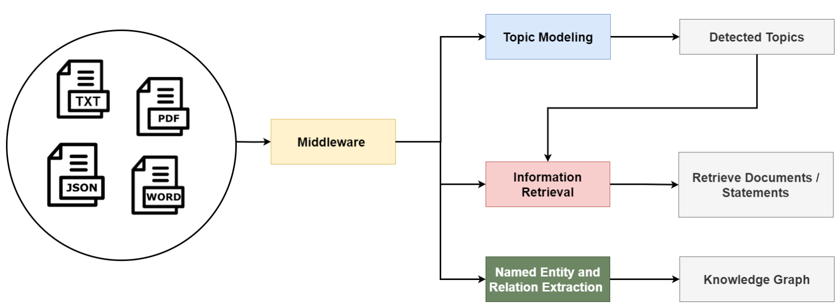


Figure 5: The DanXe Textual Analysis architecture employs topic modeling to pinpoint key discussions and optimizes information retrieval through named entity/relation extraction.

An exemplar dance use case that adopts some of the here mentioned techniques, will be detailed in the Rudolf Nureyev use case study (Section 5).

Topic modeling. Topic modeling (TM) is a technique used to analyze large text collections by identifying underlying topics. Each topic represents a semantic concept, offering a latent interpretable representation of sentences and documents. TM is pivotal for document analysis since allows one to organize and summarize documents. This specific outcome is then used to improve information retrieval, content recommendation, and sentiment analysis systems (68). As stated in (68) a topic is a cluster of constructs that together describe a pure semantic meaning. In practice, It is an idealized notion of a document that is as pure as possible. Considering the sub-category of neural topic modeling, the state-of-the-art pipeline involves using a pre-trained language model, such as BERT, and then clustering the result embeddings with approaches like HDBSCAN (53). Then, those embeddings could be first projected in a lower-dimensional space using dimensionality reduction algorithms like UMAP (69). The representation of those clusters is then reinforced by using a class-based Term Frequency-Inverse Document Frequency or cTF-IDF to obtain the words representing each topic (53). With such a pipeline, named BERTopic, we can create an organized and searchable textual knowledge base, allowing us to categorize new documents and retrieve documents that are talking about the same topic. For example, a professional from the dance world could easily recover the statements in all the available document that talks about a certain topic.

Along with topic modeling, we could re-use the knowledge extracted by pre-trained language models to predict other categories from them, like the sentiment (70). This highlights the fundamental role of pre-trained language models like BERT, GPT, or LLAMA (71).

Information Retrieval & Summarization. Information retrieval (IR) involves acquiring resources from unstructured data to meet user needs. Nowadays, with DL approaches we can learn how to construct ranking search models automatically, outperforming traditional ones(72). In particular, considering a pre-trained language model, we could extract word or sentence latent representations as vectors. Then use those representations (as singles or with an aggregation mechanism) to query another document in latent space, retrieving the most similar ones, which corresponds to the less distant in the learned latent space. In the dance domain, however, several challenges arise, among which: (i) the length of the documents; and (ii) the high variety of the adopted vocabulary. Considering (i) state-of-the-art approaches involving the usage of sentence-level, passage-level score, and passage-level representation aggregation, as possible solutions to process lengthy texts. Sentence-level score aggregation methods, divide documents into sentences and combine top scores retrieved by models like BERT (72; 73). On the other hand, passage-level scores segment long documents into overlapping passages to facilitate document-level retrieval by aggregating passage scores, which could also be combined with representations using strategies like pooling and attention mechanisms (72; 73). Considering (ii) instead, we could resort to ad hoc ranking strategies, considering also cross-domain approaches (72; 73; 74). For example, (74) introduced an ad hoc DL document retrieval approach, again using BERT, integrating sentence-level evidence for document ranking and transferring relevance models across different domains. The same approach could be adopted for fine-tuning BERT specifically for the very detailed vocabulary of the dance world (75). Using then embeddings extracted by such models on the relevant documents, we could perform automatic text summarization to provide concise and accurate summaries, thereby significantly reducing the human effort required in processing large volumes of text (76).

Named Entity Recognition and Relation Extraction. Among textual information extraction, Named Entity Recognition (NER), emerges as key and focuses on identifying named entities within text. At the same time, Relation Extraction (RE) deals with extracting relationships between these entities. For example, one could recognize subjects and their parental or friendship relationships (77). These extracted entities and relationships are fundamental for NLP tasks like machine translation, question-answering, and knowledge graph creation (KG) (78). KG is a form of structured human knowledge, composed of a structured representation of facts, consisting of entities, relationships, and semantic descriptions (78). In KGs: entities can be real-world objects and abstract concepts; relationships represent the links between entities; Semantic descriptions of entities, and their relationships contain types and properties with well-defined meanings (78). Considering the latter, authors from (79; 80) explored the challenges of KG definition from art-historical texts, focusing specifically on Giorgio Vasari's seminal work "The Lives of The Artists", by implementing an entity recognition and linking model exploiting co-references, temporal information extraction, and identifying artwork-related information. This KG could be built on top of the DL Relation Extraction Model like REBEL, an auto-regressive approach that frames relation extraction as language a seq2seq task (52). Tools like this could also be adopted in the dance world, to discover dancers, archival materials, and their relationships (e.g., dance - styles, original-interpretations, master-pupil, location - performances).

4.2.2. Vision domain: images and videos

Dance is primarily a visual art form, relying heavily on movements, gestures, and expressions to convey meaning, emotion, and narrative (18). For example, movements and gestures allow dancers to create shapes, patterns, and sequences in space communicating ideas and emotions to the audience.

The cited items amount to an intangible patrimony created by the human body and knowledge. However, we should also consider tangible ones. Firstly, dance costumes can help to establish the mood, setting, and characteristics of a dance piece/character, creating also a strong connection with fashion studies (81). At the same time, we should consider pictures depicting information related to dance shows or dancers' memories like theatre playbills, press reviews, videos/photos of the performances on stage or in the studio, etc.

Considering such variety, relying on computer vision techniques to extract knowledge and identifying patterns in images and videos depicting such intangible and tangible knowledge amounts to a fundamental step. To this date, we here consider some DL models to solve tasks that could support professionals and researchers in extrapolating knowledge from the mentioned material, considering both images and videos. In such a context, the fundamental pillars are those defined as Vision DL Backbones, i.e., those DL models that were pre-trained on images to solve particular tasks and that learned good latent representations that could be used in downstream tasks, adopting transfer knowledge techniques. The main DL backbones are those following two kinds of formalism: Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) (82; 71). Those two approaches provide different pros and cons and reach state-of-the-art results in different tasks or domains (82; 71; 83), requiring to exploitation of both to have a holistic state-of-the-art toolbox.

In particular, DanXe includes both CNN and ViT models for image/video Optical Character Recognition (OCR), Key-point Regression, tracking, and analysis, providing a baseline for Dancers Re-Identification and Segmentation. Finally, DanXe provides the baseline structure to exploit Generative DL starting from existing heritage material.

OCR. The process of analyzing such visual dance documents begins with OCR which allows the extraction of text from images, enabling the digitization of archives containing program leaflets, articles, and handwritten notes (such as the dancer's memories or choreographic scores). Once the text is extracted, it can be subjected to further analysis using several computational techniques such as text mining or natural language processing, like the ones detailed in Section 4.2.1. An underlying difficulty in developing general-purpose OCR systems lies in the lack of general guidelines to define them. OCR has to deal with a high variety of documents, considering both hand-written and digitally printed ones. Among both, there could be ligatures that can be recognized either as individual code points, characters that cannot be represented, or that are not included in the encoding standards. At the same time, images representing those documents could be taken with consumer devices, creating illumination and posing problems, which often do not appear for scanned printed documents (84; 85). The same difficulties are reflected in the dance domain, considering the aforementioned documents. Currently, the best solution amounts to adopting a DL-based OCR which was trained on the most possible varied data, adopting an *in the wild* approach, that should improve their generalization (60; 84; 86)). To this date, many OCRs are available in literature (87; 88). However, considering the analysis made in (88) we here consider that the body of developed OCRS exposes the following robust and modular DL pipeline: (i) detection component employing algorithm like CRAFT (89); (ii) recognition model exploiting approaches like Convolutional Recurrent Neural Network (90); (iii) a final decoding step utilizing algorithm like CTC (91). In such a subset, DL pipelines like EasyOCR emerge as the best option, considering also that it was trained following an *in the wild* approach (60).

Dance Human tracking: Keypoint regression/analysis and re-identification. In this paragraph, we will discuss all the key components to implement the DanXe pipeline to track and classify dancers by analyzing pictures/videos of them. In particular, we will discuss such an architecture considering the use case of choreographies, as visually reported in Figure 6.

Human tracking with deep learning involves detecting humans in images or videos using methods like object detection and Human Pose Estimation (HPE) while associating unique identifiers with each detected individual to track them across images (or video frames) (92). This information could then be adopted for making a posterior analysis of great interest for the dance world (e.g., steps classification and transcriptions, and thus choreography preservation). HPE is a fundamental challenge for computer vision and a core step for the analysis of dance (14; 93; 94). HPE involves estimating the configuration of human body parts from input data captured by sensors, in particular images and videos. HPE provides geometric and motion information about the human body that has been applied to a wide range of applications from emotional analysis to XR. This task can be also extended to multi-person full-body pose estimation. In both single and multiple cases, dance adds a challenge: a fast temporal change in human move-

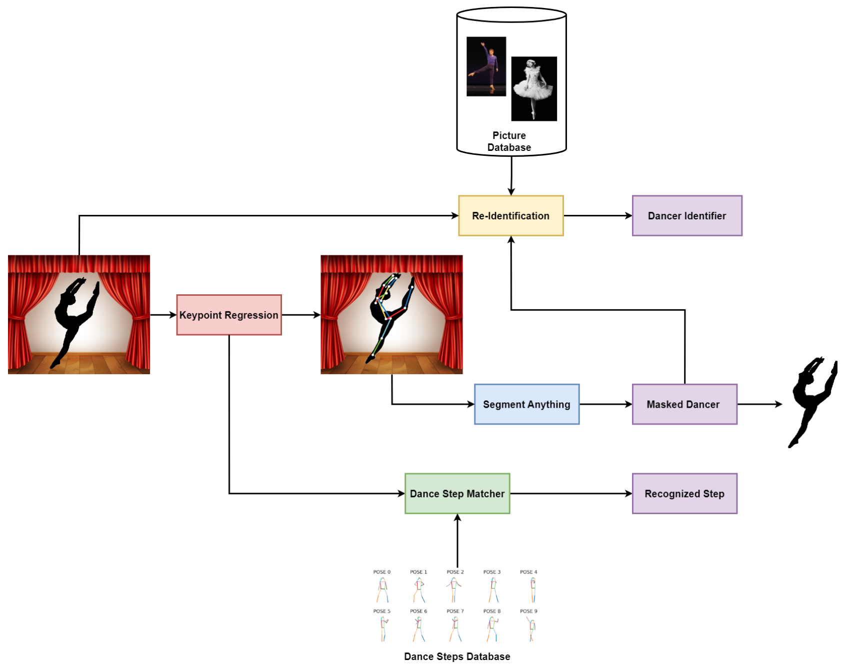


Figure 6: DanXe architecture for Dancer's tracking, re-identification, step classification. Final outputs are highlighted in purple.

ments. For such tasks, DL-based approaches outperformed classical computer vision ones (95; 94). Beyond the 2D domain, recent methods approached 3D HPE, obtaining accurate 3D pose annotations is much more difficult than its 2D counterpart, posing the basis to overcome classical Motion capture systems (38; 94).

In the arena of DL-based HPE models, an important resource amounts to the ones that were trained with an *in the wild* approach. This would be particularly feasible for the dance domain, which should be analyzed in very different settings and scenarios. Despite many works having adopted such an approach (96; 94), one emerged for its efficiency, efficacy, modularity, and ease of use: AlphaPose (93). Its authors proposed it to enhance the efficiency and reliability of top-down frameworks for efficient human pose estimation mainly by: (i) adjusting detection confidence and non-maximum suppression thresholds to provide more candidates for subsequent pose estimation; (ii) eliminating redundant detected poses; (iii) introducing a novel optimization for pose distance parameters; (iv) introducing a symmetric integral key points regression method to accurately localize key points across different scales. AlphaPose was optimized to perform not only single image 2D HPE but also videos, considering a 3D pose estimation exploiting deformations of the known Skinned Multi-Person Linear Model (SMPL) databases (93).

Having at our disposal reliable temporal estimated human key points, we could now proceed with their analysis. Several studies focus on action classification, future action prediction, and action application in human posture information, facilitating human-computer interactions across various domains like dance (97; 98; 99). For example, in dance training, research primarily addresses evaluating correct dance poses and assessing dancers' adherence to masters' instructions, often utilizing joint position and angle information (98). This step amounts to what is defined as action recognition, where actions in dance are contextualized as steps (98; 99). Steps recognition could go beyond training, facilitating researchers' and professionals' analysis of previous choreography from images and videos. For such analysis, spatiotemporal constraints are often used to track a single person while analyzing key points' positions within videos (or images with motion prediction) (99). In the dance domain, authors from (98) introduced a real-time method using smartphone cameras, to correct body ratio differences with Partial-Affine transformation and simultaneously evaluate joint positions and angles for objective feedback. This approach could be extended and generalized using DL vision backbones. For example, (100) analyzed dance choreography analysis by introducing DanceMoves, a system to directly describe dance moves through body positions and movements (firstly inferred with HPE). Exploiting data normalization, feature extraction, and similarity measures, enables interactive analysis, comparison, and search of dance poses, catering to both qualitative and quantitative analysis needs. Another work adopted a CNN approach to extracting feature maps and then classified 200 Indian classical dance poses performed by 10 dancers (101). However, they did not rely on human key points, defining a general bias towards visual aspects more than human joints themselves (102). However, hybrid approaches can performed, by adopting, for example, CNN that could learn how to discriminate among a fixed step of actions, implementing particular pooling operations or embedding layers, to let the model look at features related to estimated key points (99). All the mentioned technologies could be expanded in the 3D domains using mesh deformation predictions, allowing for a more detailed representation compared to traditional skeletal models (93; 103).

Recognizing the pose and tracking it, provide us implicit information regarding the position of the human body in space, easing masking and detection mechanisms. This information could be also exploited to recognize and mask out the pixels regarding the human body. To this aim, one could adopt DL models like

Segment Anything (58) exploiting geometric prompts querying that correspond to the coordinates of the predicted key points to improve the masking mechanism and automatically selecting only the ones we are interested in.

The masked pixels could then be fed to what is known as Mask-guided Person Re-Identification (ReID) (104; 105) to automatically detect who is the person we are attempting to track. In particular, ReID models aim at retrieving a person of interest across multiple views in different domains through a visual query mechanism (105). Modern DL approaches rely on metric learning, where a visual backbone learns meaningful cues to match images or videos of the same individuals. This is obtained by forcing the model to learn representations of data points in such a way that semantically similar points are closer together in the embedding space while dissimilar points are farther apart. A well-established approach is based on the combination of Siamese networks and the triplet loss, which analyzes triplets of images as input, where each triplet consists of images of the same person (anchor, and positive example) and different people (negative example). The network is trained to minimize the distance between embeddings of positive pairs and maximize the distance between embeddings of negative ones (105). Considering now ReID systems exploiting binary masks, authors of (104) introduced a training pipeline including the binary segmentation masks of RGB people datasets to cope with the diverse background clutters, variations in viewpoints, and body poses, reaching state-of-the-art results. Following the same approach, we could expand such ReId to classical image retrieval for many visual dance items (such as stage costumes).

The composition of all the mentioned technologies could be re-used for the dance world, in particular, to track and segment dancers and dance performances, which is the main purpose of it, considering the huge amount of work that, otherwise, professionals should do to achieve the same result. For this, we have included them in the DanXe framework.

Costume Sketch to Image Generation. Dance heritage, is made by a high amount of visual sketches, mostly related to costumes (106; 107). It is worth highlighting that Pamela Howard, a famous theatre director, and scenographer, highlighted how a costume may evolve away from the designer's initial sketch, noting that "brilliant the costume drawing, the fitting room is where the real creative costume work begins" (108). With these statements, we can open new roads to speed up the process of designing a costume by digitally generating one and letting it be worn by a model, from a single sketch. This pipeline, visually depicted in Figure 7, will involve the usage of a Sketch2Image DL model (109; 110), along with a Virtual-

Try On (VTON) module (111).

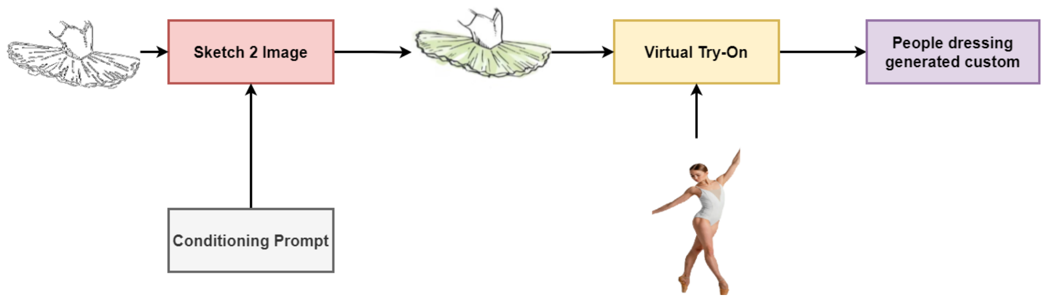


Figure 7: Generative DL pipeline: a combination of sketch2image and VTON models, allows to let a model wear a hypothetical dance costume, to speed up the design phase.

Sketch-to-image refers to the generative task of generating realistic images from rough sketches or outlines. In this approach, a generative AI model is trained on a dataset containing pairs of sketches and corresponding real images. The model learns to understand the relationship between the sketches and the desired output images, enabling it to generate realistic images from new sketches that it hasn't seen during training (109; 110). Nowadays, such sketched conditioning is paired with textual description control through a special class of generative AI models: (Latent) Diffusion (112; 59). For example, (59) added spatial control to the Stable Diffusion text-to-image model (112), allowing to provide structural conditioning image (e.g., edge maps, human pose skeletons, segmentation maps, depth, normals, etc.) for the image generation process, named ControlNet. This model provides state-of-art performance for many conditioning tasks, making it the perfect candidate for our use case: generating dance costumes from sketch and prompt conditioning.

Then, our mentioned pipeline should provide a method to let this generated dress be worn by (pictures of) dancers. To this date, a VTON model should be adopted, (111). Born in the fashion domain, image-based VTON exploits 2D people images and dresses ones to warp, and projects the latter to fit the body detailed in the former, creating photo-realistic results that maintain body pose and clothing deformation. For such tasks, different Generative AI models, like Generative Adversarial Networks (GANs) were proposed as possible solutions. However, recently, Diffusion models emerged as the best option given their generative performance (113). Among the set of diffusion models developed to solve the VTON task, a recent alternative, named OODiffusion, provided an open-source and state-of-the-art tool for the VTON generation task (114). This tool became so part of our framework.

2D to 3D synthesis. Transitioning from 2D images to 3D models involves various methods, including photogrammetry, and DL-based ones (115; 116). Considering photogrammetry, structure-from-motion (SfM) allows to reconstruct 3D geometry from a series of 2D images by analyzing their relative positions and orientations. The process involves analyzing the geometric properties of images taken from different viewpoints to reconstruct the three-dimensional structure of objects or scenes. In particular, SfM methods first detect and match key points between pairs of images. Then, camera poses are estimated and 3D coordinates are triangulated by intersecting the rays back-projected from the cameras (117; 116). Finally, they provide a dense 3D model with texture mapping. SfM is advantageous for its simplicity and ability to handle large datasets, however, it struggles with accuracy in complex scenes and requires considerable computational resources. Furthermore, multi-view stereo (MVS) methods have been developed. Different from SfM, this class of methods estimates depth information by analyzing multiple images of the same scene. MVS offers improved accuracy compared to SfM, particularly in detailed environments, but it can be computationally intensive and sensitive to image quality and viewpoint variations (117; 116). However, classical methods are sensitive to factors like lighting conditions, camera calibration, and image quality, which may affect the accuracy of the resulting models.

For this reason, novel paradigms based on DL, like Neural Rendering (NR) are rapidly gaining popularity (118). In such a domain, Neural Radiance Fields (NeRF) use a neural network to learn implicitly a continuous volumetric representation from 2D images taken from a sparse set of camera poses, achieving photorealistic reconstructions (119; 118). In the realm of NR, 3D Gaussian splatting (3DGS) corresponds to an explicit radiance field technique. 3DGS provides a representation based on differentiable 3D Gaussian-shaped primitives for modeling 3D structures (120). These primitives are used to parameterize the radiance field and can be rendered to produce novel views. In contrast to NeRF, which relies on computationally expensive volumetric ray sampling, 3DGS achieves real-time rendering through a tile-based rasterizer. As a downside, it is more challenging to extract a mesh from 3DGS concerning NeRFs. However, for both paradigms, this is still an open challenge (121; 118). However, recent 3DGS paradigms like (64) exhibited the ability of 3D reconstruction from a single or few images. Differently from NR, different DL-based methods were designed to directly reconstruct the underlining 3D geometry and texture of certain classes of objects, like PIFu (122).

PIFu, which stands for pixel-aligned implicit function, employs a neural network architecture to predict a voxelized representation of the 3D object directly from a set of 2D input images, depicting a clothed human. This method generates highly detailed and accurate 3D reconstructions even with one single image, capturing fine-grained surface details through implicit functions (122).

Considering the large amount of dance heritage that is nowadays stored in few images in archives, often taken in an “in the wild” setting (123), we integrated in DanXe, DL-based paradigms. Specifically, we adopted 3DGS and PIFu paradigms to generate customers and dancers, considering their ability to well approximate 3D reconstructions with a single or few images. However, it’s worth noting that DanXe could straightforwardly integrate all the cited methods for 3D reconstruction, depending on the specific use case. Based on the latter, it would be necessary to establish a protocol for acquiring the necessary data sources (124).

4.2.3. Sound Analysis

DL also exposes SOTA performances for analyzing historical recordings stored as audio data. These techniques offer a framework for delving into the complexities of dance heritage, in understanding how music influences dance performances and how to analyze spoken narratives and interviews. To this date, we here introduce the audio pipeline depicted in Figure 8 for musical and interview audio analysis.

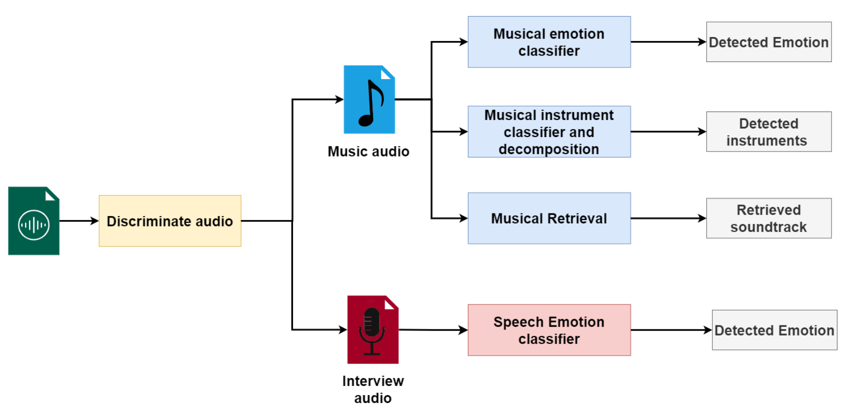


Figure 8: DanXe architecture for Musical and Interviews. It first adopts a discriminator to classify musicals from interviews and audio and then provides different kinds of analysis for both of them.

Music Analysis. Music Information Retrieval (125) is a field of study that focuses on the extraction of useful information from music data. It involves the development and application of computational techniques to analyze, organize, search, and understand music content. Related to dance, methods such as Emotion Recognition in Music (61) can help researchers analyze the emotional content of music to understand how it might influence the mood and expression of a dance performance. Additionally, Musical Instrument Recognition techniques (62) add another dimension to this analysis, by identifying the instruments being used in a piece of music, which can help choreographers match movements to specific sounds or rhythms, and thus better adhere to the original score (both choreographically and musically).

Audio Signal Analysis. To analyze audio sources related to dance, we need to refer to the general domain of audio signal analysis. Transcribing music and synchronizing audio with dance movements can enhance different aspects of dance material analysis, not only related to music analysis. In interactive dance performances or installations, audio synchronization applications can enable real-time interactions between dancers and sound. Moreover, music transcription tools can be valuable resources for dance educators and dance history and preservation efforts: music transcription and audio synchronization applications can help recreate and document historical dance performances. Aubio[2] (63) is widely used in music transcription and audio synchronization applications. Methods for audio analysis facilitate the integration with choreographic processes through techniques such as musical instrument recognition, music genre classification, and audio transcription and synchronization.

Audio-to-Text Conversion and Sentiment Audio Analysis. Mostly related to our purposes of preserving and understanding dance cultural heritage through audio analysis, recent sophisticated methods such as Audio-to-Text Conversion (126) can transcribe spoken narratives, interviews, and oral histories into written form. This facilitates the preservation and digitization of auditory records, including discussions on dance traditions, techniques, and personal experiences.

Once transcribed, the textual data undergoes computational examination, employing techniques like sentiment analysis and thematic categorization. In particular, the possibility to transcribe or recognize emotions of spoken narratives or interviews allows us to better understand the public images that artists have transmitted to the general public, an essential point to objectively analyze their media impact. To carry out such analysis we could resort to a state-of-the-art model,

[2] `https://github.com/aubio/aubio`

SpeechBrain (127) trained on the IEMOCAP training data[3] (128). This model proposes a method for speech emotion recognition where features extracted from pre-trained wav2vec models are modeled using simple neural networks. By using this tool in our case study (see Figure 8), we analyzed the resulting emotion of Dancer's interviews, available as audio speech. The results allow for a deeper comprehension of the public images that the artist transmitted to the general audience.

4.3. Digital Twins and Immersive Manipulations

DanXe first aims at digitizing tangible and intangible material to capture it in different kinds of data structures, then analyzing those through an intelligent layer, composed of several AI-based algorithms. However, the outcomes of such analyses would be hard to organize and visualize in an effective way (e.g., to discover data correlation). This phenomenon is emphasized in the dance domain, considering the multi-modal and temporal nature of its sources (12).

A solution to gather, maintain, and visualize such kind of complex information, amounts to the usage of DT and their immersive visualization employing XR paradigms (129; 130; 131; 16; 132; 11). As highlighted by (133; 16) it is possible to define a five-dimensional DT model that is adaptable for different domains, data, and contexts. In our case, we want to put Dance at its center, defining a flow that comprises not only a robust DT model but also smart immersive interfaces to augment the experience of accessing and manipulating such data, extending the one introduced in (16). The main motivations to implement such immersive and modular interfaces rely on their capability of easing the visualization of complex data while improving human data interaction and manipulation capabilities (130; 16; 131; 11). The architecture designed for this aim is visually depicted in Figure 9.

The first component amounts to an extension of the five-dimensional DT model introduced in (133) which is composed of the following components: (i) a Physical Environment (PE) that includes a series of entities, such as humans, machines, and documents; (ii) a Virtual Environment (VE) consisting of models built in multiple dimensions, including geometry, physics, behavior, and rules, evolving according to the PE, and the (iii) Activity Service System (ASSYST). ASSYST amounts to an integrated service module, which encapsulates the functions of data analytics, models, algorithms, etc., into sub-services, and combines them to form

[3]`https://huggingface.co/speechbrain/emotion-recognition-wav2vec2-IEMOCAP`

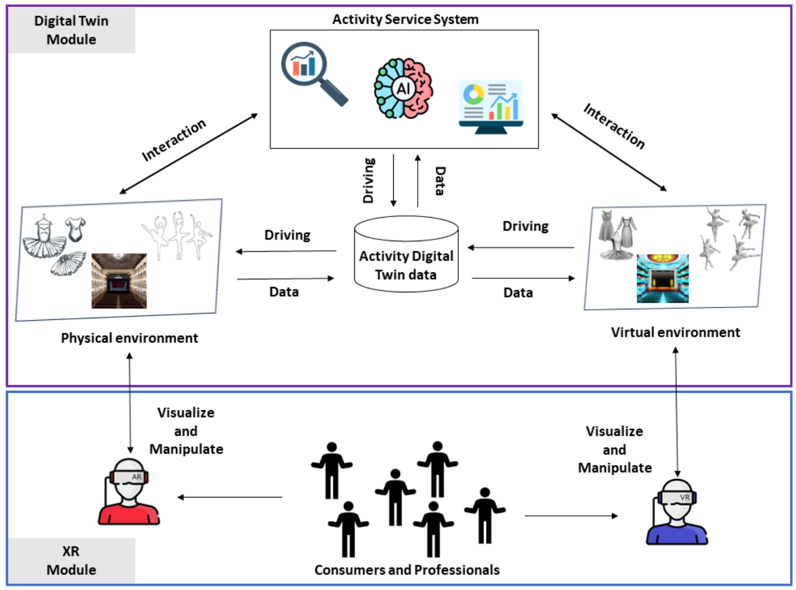


Figure 9: DanXe XR-DT architecture: a five-dimensional DT model with an XR plug-in module was adapted from (133; 16) to twin Dance Heritage.

composite services for specific demands from the PE and the VE. Finally, (d) the Activity Digital Twin Data (ADTD) includes the PE, VE, and ASSYST data, their aggregations, and the existing modeling methods, optimizing and empowering both the PE and VE. The data in ADTD communicates in real-time with all other modules eliminating possible isolated information to provide a synchronized and consistent vision. The second component corresponds instead to an XR plug-in used to visualize and manipulate information related to the DT itself (which is here defined as the DT of the dance world). In practice, each user that interacts with the PE or the VE can exploit AR, MR, or VR devices to fetch, visualize, and manipulate the information flowing from the physical realm to the virtual one and vice versa. The actions taken in AR, MR, and VR are processed by the ADTD and the ASSYST modules that will manage the manual update of the DT made by humans with an in-the-loop approach (16).

Such an architecture will implement the core activity of DanXe: a bridge between reality and virtuality, which exploits digitization, intelligent analysis, and smart visualizations for dance heritage. In practice, the Digitization module introduced in Section 4.1 will project the physical environment into the virtual one while the analysis one, introduced in Section 4.2 will cover the role of the ASSYST module. In the following, we will discuss how to implement the XR immersive module to access multimedia, multimodal, and also immersive information within each degree of XR.

Multi-modal immersive XR visualizations. Considering a single Dance "centroid" (e.g., a dancer, a dance style, a step) it is possible to correlate, link, and connect multi-domain and multi-modal digital knowledge to it. In such a context, immersive visualizations play the key role of providing ergonomic, modular, and accessible interfaces to that multi-modal graphs (134; 135; 136; 137). However, a technological pipeline should be carefully composed to combine three different factors: (i) exploit validated UI elements and layouts; (ii) seamlessly integrate any kind of multi-media data; (iii) render those graphical content in any possible degree of reality; (iv) provide the chance to collaborate. Nowadays, we can holistically obtain all those features, considering the immersive web and, in particular, the WebXR framework (138; 139). Web technologies along with WebXR APIs and web frameworks providing interfaces to them (like Tree.js [4]) provide the chance to develop an all-in-one platform, compatible with all XR devices implementing those interfaces. WebXR can modularly discriminate different supported XR devices, allowing them to steer the flow of the simulation and furnish the interface that corresponds to the target device. At the same time, web networking frameworks, like Socket.io [5], and Photon [6] could exploited to define a distributed architecture to let users being immersed within a collaborative and real-time multiplayer immersive experience, to visualize and manipulate such digital information, like the open-source frameworks introduced in (138; 139). To provide the first version of such a system, we here introduced a module developed in Unity, which consists of an immersive environment that provides multi-modal connections among the "centroid" items and all their linked information. In practice, this module represents an instance of the process reported in Figure 9.

5. DanXe applied to Rudolf Nureyev Heritage

Studying Nureyev's heritage entails several criticalities, including (i) Lack of scientific studies and Absence of theoretical material produced by Rudolf Nureyev; (ii) Massive mediatization; (iii) Geographical dispersion of sources; (iv) Peculiarities of choreographic language; (v) Fragility of media supports; (vi) Delay in the

[4] `https://threejs.org/`
[5] `https://socket.io/`
[6] `https://www.photonengine.com/`

process of digitizing documents.

Moreover, the polarization of sources around the two previously identified cores (memory as trace, embodied memory) is particularly relevant and cannot be approached by considering only one of the two dimensions. The artist and the myth, the works, and the impact are poles that feed into and influence each other and, necessarily, must be investigated from a comparative perspective.

Challenging and complex, the 'Nureyev case' nonetheless offers significant opportunities for access from the standpoint of digital technologies, aligning closely with all the identified peculiarities in the preservation of performing arts and, more broadly, Intangible Cultural Heritage. Applied to the case study, DanXe has proven to be a valuable tool for addressing criticalities and approaching the entire range of documented resources, depicted in Figure 10. In the following, we will describe how DanXe was applied to all the here listed use cases, exploiting a subset of its functionalities.

Figure 10: DanXe's coverage of Dance Cultural Heritage.

	OCR	TOPIC MODELING	VIDEO PROCESSING	SKETCH TO COSTUME	2D-3D	AUDIO PROCESSING
PAPER BASED DOCUMENTS	✓	✓		✓	✓	
AUDIOVISUALS				✓	✓	
TEXTILES AND ACCESSORISES					✓	
EMBODIED KNOWLEDGE			✓			✓

Text Extraction with OCR. Let's take the case of paper documents, which are among the most numerous resources in dance heritage studies. Through analysis techniques such as OCR (60), it's possible to immediately process textual content, transcribe it, conduct macro and micro analyses, and synthetically visualize the derived results, connecting and putting them into communication. This automatic process, included in DanXe, was applied to the case of a billboard, as visually depicted in Figure 11.

From such information, it's possible to extract the names of individual performers, authors, and staff, gather information about performance venues, cast

(a) Words bounding box detection.

(b) Words description with selection by returned classification confidence.

Figure 11: Sentence recognition on a section of a Theatre Poster from "Teatro Comunale di Bologna". ©Archivio Storico/Museo Internazionale e Biblioteca della Musica di Bologna.

lists for specific dates, relevant seasons, and ticket costs. These operations, usually entrusted to manual transcription, are of primary importance for contributing to the improvement of meta-descriptions of archival assets, essential for making the preserved heritage concretely accessible. Although prevalent in the use of musical and operatic materials (140), there's still a lack of consistent literature applied to dance documents, often influenced by the skills and specificity of the compiler to whom they are entrusted.

Topic extraction. Still regarding the treatment of textual documents, particularly those already digitized, the use of resources like those offered by topic modeling objectively highlights research trends and areas of interest, allowing for a comprehensive verification of the progress of studies on a specific topic. In the case of Nureyev, this type of research can be essential for verifying how his image was communicated by the press and thus analyzing punctually the communicative strategies employed to convey his public persona. For instance, the Topic Map of Figure 12 highlights the centrality of the Cold War in the dissemination of associated content (extracted with BERTopic (53)). In addition to "Topic 9," naturally, the many others that, to a greater or lesser extent, are highlighted on the

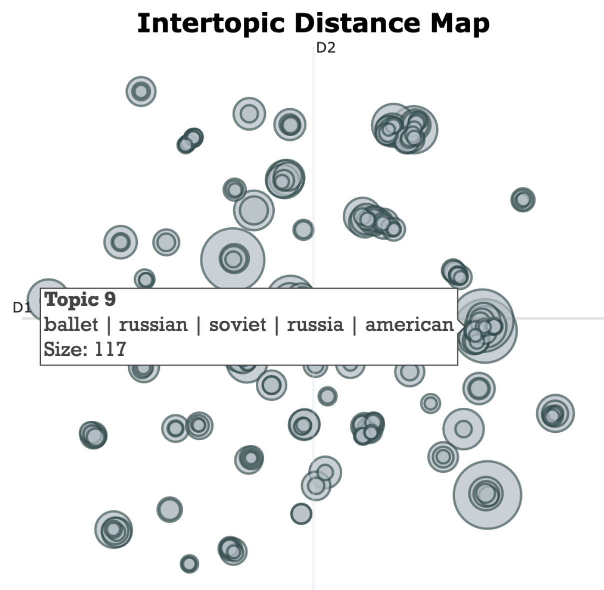


Figure 12: Topic modeling with BERTopic, among the different topics that emerged from the considered documents, we highlighted one of the main discussed trends behind Rudolf Nureyev.

map should be added. These can emerge and take shape differently depending on the input content, allowing the scholar to have a tool to verify the trends of the documents they are referring to in their analysis.

Sentiment Audio analysis. In terms of media impact analysis, particular interest has been shown in the use of audio processing to measure speech emotion. The graph in Figure 13 shows the result of the analysis process of an interview recorded by the British broadcaster Thames Television on June 17, 1961, within the program “Afternoon Plus”.

In this interview, among the most viewed online according to YouTube trends, Nureyev exhibits a relaxed and calm tone of voice. The dancer is indeed extremely comfortable in front of the camera, a fact that, when compared with the results emerging from analyses of similar television moments, can provide scientifically valid and objective data for the study of celebrity and its mechanisms. The technology applied to television broadcasting can be extended to all audio contributions and thus provide possibilities for the study and analysis of documents of primary importance, such as oral testimonies, usually treated for transcription

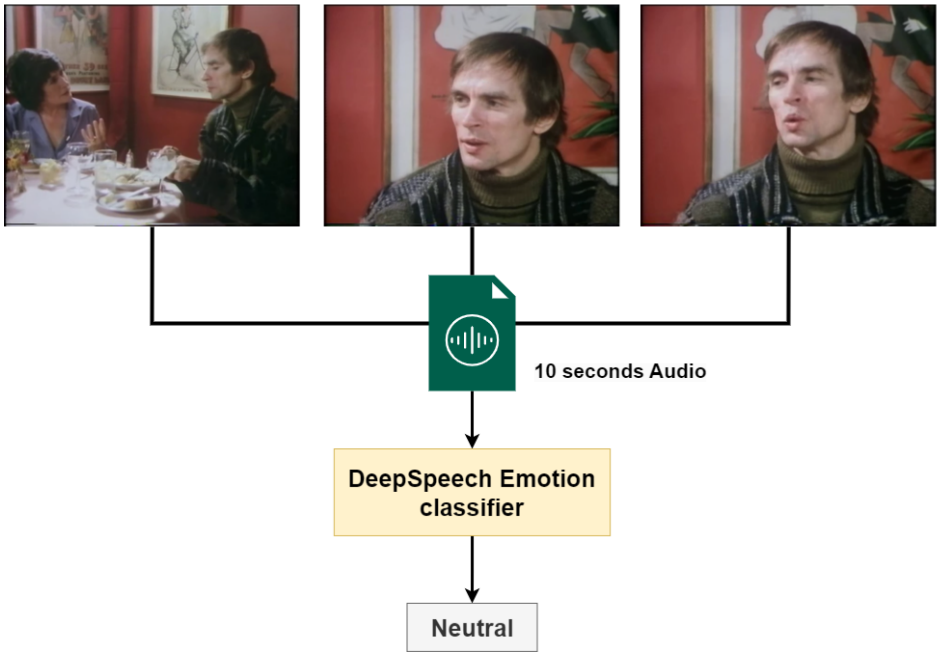


Figure 13: Speech emotion performed on Rudolf Nureyev's interview (1981) ©YouTube/ThamesTv.

purposes.

3D reconstruction. In the perspective of covering the entire range of documentary assets associated with the dancer, DanXe provides important tools for the treatment of stage costumes. Nureyev's interest in clothing (both theatrical and non-theatrical) is well known (141). His distinguishable and iconic costumes were expressions of his taste and physique, capable of recounting peculiarities and aesthetics. Often preserved in theater archives, stage costumes are largely communicated to us through performance photos and sketches, 2D formats to which, as highlighted in Figure 14, it's possible to restore three-dimensionality, preserving the assets from degradation while providing a rendering that allows users to appreciate their volumes (3d view in 14) by adopting a Gaussian Splatting based-method included in DanXe, named Triplane Gaussian (64).

Sketch to Image. For sketches instead, we picked the mentioned ControlNet fine-tuned with edge map images calculated from the canny filter (59). We then adopted such a model to generate proposals from a gray-scaled sketch using the prompt "White and gold stripes". From such proposals, we picked one and then

Figure 14: Doublet for the role of Basilio in *Don Quixote*, 1979. Costume by Nicholas Georgiadis. © CNCS / Photo Pascal Francois.

passed such generated costume along with the image of a dancer to a Virtual Try-on Module (OOTDiffusion) to generate the image of such a dancer who wearing the synthetically generated costume. This process is visually depicted in Figure 15.

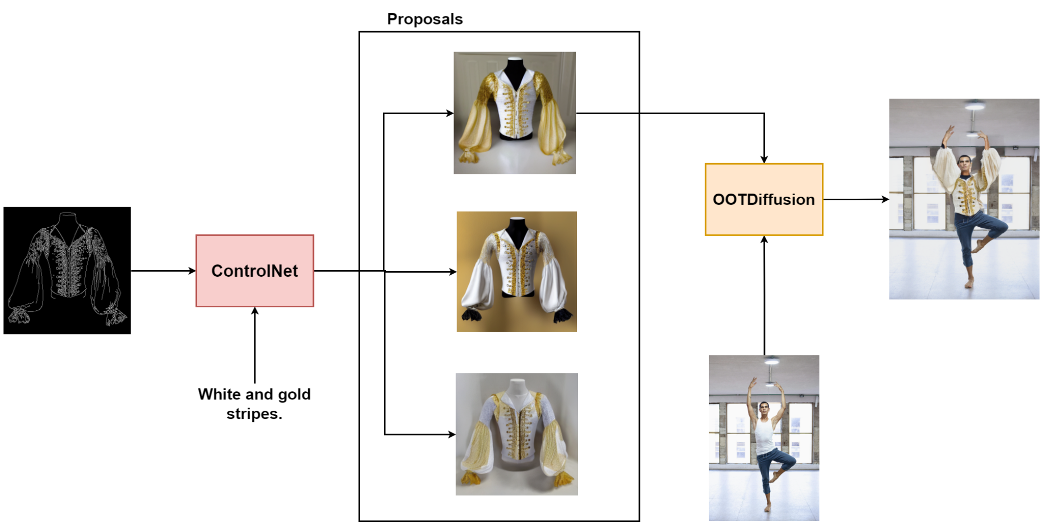


Figure 15: ControlNet applied with OOTDiffusion.

Keypoint regression and Masking. As highlighted at the outset, the cultural heritage of dance, to be fully restored, cannot disregard moving bodies, stories, and knowledge they can convey. In the case of a highly mediated dancer like Nureyev, for example, there are numerous audiovisual materials that, together with witnesses, play an important role in preserving original choreographies. If analyzed through systems like key-point detection, these resources can offer new perspectives in dance studies, providing direct access to the choreographic language, enhancing comprehension and analysis of the repertoire, preserving the authenticity of works, and improving accessibility to a cultural heritage that remains tied to specialized users. Figure 16 shows Nureyev in some frames from *Swan Lake* (1966).

Figure 16: Combo AlphaPose with segment anything furnishing key points as geometrical features query. (Rudolf Nureyev as Siegfried in *Swan Lake*, Act I, 1966 ©YouTube/JohnHall.

Once the positions of points on the body are identified and fixed patterns associated with steps or specific poses are memorized, it's possible to analyze the choreography, transcribe it, and verify recurring features. We also exploited those key points to mask out only those pixels related to Rudolf Nureyev. In essence, ensuring its safeguarding for future generations of scholars and users, including dancers and their bodies in the processes. Applied to the Nureyev case, DanXe provides evidence of how necessary a comparative and multimodal approach is to the reading and storytelling of an artistic heritage and how dance research can provide suitable use cases for multidisciplinary projects involving resources and implications.

Immersive visualizations. Considering our final goal of defining the Digital Twin of the Dance world, we here apply the architecture detailed in Section 4.3 to visualize the analysis and the outcomes performed with our AI-based intelligent layer. In particular, we here implement an instance of it by rendering a 3D that contains: a central 3D model of Rudolf Nureyev (low quality) automatically generated by the PiFU model, and various multi-modal documents linked to the scene it represents: the original masked picture, the theater billboard and the topics emerged analyzing documents connected to Nureyev. Such a 3D environment, depicted in Figure 17, was prototyped with the Unity Game Engine [7], preceding its deployment in the WebXR framework cited in Section 4.3. All the 3D rendered elements

[7] https://unity.com/

are connected by a straight line from the center of the 3D Nureyev's model to all the cited elements.

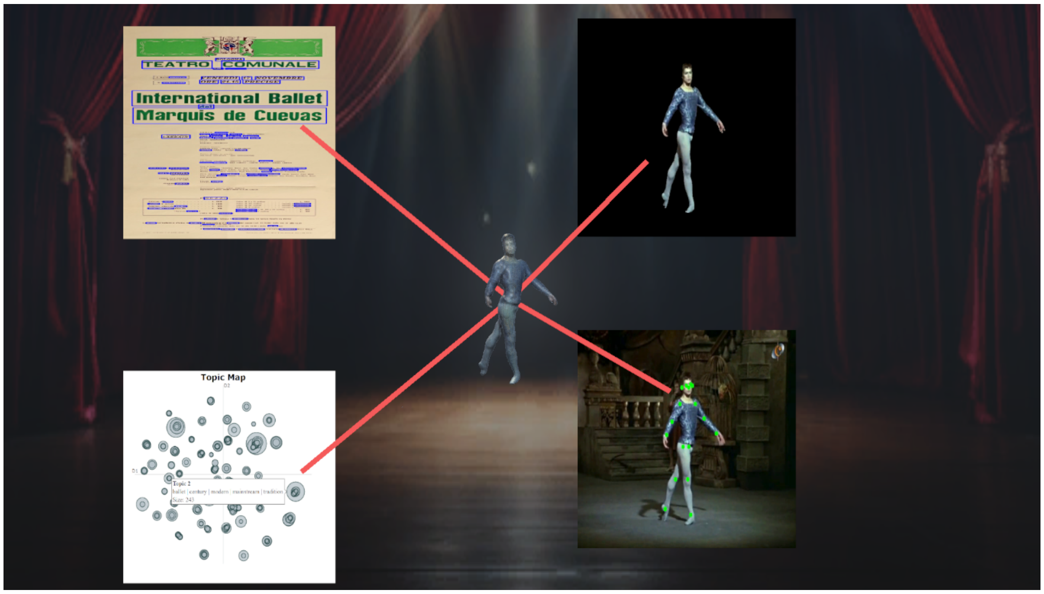


Figure 17: Multi-modal visualizations mockup made in Unity.

6. Discussions and Conclusions

In 2003, the UNESCO General Conference signed the *Convention for the Safeguarding of Intangible Cultural Heritage*, aimed to recognize, promote, and preserve intangible cultural heritage, encompassing traditions and practices passed down through generations. Among the motivations prompting Member States to ratify it was the lack of adequate tools for safeguarding this valuable and special type of heritage.

Dance holds a prominent position within this *Convention*, reflecting its importance as a form of intangible cultural heritage. Coupled with this aspect, especially in the case of theatrical dance, is a strong tangible dimension due to the presence of a centuries-old documentary heritage consisting of writings, theories, and documentation of performances. This dual dimension, comprising both tangible and intangible aspects, characterizes the dance heritage, necessitating a multi-modal approach to its study, storytelling, and preservation. Despite a tradition of significant exchange and communication between dance and information technologies in the field of creation, there's a lack of approaches in terms of safeguarding and promotion. However, advancements in AI and XR, offer new opportunities for the digitization, analysis, and visualization of dance heritage, leading to the creation of a hypothetical DT of the dance world. To this date, we have introduced a multidisciplinary framework called DanXe, which aims to preserve and analyze the entire spectrum of documentary assets associated with dance heritage.

DanXe combines, in particular, AI strategies for digitization and analysis with XR solutions for immersive visualization, paving the way for the definition of an Extended Artificial Intelligence tool. This framework not only enriches the storytelling of dance traditions but also enhances accessibility and provides valuable insights through automated data analysis. To demonstrate its multi-modal flexibility, we took as a case study the artistic legacy of Rudolf Nureyev. Overall, the multidisciplinary approach presented offers a comprehensive solution for safeguarding and promoting intangible cultural heritage, with dance being a central focus.

However, DanXe presents some limits: notably, it lacks a deployment mechanism for real-time data collection and analysis, which could hinder its effectiveness in capturing dynamic aspects of dance performances. Additionally, while the framework is essential in digitization and analysis, its integration with traditional preservation methods may require further development. Also, the accessibility of DanXe to smaller cultural institutions or independent researchers might be constrained by technical expertise or resource requirements. Furthermore, the ethical considerations surrounding data privacy and ownership in digitization efforts need careful attention (4). Clear protocols for consent, intellectual property rights, and data sharing must be established to support the integrity of the communities whose heritage is being digitized.

It is worth highlighting, that the framework's implications go beyond the dance domain, impacting various creative industries and cultural heritage preservation efforts. The knowledge and methodologies developed through DanXe can be repurposed and adapted across various domains, including the fashion industry, archival science, philology, museography, and other creative industries. By leveraging AI and XR technologies to digitize and analyze dance heritage, DanXe opens avenues for enhancing accessibility, fostering interdisciplinary collaborations, and enriching scholarly research, whether it's applied to dance, fashion design, historical documents, or curating immersive museum exhibitions. Future works will explore methods and technologies to overcome DanXe limitations while at the same time integrating multi-modal approaches in its analysis and visualization components.

Acknowledgements

This work was partly funded by: (i) the PNRR - M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - ”FAIR - Future Artificial Intelligence Research” - Spoke 8 ”Pervasive AI”, funded by the European Commission under the NextGeneration EU program; (ii) MICS (Made in Italy – Circular and Sustainable) Extended Partnership and received funding from the European Union Next-Generation EU (PNRR – M4C2, Investimento 1.3 - D.D. 1551.11-10-2022, Partenariato Esteso PE00000004); (iii) a Ph.D. scholarship funded by the “PON Ricerca e Innovazione 2014-2020” project.

References

[1] Convention concerning the protection of the world cultural and natural heritage, https://unesdoc.unesco.org/ark:/48223/pf0000002091 (2024).

[2] Convention for the safeguarding of the intangible cultural heritage, https://ich.unesco.org/en/convention (2003).

[3] Browse the lists of intangible cultural heritage and the register of good safeguarding practices, https://ich.unesco.org/en/lists? (2024).

[4] S. Pansoni, S. Tiribelli, M. Paolanti, E. Frontoni, B. Giovanola, Design of an ethical framework for artificial intelligence in cultural heritage, in: 2023 IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS), IEEE, 2023, pp. 1–5.

[5] Operational directives for the implementation of the convention for the safeguarding of the intangible heritage, https://ich.unesco.org/en/directives (2024).

[6] M. Casari, et al., Antropologia e teatro. rivista di studi, ANTROPOLOGIA E TEATRO (2022).

[7] E. Cervellati, et al., La danza in scena. Storia di un’arte dal Medioevo a oggi, Bruno Mondadori, 2009.

[8] L. Stacchio, A. Angeli, G. Lisanti, G. Marfia, Applying deep learning approaches to mixed quantitative-qualitative analyses, in: Proceedings of the 2022 ACM Conference on Information Technology for Social Good, 2022, pp. 161–166.

[9] W. Alma'aitah, A. Z. Talib, M. A. Osman, Opportunities and challenges in enhancing access to metadata of cultural heritage collections: a survey, Artificial Intelligence Review 53 (5) (2020) 3621–3646.

[10] C. Wienrich, M. E. Latoschik, extended artificial intelligence: New prospects of human-ai interaction research, Frontiers in Virtual Reality 2 (2021) 686783.

[11] T. Hirzle, F. Müller, F. Draxler, M. Schmitz, P. Knierim, K. Hornbæk, When xr and ai meet-a scoping review on extended reality and artificial intelligence, in: Proceedings of the 2023 CHI Conference on Human Factors in ComputingmSystems, 2023, pp. 1–45.

[12] K. El Raheb, Y. Ioannidis, Dance in the world of data and objects, in: International Conference on Information Technologies for Performing Arts, Media Access, and Entertainment, Springer, 2013, pp. 192–204.

[13] S. Jürgens, Dance— data— storytelling, Dance data, cognition, and multimodal communication (2022) 38–54.

[14] K. El Raheb, N. Papapetrou, V. Katifori, Y. Ioannidis, Balonse: Ballet ontology for annotating and searching video performances, in: Proceedings of the 3rd International Symposium on Movement and Computing, 2016, pp. 1–8.

[15] B. R. Barricelli, E. Casiraghi, D. Fogli, A survey on digital twin: Definitions, characteristics, applications, and design implications, IEEE access 7 (2019) 167653–167671.

[16] L. Stacchio, A. Angeli, G. Marfia, Empowering digital twins with extended reality collaborations, Virtual Reality & Intelligent Hardware 4 (6) (2022) 487–505.

[17] I. Rallis, A. Voulodimos, N. Bakalos, E. Protopapadakis, N. Doulamis, A. Doulamis, Machine learning for intangible cultural heritage: a review of techniques on dance analysis, Visual Computing for Cultural Heritage (2020) 103–119.

[18] J. Adshead-Lansdale, J. Layson, Dance history: An introduction, Routledge, 2006.

[19] M. De Marinis, Il corpo dello spettatore. performance studies e nuova teatrologia, Sezione di Lettere (2014) 188–201.

[20] V. Gravano, E. Pitozzi, A. Sacchi, et al., B. Motion. Spazio di riflessione fuori e dentro le arti performative, Costa & Nolan, 2008.

[21] E. Randi, Primi appunti per un progetto di edizione critica coreica, SigMa-Rivista di Letterature comparate, Teatro e Arti dello spettacolo 4 (2020) 755–771.

[22] E. Giannasca, Dance in the ontological perspective of a document theory of art, Danza e ricerca. laboratorio di studi, scritture, visioni 10 (10) (2018) 325–346.

[23] S. Franco, Corpo-archivio: mappatura di una nozione tra incorporazione e pratica coreografica (2019).

[24] C. Barnes, Nureyev, New York, NY: Helene Obolensky Enterprises, 1982.

[25] F. S. Saunders, Who paid the piper?: the CIA and the cultural Cold War, Granta Books, 2000.

[26] E. Jacobs, When Merce Cunningham Took On a New Collaborator: The Computer (Published 2020) — nytimes.com, `https://www.nytimes.com/2020/09/01/arts/dance/merce-cunningham-computer-LifeForms.html` (2020).

[27] M. Cunningham, F. Starr, Changes: Notes on choreography, (No Title) (1968).

[28] J. Bradley, Learning the grammar of dance, in: In Proceedings of the International Conference on Machine Learning (ICML, page 547555. Morgan Kaufmann, 1998.

[29] V. Singh, C. Latulipe, E. Carroll, D. Lottridge, The choreographer's notebook: a video annotation system for dancers and choreographers, in: Proceedings of the 8th ACM Conference on Creativity and Cognition, 2011, pp. 197–206.

[30] A. Aristidou, A. Shamir, Y. Chrysanthou, Digital dance ethnography: Organizing large dance collections, Journal on Computing and Cultural Heritage (JOCCH) 12 (4) (2019) 1–27.

[31] R. E. Cisneros, K. Stamp, S. Whatley, K. Wood, Wholodance: digital tools and the dance learning environment, in: Dance, Professional Practice, and the Workplace, Routledge, 2020, pp. 88–108.

[32] S. Senecal, N. A. Nijdam, A. Aristidou, N. Magnenat-Thalmann, Salsa dance learning evaluation and motion analysis in gamified virtual reality environment, Multimedia Tools and Applications 79 (2020) 24621–24643.

[33] T. Shi, Application of vr image recognition and digital twins in artistic gymnastics courses, Journal of Intelligent & Fuzzy Systems 40 (4) (2021) 7371–7382.

[34] X. Li, The art of dance from the perspective of artificial intelligence, in: Journal of Physics: Conference Series, Vol. 1852, IOP Publishing, 2021, p. 042011.

[35] Q. Zhou, J. Wang, P. Wu, Y. Qi, Application development of dance pose recognition based on embedded artificial intelligence equipment, in: Journal of Physics: Conference Series, Vol. 1757, IOP Publishing, 2021, p. 012011.

[36] W. Xu, Q.-W. Xing, J.-D. Zhu, X. Liu, P.-N. Jin, Effectiveness of an extended-reality interactive learning system in a dance training course, Education and Information Technologies (2023) 1–31.

[37] M. Joshi, S. Chakrabarty, An extensive review of computational dance automation techniques and applications, Proceedings of the Royal Society A 477 (2251) (2021) 20210071.

[38] M. Reshma, B. Kannan, V. J. Raj, S. Shailesh, Cultural heritage preservation through dance digitization: A review, Digital Applications in Archaeology and Cultural Heritage (2023) e00257.

[39] A. Ramtohul, K. K. Khedo, Augmented reality systems in the cultural heritage domains: A systematic review, Digital Applications in Archaeology and Cultural Heritage (2024) e00317.

[40] N. Magnenat-Thalmann, D. Protopsaltou, E. Kavakli, Learning how to dance using a web 3d platform, in: Advances in Web Based Learning–ICWL 2007: 6th International Conference Edinburgh, UK, August 15-17, 2007 Revised Papers 6, Springer, 2008, pp. 1–12.

[41] I. Kico, N. Grammalidis, Y. Christidis, F. Liarokapis, Digitization and visualization of folk dances in cultural heritage: a review, Inventions 3 (4) (2018) 72.

[42] D. Chapinal-Heras, C. Díaz-Sánchez, A review of ai applications in human sciences research, Digital Applications in Archaeology and Cultural Heritage (2023) e00288.

[43] K. Kingsland, Comparative analysis of digital photogrammetry software for cultural heritage, Digital Applications in Archaeology and Cultural Heritage 18 (2020) e00157.

[44] E. Nofal, M. Reffat, A. Vande Moere, Phygital heritage: An approach for heritage communication, in: Immersive learning research network conference, Verlag der Technischen Universität Graz; Graz, Austria, 2017, pp. 220–229.

[45] P. Jin, Y. Liu, Fluid space: Digitisation of cultural heritage and its media dissemination, Telematics and Informatics Reports 8 (2022) 100022.

[46] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, J. Gain, A survey of augmented, virtual, and mixed reality for cultural heritage, Journal on Computing and Cultural Heritage (JOCCH) 11 (2) (2018) 1–36.

[47] E. Balloni, L. Gorgoglione, M. Paolanti, A. Mancini, R. Pierdicca, Few shot photogrametry: A comparison between nerf and mvs-sfm for the documentation of cultural heritage, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 48 (2023) 155–162.

[48] C. Fernandes, S. P. Coelho, A. B. Vieira, Dance and the (digital) archive: A survey of the field, Dance research 38 (2) (2020) 271–288.

[49] S. Karumuri, R. Niewiadomski, G. Volpe, A. Camurri, From motions to emotions: classification of affect from dance movements using deep learning, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–6.

[50] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, L. He, A survey of human-in-the-loop for machine learning, Future Generation Computer Systems 135 (2022) 364–381.

[51] P. Xu, X. Zhu, D. A. Clifton, Multimodal learning with transformers: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence (2023).

[52] P.-L. H. Cabot, R. Navigli, Rebel: Relation extraction by end-to-end language generation, in: Findings of the Association for Computational Linguistics: EMNLP 2021, 2021, pp. 2370–2381.

[53] M. Grootendorst, Bertopic: Neural topic modeling with a class-based tf-idf procedure, arXiv preprint arXiv:2203.05794 (2022).

[54] H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale, et al., Llama 2: Open foundation and fine-tuned chat models, arXiv preprint arXiv:2307.09288 (2023).

[55] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li, C. Lu, Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time, IEEE Transactions on Pattern Analysis and Machine Intelligence (2022).

[56] P. Schneider, R. Memmesheimer, I. Kramer, D. Paulus, Gesture recognition in rgb videos using human body keypoints and dynamic time warping, in: RoboCup 2019: Robot World Cup XXIII 23, Springer, 2019, pp. 281–293.

[57] K. Zhou, Y. Yang, A. Cavallaro, T. Xiang, Learning generalisable omni-scale representations for person re-identification, TPAMI (2021).

[58] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, et al., Segment anything, arXiv preprint arXiv:2304.02643 (2023).

[59] L. Zhang, A. Rao, M. Agrawala, Adding conditional control to text-to-image diffusion models, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 3836–3847.

[60] JaidedAI, EasyOCR: Ready-to-use OCR with 80+ supported languages and all popular writing scripts including Latin, Chinese, Arabic, Devanagari, Cyrillic and etc., `https://github.com/JaidedAI/EasyOCR` (Accessed 2024).

[61] R. Panda, R. Malheiro, R. P. Paiva, Novel audio features for music emotion recognition, IEEE Transactions on Affective Computing 11 (4) (2018) 614–626.

[62] A. Lerch, An introduction to audio content analysis: Music Information Retrieval tasks and applications, John Wiley & Sons, 2022.

[63] T. Giannakopoulos, pyaudioanalysis: An open-source python library for audio signal analysis, PLoS ONE 10 (2015).

[64] Z.-X. Zou, Z. Yu, Y.-C. Guo, Y. Li, D. Liang, Y.-P. Cao, S.-H. Zhang, Triplane meets gaussian splatting: Fast and generalizable single-view 3d reconstruction with transformers, arXiv preprint arXiv:2312.09147 (2023).

[65] M. Nakamura, K. Hachimura, An xml representation of labanotation, labanxml, and its implementation on the notation editor labaneditor2, Review of the National Center for Digitization (Online Journal) 9 (2006) 47–51.

[66] J. Lansdale, Intertextual narratives in dance analysis, in: Decentring Dancing Texts: The Challenge of Interpreting Dances, Springer, 2008, pp. 1–20.

[67] D. W. Otter, J. R. Medina, J. K. Kalita, A survey of the usages of deep learning for natural language processing, IEEE transactions on neural networks and learning systems 32 (2) (2020) 604–624.

[68] A. Abdelrazek, Y. Eid, E. Gawish, W. Medhat, A. Hassan, Topic modeling algorithms and applications: A survey, Information Systems 112 (2023) 102131.

[69] L. Stacchio, A. Angeli, G. Lisanti, G. Marfia, Analyzing cultural relationships visual cues through deep learning models in a cross-dataset setting, Neural Computing and Applications (2023) 1–16.

[70] Y. Cai, F. Wang, H. Wang, Q. Qian, Public sentiment analysis and topic modeling regarding chatgpt in mental health on reddit: Negative sentiments increase over time, arXiv preprint arXiv:2311.15800 (2023).

[71] J. Wu, S. Yang, R. Zhan, Y. Yuan, D. F. Wong, L. S. Chao, A survey on llm-gernerated text detection: Necessity, methods, and future directions, arXiv preprint arXiv:2310.14724 (2023).

[72] J. Wang, J. X. Huang, X. Tu, J. Wang, A. J. Huang, M. T. R. Laskar, A. Bhuiyan, Utilizing bert for information retrieval: Survey, applications, resources, and challenges, ACM Computing Surveys (2024).

[73] W. Yang, H. Zhang, J. Lin, Simple applications of bert for ad hoc document retrieval, arXiv preprint arXiv:1903.10972 (2019).

[74] Z. A. Yilmaz, W. Yang, H. Zhang, J. Lin, Cross-domain modeling of sentence-level evidence for document retrieval, in: Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), 2019, pp. 3490–3496.

[75] A. Clarance, A proposal for the creation of a dance ontology, in: Arts and Technology: Fourth International Conference, ArtsIT 2014, Istanbul, Turkey, November 10-12, 2014, Revised Selected Papers 4, Springer, 2015, pp. 86–99.

[76] H. Jin, Y. Zhang, D. Meng, J. Wang, J. Tan, A comprehensive survey on process-oriented automatic text summarization with exploration of llm-based methods, arXiv preprint arXiv:2403.02901 (2024).

[77] Z. Nasar, S. W. Jaffry, M. K. Malik, Named entity recognition and relation extraction: State-of-the-art, ACM Computing Surveys (CSUR) 54 (1) (2021) 1–39.

[78] S. Ji, S. Pan, E. Cambria, P. Marttinen, S. Y. Philip, A survey on knowledge graphs: Representation, acquisition, and applications, IEEE transactions on neural networks and learning systems 33 (2) (2021) 494–514.

[79] C. Santini, M. A. Tan, T. Tietz, O. Bruns, E. Posthumus, H. Sack, Knowledge extraction for art history: the case of vasari's the lives of the artists (1568), in: Qurator 2022, Third Conference on Digital Curation Technologies, Sept. 19th-23rd, 2022, Berlin, Germany, Vol. 3234, Aachen, Germany: RWTH Aachen, 2022, p. 7.

[80] C. Santini, N. Garay, E. Posthumus, H. Sack, The art of relations, DHd 2024 Quo Vadis DH (DHd2024), Passau, Deutschland (febbraio 21 2024). URL https://doi.org/10.5281/zenodo.10698246

[81] M. Bahar, M. Hartati, et al., Jambi malay dance fashion aesthetics, Mudra Jurnal Seni Budaya 37 (4) (2022) 446–454.

[82] C. Li, C. Zhang, Cnn or vit? revisiting vision transformers through the lens of convolution, arXiv preprint arXiv:2309.05375 (2023).

[83] P. Dierickx, A. Van Damme, N. Dupuis, O. Delaby, Comparison between cnn, vit and cct for channel frequency response interpretation and application to g. fast, IEEE Access 11 (2023) 24039–24052.

[84] Z. Raisi, M. A. Naiel, P. Fieguth, S. Wardell, J. Zelek, Text detection and recognition in the wild: A review, arXiv preprint arXiv:2006.04305 (2020).

[85] C. Neudecker, K. Baierer, M. Gerber, C. Clausner, A. Antonacopoulos, S. Pletschacher, A survey of ocr evaluation tools and metrics, in: The 6th International Workshop on Historical Document Imaging and Processing, 2021, pp. 13–18.

[86] A. Angeli, L. Stacchio, L. Donatiello, A. Giacchè, G. Marfia, Making paper labels smart for augmented wine recognition, The Visual Computer (2023) 1–13.

[87] A. Singh, K. Bacchuwar, A. Bhasin, A survey of ocr applications, International Journal of Machine Learning and Computing 2 (3) (2012) 314.

[88] K. Smelyakov, A. Chupryna, D. Darahan, S. Midina, Effectiveness of modern text recognition solutions and tools for common data sources., in: COLINS, 2021, pp. 154–165.

[89] Y. Baek, B. Lee, D. Han, S. Yun, H. Lee, Character region awareness for text detection, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 9365–9374.

[90] B. Shi, X. Bai, C. Yao, An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition, IEEE transactions on pattern analysis and machine intelligence 39 (11) (2016) 2298–2304.

[91] A. Graves, S. Fernández, F. Gomez, J. Schmidhuber, Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks, in: Proceedings of the 23rd international conference on Machine learning, 2006, pp. 369–376.

[92] D. R. Beddiar, B. Nini, M. Sabokrou, A. Hadid, Vision-based human activity recognition: a survey, Multimedia Tools and Applications 79 (41-42) (2020) 30509–30555.

[93] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li, C. Lu, Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time, IEEE Transactions on Pattern Analysis and Machine Intelligence (2022).

[94] C. Zheng, W. Wu, C. Chen, T. Yang, S. Zhu, J. Shen, N. Kehtarnavaz, M. Shah, Deep learning-based human pose estimation: A survey, ACM Computing Surveys 56 (1) (2023) 1–37.

[95] Y. Xu, J. Zhang, Q. Zhang, D. Tao, Vitpose: Simple vision transformer baselines for human pose estimation, Advances in Neural Information Processing Systems 35 (2022) 38571–38584.

[96] W. Zhu, X. Ma, Z. Liu, L. Liu, W. Wu, Y. Wang, Motionbert: A unified perspective on learning human motion representations, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 15085–15099.

[97] H.-B. Zhang, Y.-X. Zhang, B. Zhong, Q. Lei, L. Yang, J.-X. Du, D.-S. Chen, A comprehensive survey of vision-based human action recognition methods, Sensors 19 (5) (2019) 1005.

[98] J.-H. Choi, J.-J. Lee, A. Nasridinov, Dance self-learning application and its dance pose evaluations, in: Proceedings of the 36th annual acm symposium on applied computing, 2021, pp. 1037–1045.

[99] R. Hachiuma, F. Sato, T. Sekii, Unified keypoint-based action recognition framework via structured keypoint pooling, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 22962–22971.

[100] V. Arpatzoglou, A. A. Kardara, A. Diehl, B. Flueckiger, S. Helmer, R. Pajarola, Dancemoves: A visual analytics tool for dance movement analysis, EUROVIS (2021).

[101] P. Kishore, K. Kumar, E. Kiran Kumar, A. Sastry, M. Teja Kiran, D. Anil Kumar, M. Prasad, et al., Indian classical dance action identification and classification with convolutional neural networks, Advances in Multimedia 2018 (2018).

[102] K. Lai, S. N. Yanushkevich, V. Shmerko, M. Hou, Capturing causality and bias in human action recognition, Pattern Recognition Letters 147 (2021) 164–171.

[103] F. Baradel, M. Armando, S. Galaaoui, R. Brégier, P. Weinzaepfel, G. Rogez, T. Lucas, Multi-hmr: Multi-person whole-body human mesh recovery in a single shot, arXiv preprint arXiv:2402.14654 (2024).

[104] C. Song, Y. Huang, W. Ouyang, L. Wang, Mask-guided contrastive attention model for person re-identification, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 1179–1188.

[105] M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, S. C. Hoi, Deep learning for person re-identification: A survey and outlook, IEEE transactions on pattern analysis and machine intelligence 44 (6) (2021) 2872–2893.

[106] R. Fensham, Designing for movement: Dance costumes, art schools and natural movement in the early twentieth century, Journal of Design History 28 (4) (2015) 348–367.

[107] K. E. Barford, et al., Drawing, interpretation and costume design: a study of the costumed body informed by watching'tanztheater wuppertal'in rehearsal and performance, Ph.D. thesis, University of the Arts London (2016).

[108] P. Howard, What is Scenography?, 1st Edition, Routledge, London, 2001.

[109] B. Liu, Y. Zhu, K. Song, A. Elgammal, Self-supervised sketch-to-image synthesis, in: Proceedings of the AAAI conference on artificial intelligence, Vol. 35, 2021, pp. 2073–2081.

[110] P. Xu, T. M. Hospedales, Q. Yin, Y.-Z. Song, T. Xiang, L. Wang, Deep learning for free-hand sketch: A survey, IEEE transactions on pattern analysis and machine intelligence 45 (1) (2022) 285–312.

[111] X. Han, Z. Wu, Z. Wu, R. Yu, L. S. Davis, Viton: An image-based virtual try-on network, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7543–7552.

[112] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 10684–10695.

[113] J. Gou, S. Sun, J. Zhang, J. Si, C. Qian, L. Zhang, Taming the power of diffusion models for high-quality virtual try-on with appearance flow, in: Proceedings of the 31st ACM International Conference on Multimedia, 2023, pp. 7599–7607.

[114] Y. Xu, T. Gu, W. Chen, C. Chen, Ootdiffusion: Outfitting fusion based latent diffusion for controllable virtual try-on, Xiaomi Research (2024).

[115] S. Qi, X. Ning, G. Yang, L. Zhang, P. Long, W. Cai, W. Li, Review of multi-view 3d object recognition methods based on deep learning, Displays 69 (2021) 102053.

[116] M. Pepe, V. S. Alfio, D. Costantino, Uav platforms and the sfm-mvs approach in the 3d surveys and modelling: A review in the cultural heritage field, Applied Sciences 12 (24) (2022) 12886.

[117] B. U. Meinen, D. T. Robinson, Mapping erosion and deposition in an agricultural landscape: Optimization of uav image acquisition schemes for sfm-mvs, Remote Sensing of Environment 239 (2020) 111666.

[118] F. Tosi, Y. Zhang, Z. Gong, E. Sandström, S. Mattoccia, M. R. Oswald, M. Poggi, How nerfs and 3d gaussian splatting are reshaping slam: a survey, arXiv preprint arXiv:2402.13255 (2024).

[119] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, Nerf: Representing scenes as neural radiance fields for view synthesis, Communications of the ACM 65 (1) (2021) 99–106.

[120] B. Kerbl, G. Kopanas, T. Leimkühler, G. Drettakis, 3d gaussian splatting for real-time radiance field rendering, ACM Transactions on Graphics 42 (4) (2023) 1–14.

[121] A. Guédon, V. Lepetit, Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering, arXiv preprint arXiv:2311.12775 (2023).

[122] S. Saito, Z. Huang, R. Natsume, S. Morishima, A. Kanazawa, H. Li, Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization, in: Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 2304–2314.

[123] R. Martin-Brualla, N. Radwan, M. S. Sajjadi, J. T. Barron, A. Dosovitskiy, D. Duckworth, Nerf in the wild: Neural radiance fields for unconstrained photo collections, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 7210–7219.

[124] S. Bi, Z. Xu, P. Srinivasan, B. Mildenhall, K. Sunkavalli, M. Hašan, Y. Hold-Geoffroy, D. Kriegman, R. Ramamoorthi, Neural reflectance fields for appearance acquisition, arXiv preprint arXiv:2008.03824 (2020).

[125] F. Simonetta, S. Ntalampiras, F. Avanzini, Multimodal music information processing and retrieval: Survey and future challenges, in: 2019 international workshop on multilayer music representation and processing (MMRP), IEEE, 2019, pp. 10–18.

[126] A. L. Haz, E. D. Fajrianti, N. Funabiki, S. Sukaridhoto, A study of audio-to-text conversion software using whispers model, in: 2023 Sixth International Conference on Vocational Education and Electrical Engineering (ICVEE), IEEE, 2023, pp. 268–273.

[127] M. Ravanelli, T. Parcollet, P. Plantinga, A. Rouhe, S. Cornell, L. Lugosch, C. Subakan, N. Dawalatabad, A. Heba, J. Zhong, J.-C. Chou, S.-L. Yeh, S.-W. Fu, C.-F. Liao, E. Rastorgueva, F. Grondin, W. Aris, H. Na, Y. Gao, R. D. Mori, Y. Bengio, SpeechBrain: A general-purpose speech toolkit, arXiv:2106.04624 (2021). `arXiv:2106.04624`.

[128] L. Pepino, P. Riera, L. Ferrer, Emotion recognition from speech using wav2vec 2.0 embeddings, arXiv preprint arXiv:2104.03502 (2021).

[129] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, T. Dwyer, Data visceralization: Enabling deeper understanding of data using virtual reality, IEEE Transactions on Visualization and Computer Graphics 27 (2) (2020) 1095–1105.

[130] S. Alizadehsalehi, I. Yitmen, Digital twin-based progress monitoring management model through reality capture to extended reality technologies (drx), Smart and Sustainable Built Environment 12 (1) (2021) 200–236.

[131] R. Liu, M. Gao, L. Wang, X. Wang, Y. Xiang, A. Zhang, J. Xia, Y. Chen, S. Chen, Interactive extended reality techniques in information visualization, IEEE Transactions on Human-Machine Systems 52 (6) (2022) 1338–1351.

[132] C. Scorolli, E. N. Grasso, L. Stacchio, V. Armandi, G. Matteucci, G. Marfia, Would you rather come to a tango concert in theater or in vr? aesthetic emotions & social presence in musical experiences, either live, 2d or 3d, Computers in Human Behavior 149 (2023) 107910.

[133] F. Tao, M. Zhang, Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing, IEEE access 5 (2017) 20418–20427.

[134] X. Wang, S. Ong, A. Y.-C. Nee, Multi-modal augmented-reality assembly guidance based on bare-hand interface, Advanced Engineering Informatics 30 (3) (2016) 406–421.

[135] A. Holzinger, B. Malle, A. Saranti, B. Pfeifer, Towards multi-modal causability with graph neural networks enabling information fusion for explainable ai, Information Fusion 71 (2021) 28–37.

[136] M. Chikobava, A. Moisieiev, T. A. Schmeyer, P. Poller, M. Deru, A. Ndiaye, A. Klimenko, C. Braun, J. Baus, B. Brandherm, Multimodal interactive system for visualization of energy data in extended reality (xr) settings, in: International Conference on Human-Computer Interaction, Springer, 2023, pp. 163–168.

[137] B. Marques, S. Silva, R. Maio, J. Alves, C. Ferreira, P. Dias, B. S. Santos, Evaluating outside the box: Lessons learned on extended reality multimodal experiments beyond the laboratory, in: Proceedings of the 25th International Conference on Multimodal Interaction, 2023, pp. 234–242.

[138] B. Fanini, D. Ferdani, E. Demetrescu, S. Berto, E. d'Annibale, Aton: An open-source framework for creating immersive, collaborative and liquid web-apps for cultural heritage, Applied Sciences 11 (22) (2021) 11062.

[139] L. Stacchio, G. Vallasciani, G. Augello, S. Carradori, P. Cascarano, G. Marfia, Wixard: Towards a holistic distributed platform for multi-party and cross-reality webxr experiences, in: 2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), to appear, IEEE, 2024.

[140] P. Bonora, A. Pompilio, Corago in lod: the debut of an opera repository into the linked data arena, Corago in LOD: The debut of an Opera repository into the Linked Data arena (2021) 54–72.

[141] M. Kahane, Rudolf Noureev 1938-1993. Costumes et photographies., CNCS Editions du Me'ce'ene, 2009.