

# PREMIERE

Performing arts in a new era

## D6.1 - Data Indexing

31 July 2023- M10

**Document identifier:** PRMR-D6.1-Data Indexing

**Version:** 1

**Author:** Aggelos Gkiokas, Maximos Kalliakatsos, Grigoris Bastas (ARC), Suzan Tunca (ICK), Muriel Romero, Daniel Bisig (STO), Philippe Colantoni (UJM), Bruno Moreira (FIT)

**Dissemination status:** PU

|                                 |  |
|---------------------------------|--|
| <b>Grant Agreement n°</b>       | 101061303  |
| <b>Project acronym</b>          | PREMIERE   |
| <b>Project title</b>            | Performing arts in a new era: AI and XR tools for better understanding, preservation, enjoyment and accessibility  |
| <b>Funding Scheme</b>           | HORIZON-CL2-2021-HERITAGE-01 (HORIZON Research and Innovation Actions)   |
| <b>Project Duration</b>         | 01/10/2022 – 30/09/2025 (36 months)  |
| <b>Coordinator</b>              | Athena Research Center (ARC)   |
| <b>Associated Beneficiaries</b> | <ul style="list-style-type: none"> <li>• Stichting Amsterdamse Hogeschool voor de Kunsten (AHK)</li> <li>• Forum Dança - Associação Cultural (FDA)</li> <li>• Tempesta Media SL (TMP)</li> <li>• Cyens - Centre of Excellence (CNS)</li> <li>• Kallitechniki Etaireia Argo (ARG)</li> <li>• Medidata.Net - Sistemas de Informação para Autarquias SA (MED)</li> <li>• Fitei Festival Internacional Teatro Expressão Ibérica Cri (FIT)</li> <li>• Instituto Stocos (STO)</li> <li>• Université Jean Monnet Saint-Etienne (UJM)</li> <li>• Associação dos Amigos do Coliseu Do Porto (COL)</li> <li>• Stichting International Choreographic Arts Centre (ICK)</li> </ul> |

# Project no. 101061303

## PREMIERE

Performing arts in a new era: AI and XR tools for better understanding, preservation,  
enjoyment and accessibility

HORIZON-CL2-2021-HERITAGE-01

**Start date of project:** 01/10/2022

**Duration:** 36 months

| History Chart |            |  |                         |  |
|---------------|------------|--|-------------------------|--|
| Issue         | Date       | Changed page(s)                                  | Cause of change         | Implemented by   |
| 0.1           | 04/07/2023 | -  | 1 <sup>st</sup> outline | Aggelos Gkiokas (ARC)                                  |
| 0.2           | 10/07/2023 | ALL  | additions               | Daniel Bisig (STO)<br>Maximos Kalliakatsos (ARC)       |
| 0.3           | 14/07/2023 | ALL  | additions               | Muriel Romero (STO)<br>Suzan Tunca (ICK)               |
| 0.4           | 18/07/2023 | ALL  | additions               | Aggelos Gkiokas (ARC)<br>Daniel Bisig (STO)            |
| 0.5           | 22/07/2023 | ALL  | additions               | Maximos Kalliakatsos (ARC)<br>Philippe Colantoni (UJM) |
| 0.6           | 26/07/2023 | ALL  | additions               | Muriel Romero (STO)<br>Suzan Tunca (ICK)               |
| 0.7           | 29/07/2023 | ALL  | Final corrections       | Aggelos Gkiokas (ARC)                                  |
| 1.0           | 31/07/2023 | ALL  | Final version           | Aggelos Gkiokas (ARC)                                  |
| Validation    |            |  |                         |  |
| No.           | Action     | Beneficiary                                      | Date                    |  |
| 1             | Prepared   | Ezequiel Santos (FD)                             | 29/07/2023              |  |
| 2             | Approved   | Ezequiel Santos (FD)<br>Kalli Foundoulidou (CNS) | 31/07/2023              |  |

|   |          |  |            |
|---|----------|--|------------|
| 3 | Released | Ezequiel Santos (FD)<br>Kalli Foundoulidou (CNS) | 31/07/2023 |
|---|----------|--|------------|

Disclaimer: The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

**All rights reserved.**

The document is proprietary of the PREMIERE consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

# Table of Contents

|   |    |
|---|----|
| Executive Summary.....                                    | 7  |
| Acronyms and abbreviations.....                           | 11 |
| 1. Introduction.....                                      | 12 |
| 2. Data selection.....                                    | 12 |
| 2.1. Selection criteria.....                              | 12 |
| 2.2. Dance selection.....                                 | 13 |
| 2.3. Theatre selection.....                               | 15 |
| 2.4. Copyright management.....                            | 17 |
| 3. Data schema.....                                       | 19 |
| 3.1. Introduction.....                                    | 19 |
| 3.2. Level 1 schema.....                                  | 19 |
| 3.3. Level 2 schema.....                                  | 24 |
| 3.4. Level 3 schema.....                                  | 29 |
| 3.4.1. Theatre schema.....                                | 29 |
| 3.4.2. Dance schema.....                                  | 31 |
| 3.4.2.1. Dance qualities categories.....                  | 31 |
| 3.4.2.2. Universals.....                                  | 34 |
| 3.4.2.3. Signature specific.....                          | 35 |
| 4. Data processing.....                                   | 38 |
| 4.1. Audio and text related data.....                     | 38 |
| 4.2. Video related data.....                              | 38 |
| 5. Data annotation.....                                   | 40 |
| 5.1. Annotation software.....                             | 40 |
| 5.2. Audio annotation.....                                | 41 |
| 5.3. Video annotation.....                                | 44 |
| 6. Recommendations for the Content Management System..... | 46 |
| 6.1. Recommendation for the CMS backend.....              | 46 |

6.2. Recommendation for the CMS front-end..... 46

7. Conclusions .....48

## Executive Summary

This deliverable describes the data considered in 1st use case **T6.2 - Performing Arts Archives browsing enhanced with AR/VR technologies** and how this data will be prepared and loaded in the Content Management System. Thus, it is related to **D5.1 - PREMIERE data and streaming infrastructure v1** and the use case 1 corresponding deliverables.

As a pool of choice, we mainly considered the archives of PREMIERE partners FITEI and Forum Danca. We took into account several criteria for selecting the first set of pieces that we will use in the archives use case. These criteria are related based on the artistic importance, while we tried to keep a large artistic style variety. We also considered videos that have a variety on the “difficulty” of the content with respect to the performance of video and audio analysis algorithms, so that we can understand the performance of the state-of-the-art algorithms in various data settings as well as challenge the SOTA. To do so, we carried out several iterations of a) selecting videos by artistic experts, and b) reviewing and providing feedback on the videos by technical experts of PREMIERE.

This selection process based on the above criteria resulted in a total of 11 performances (6 for theatre, 5 for dance) as the baseline for the pilot 1. However, as the project progresses, we will consider more videos. The final selection of the videos will be reported in the final deliverable of the pilot, **D6.4/D6.5 - Working prototype for enhanced archive browsing (v1/v2)**. Technical details of the selected performances can be found in **Deliverable D2.3 - Specifications for Archived Content**.

Aiming to grant to the project, the use rights of the theatre and dance pieces selected and establishing the legal terms and conditions related to the existing Intellectual Properties (IP) of these works, a copyright license agreement template was drafted and will be signed between the copyright holders of the works and the archive partners (FDA, FIT).

Data are annotated with information that comes from both human annotators and algorithms that are developed by members of the consortium. Additionally, some annotations are uniquely pertinent to the use case of dance or of theatre, while others are common to both.

Human-generated information mainly concern bibliographical information and metadata, e.g., about when a performance was performed, who are the contributors in terms of artistic creation, performance, technical aspects and production, among others. This category of information is captured in an “annotation schema” that lends many characteristics from schema.org; this level of information is referred to by the consortium as “**Level 1**” schema. Level 1 consists of entities such as videos, articles, persons, events, etc. For developing this schema, we started from the schema.org schema which is enriched with some of the Dublin Core Metadata properties. Afterwards we adapted the schema to visit PREMIERE needs. This was done in several meetings within the consortium between scholar/artistic/academic partners and technical partners. We ended up with a rather simple schema for the sake of simplicity and intuitiveness. We consider seven basic types of information:

“**Level 2**” annotations incorporate information about components or “annotations” that are common in dance and theatre and can be extracted from the video, audio or text (extracted from audio) of a performance recording. Annotations of these components refer to “quantifiable” aspects that can be located in time and/or space on stage or on screen and can either be a) extracted by **algorithmic analysis** on video, audio and text b) **manually** annotated by humans, or c) annotated using a hybrid approach, i.e. a first automated annotation step, followed by human correction of verifications. The annotation for this level is computed on a time basis. The overall schema for Level 2 must accommodate different

types of information about identified components or annotations, depending on the type of information that each annotation describes. For example, audio-related algorithms can identify that music starts at some specific time instance, or video-related algorithms can identify that the lighting conditions on stage have changed at a specific time instance; in both examples, there is no concept of “position on stage” for the identified events. In other examples, however, “position on stage” information may be meaningful, e.g., in the case where a video algorithm detects motion of a dancer or actor who moves from one place to another on stage. Additionally, the position of a dance or actor, object or event on stage, may be described in different ways that are meaningful for different applications. For instance, the position of an object on stage would be better to be described by a 2-dimensional or 3-dimensional bounding box, in the cases of video and 3D virtual theatre respectively. Similarly, the body of a dancer or actor would be better described by demarcating the exact shape of the part of body that is visible to the viewer (either on video or on a 3D virtual theatre setting).

“Level 3” annotations are quite similar with level 2 but place a stronger emphasis on qualitative aspects of performance works that are of greater interest to scholars and artists; these aspects are specialised for dance and theatre. They are usually computed on a time basis, e.g., the values of the various characteristics are evolving over time. They can be either annotated automatically or manually. “Level 3” annotations are quite similar with level 2, with the difference that they are specialised for either dance or theatre. As in the case of level 2 annotations, they are usually considered on a time basis and can be either computed by algorithms or manually annotated.

The Theater schema is, above all, related to the form of the performers' actions on stage. The voice and the body are the actor's main tools on the scene and supplemented with the text of the play they are the starting point for these annotations. Mainly in post-dramatic and physical theatre, the movement of the performers provides us with a set of complementary information to the rest of the content of the scene that we consider important to highlight. The use of repetition, movement and immobilization, the tension of the bodies, the positioning of the bodies in the scene, all of this is relevant for the spectator and, many times, contemplated in the dramaturgy of the performance. The fact that those actions are almost always interrelated with the Director's choice on the essence of the performance, constitutes one more reason to be incorporated to Level 3 Theatre Schema. All basic actions such as walking from A to B or turning your face in a certain way, manipulating objects and scene props and the way this is done, are also important.

Dance schema mainly consists of movement qualities. These qualities are usually present in a time basis. The first ten months of PREMIERE, a great effort has been dedicated by the dance specialist members of the consortium (ICK, STO, FD) to finding a common annotation schema for dance. This effort had the difficulty of the non-existence of a globally accepted dance movement/qualities vocabulary, or in cases that there is one (e.g., the Laban glossary), it is not complete in relation to contemporary developments in dance and needs additional specifications.

Although we aimed on working towards finding a common and globally accepted dance schema, this very optimistic goal requires much more time and effort that may fall out of the scope of PREMIERE. For the needs of the project, we decided to move on by defining a fragmented collection of dance annotation sub-schemas. This fragmentation reflects the current trends in the dance scientific community of a non-agreement, but on the other hand is descriptive enough for the sake of PREMIERE. We settled on three main fragments of ontologies. The first one consists of a general hierarchical structure of broader categories. The other two consist of specific movement qualities ontologies, a universal one based on



Laban analysis, and a signature-specific one based on the research made by STO, ICK and FD.

The information that is indexed with audio-related tools, from the audio that is retrieved from an archived video (in the case of Use Case 1). In use cases that require real-time processing, the same techniques are applied to buffered data. Data processing of audio extracts features that are related to music, voice and audio events, as described in detail in Section 3.3.

The audio modality also provides information about text-related data. Text is extracted from speech audio using Automatic Speech Recognition models. To increase the accuracy of speech-to-text transcription, Voice Activity Detection is first applied for removing long pauses of speech or the presence of music. Given a clear transcription, the features described for the text modality in Section 3.3 are extracted. Entities that are identified in the text modality, are combined with metadata information and are used to retrieve information from online resources (e.g., Wikipedia).

Data processing of video extracts features that are related to frame (image) and multiple frames events. This video needs to be split into sequences which break down the video into meaningful units, allowing for efficient organization, indexing, and analysis of video data. It is then possible to break down these scenes into image sequences (frames) to automatically extract low level features that are then used to determine high level features.

For manually or semi-automatically annotating the data, we will deploy one or more open source available tools. These should be integrated in some way with the CMS, so that these annotations tools can accessed by the annotators.

We propose a series of recommendations for the CMS backend based on the contents of this deliverable. These should be taken into account for the development of the CMS, in conjunction with the user requirements presented in Deliverable D2.1 - User Requirements v1.

Data indexing and representation:

- All items of the CMS should be represented by an item belonging to the ontology described in Section 3.
  - All items should have a unique identifier
  - CMS should contain all the information between relations of items (e.g., Person A belongs to Organisation B Person C is the director of the play D).
  - Data storing and access
- Different data items should be potentially stored to different locations that are accessible through a URI. This allows a scalable and distributed data storage workload.
  - Similarly, data and the extracted metadata (e.g., a video, along with the extracted features such as performers position).
  - CMS should be able to handle permissions and access content layers
  - Compatibility and Interoperability
- Data formats stored in the CMS should be compatible with the data formats used in the annotation tool (Section 5)
  - Data stored in the CMS should be compatible and linked to the various algorithms from WP3 and WP4.
  - Data stored in the CMS should be compatible and linked to the various algorithms from WP3 and WP4.

The main recommendation of the front-end is that the data should be visualised in a uniform, user friendly and coherent way. Moreover, these interfaces should be in line with back-end recommendations mentioned before. For example, they should allow to correct to some extent and add/remove new annotations from all three levels of the data annotation schema. It is recommendable, that for each of the main data types of the Level 1 schema, we have a dedicated template page to visualise this data.

## Acronyms and abbreviations

| <b>Abbreviatio</b> | <b>Description</b>           |
|--------------------|------------------------------|
| CMS                | Content Management System    |
| SOTA               | state-of-the-art             |
| IP                 | Intellectual Properties      |
| JSON               | JavaScript Object Notation   |
| LMA                | Laban Movement Analysis      |
| ASR                | Automatic Speech Recognition |
| 2D                 | 2 dimensional                |
| 3D                 | 3 dimensional                |
| AI                 | Artificial Intlligence       |

## 1. Introduction

This deliverable describes the data considered in 1st use case T6.2 - **Performing Arts Archives browsing enhanced with AR/VR technologies** and how this data will be prepared and loaded in the Content Management System (CMS). Thus, it is related to D5.1 - **PREMIERE data and streaming infrastructure v1** and the use case 1 corresponding deliverables. In the next section of this deliverable (Section 2), we will describe the data that we consider for the use case and the criteria that we took into account for this selection. Next (Section 3), we will describe the underlying data schema that we take into account. This data schema designed in the first period of the project, comprises of elements taken from existing well-studied schemas such as the schema.org and the Core-Doubling and is adapted based on the needs of PREMIERE and by considering domain knowledge for theatre and dance archives. Section 4 describes some baseline processing of the data for various modalities and a preliminary sketch on how the data will be organised to the CMS. Section 5 describes the data annotation tools, and section 5 provides recommendations that should be taken into account for the development of the CMS. Section 6 concludes this deliverable.

## 2. Data selection

In this section we describe the data selected as a preliminary set of videos that will be loaded into the CMS for the Use case 1. This selection is subject to change, and it might be enriched till the end of the project.

### 2.1. Selection criteria

As a pool of choice, we mainly considered the archives of PREMIERE partners FITEI and Forum Danca. We took into account several criteria for selecting the first set of pieces that we will use in the archives use case. These criteria are related based on the artistic importance, while we tried to keep a large artistic style variety. We also considered videos that have a variety on the “difficulty” of the content with respect to the performance of video and audio analysis algorithms, so that we can understand the performance of the state-of-the-art (SOTA) algorithms in various data settings as well as challenge the SOTA. To do so, we carried out several iterations of a) selecting videos by artistic experts, b) reviewing and providing feedback on the videos by technical experts of PREMIERE. The selected criteria are summarised as following:

- **Artistic importance:** We selected performances that are important and have an historical value in the field. This was mainly accomplished by the specialised partners FITEI, Forum Danca and ICK.
- **Variety and balance:** The selected performances were selected such that we have a variety and balance with respect to several aspects such as **epoch, artistic styles, genre, language and country of origin.**
- **Technical challenges:** We considered performances that show a variety of technical challenges for their processing, such as **number of cameras, camera resolution, lighting conditions, use of props, contrast, maximum number of performers in scene** and other difficult situations such as **performers bodies very close together (occlusions), extreme body poses, scenes with actors speaking concurrently,** and other artefacts such as **shadows.**
- **Copyright permissions:** We also considered performances for which it is feasible to acquire copyright permissions. This is an ongoing process, described in section 2.4.

This selection process based on the above criteria resulted in a total of 11 performances (6 for theatre, 5 for dance) as the baseline for the pilot 1. However, as the project progresses, we will consider more videos. The final selection of the videos will be reported in the final deliverable of the pilot, **D6.4/D6.5 - Working prototype for enhanced archive browsing (v1/v2)**. Technical details of the selected performances can be found in **Deliverable D2.3 - Specifications for Archived Content**.

## 2.2. Dance selection

Here we list all the dance pieces selected and some information about them. In the description field we provide insights on the selection of each performance as well as some bibliographic information.

|                    |   |
|--------------------|---|
| <b>Title</b>       | Extra Dry   |
| <b>Creator</b>     | Emio Greco / Pieter C. Scholten   |
| <b>Year</b>        | 1999  |
| <b>Country</b>     | Netherlands   |
| <b>Description</b> | <p><u>Vocabulary</u>: body centre-periphery dialogue is intense; dominant effort-shape qualities of time and flow; intense quality of floor and air drawings with intense curves and spirals; movement performed in diagonal and out of balance; contralateral isolation of limbs; body awareness of pre-movement (or the intention before movement that becomes visible) is very present and cohesive; sats quality (Barba); body fluids and bone structure are strong movement ignitors. Animus energetic quality is very present. Bodies are expressive from head to toes. The performance is physically intense.</p> <p><u>Interest to viewers</u>: the scenic stage mirrors the body inner space and becomes very alive and organic; the question of identity: it is a duet form, but are they two entities or just the same body? Time and its circularity are amplified by the diverse musical score. The physical performance reveals several technical (historical) layers and constitutes a document by itself.</p> <p><u>Interest to artists as users</u>: The performance is vibrant, bodies are in energetic conditions that allow us to perceive movements before they even come out; therefore, the performance may keep its ontological condition, which survives albeit the documentation/notation form.</p> <p><u>Interest to teachers as users – key ideas</u>: Connects to dance and body representation in Renaissance: the body seeks its manner of expressing its interiority and occupies geometric space as if in pursue of its place in the universe; bodies are in-between states (Philosophy and energy).</p> |

|                    |  |
|--------------------|--|
| <b>Title</b>       | Perhaps she could dance first and think afterwards   |
| <b>Creator</b>     | Vera Mantero   |
| <b>Year</b>        | 1991   |
| <b>Country</b>     | Portugal   |
| <b>Description</b> | <p>Premiered at Klapstuk 1991, Leuven (B), it is a performance of good reputation in European dance. Here a structured improvisation is used to process the idea of impossibility (and what one can express with dance as a medium; what is dance, after all?). The body shows the record of different dance techniques, as it so happens in the case of a literary palimpsest. This piece is simultaneously intense and intimate.</p> |

|                    |   |
|--------------------|---|
| <b>Title</b>       | Stocos  |
| <b>Creator</b>     | Pablo Palacio y Muriel Romero.  |
| <b>Year</b>        | 2011  |
| <b>Country</b>     | Spain   |
| <b>Description</b> | <p><u>Interest to viewers:</u> Visually, acoustically and physically very strong with intensity shades; there is a sense of synaesthesia as movement/sound/image blend easily.</p> <p><u>Interest to artists as users:</u> Knowledge of the technologies under use; example of efficiency and transparency in the use of technologies which are co-creators of dramaturgy; the use of body, movement and sound as notions to create the idea of a living scenic ecology.</p> <p><u>Interest to teachers as users – key ideas:</u> illustration of the use of AI technologies and the construction of dramaturgies; foster the discussion on the aesthetics of dance (body movement as an interface to hidden forces); psychological principles of active imagination, self, shadow and revelation; the self that goes through a process of Alchemy.</p> |

|                    |   |
|--------------------|---|
| <b>Title</b>       | Guintche  |
| <b>Creator</b>     | Marlene Monteiro Freitas  |
| <b>Year</b>        | 2010  |
| <b>Country</b>     | Portugal  |
| <b>Description</b> | <p>Choice based on artistic value and actuality - originally from Cape Verde and being raised in Portugal, dancer/choreographer Marlene Monteiro Freitas is an established choreographer in the European and Western space having received the silver Venice Biennale award in 2018. Across her works, one can tell that African episodes and imagery are linked to current political and social global issues and are worked upon a strong corporal investment. Her artistic language is very Dionysian and original, comprised of total body expressiveness and very sudden and controlled movements that give an impression of watching old cinematography. African spirit comes out along the repetitive movements, in voice sonorities, in the hips, arms and the mask-like faces. This piece is selected out of Marlene Monteiro Freitas' repertoire, because it is scenically simple - with few performers in the cast and little scenography.</p> |

|                    |   |
|--------------------|---|
| <b>Title</b>       | Fecundação e Alívio Neste Chão Irredutível onde Com Gozo Me Insurjo   |
| <b>Creator</b>     | Joana von Mayer Trindade & Hugo Calhim Cristóvão  |
| <b>Year</b>        | 2021  |
| <b>Country</b>     | Portugal  |
| <b>Description</b> | <p>Energetic, sometimes trance-like with long sequences of solos and duets conducive to the capture and clear reading of 2D movement (cf. trailer). Inspired by the late Portuguese visual artist Ana Hatherly, the music as well as the power relation between bodies reflect ambiances from her literary and visual production. This recording may be interesting to use regarding gender studies, because the woman has the dominant role with the male dancer remaining as background and</p> |

|  |   |
|--|---|
|  | submissive figure. Some sequences of contact between the male and female dancers refer to androgyny, and others may refer to the crushing of a masculine archetype. The piece is supported by an extensive philosophical and scientific research, which resulted in a book (printed and accessible online). |
|--|---|

### 2.3. Theatre selection

For theatre we selected plays that were presented in the FITEI international festival. In the tables below we list all the theatre pieces selected and some information about them. In the description field we provide insights on the selection of each performance as well as some bibliographic information.

|                    |   |
|--------------------|---|
| <b>Title</b>       | "Noite de Reis" - William Shakespeare (Twelfth Night, Or What You Will)   |
| <b>Creator</b>     | Ricardo Pais  |
| <b>Year</b>        | 1998  |
| <b>Country</b>     | Portugal  |
| <b>Description</b> | In 1998, Ricardo Pais took Noite de Reis, one of William Shakespeare's most popular comedies, to the stage of Teatro São João, based on a translation by António M. Feijó.<br>"Noite de Reis" is a comedy about love. In the kingdom of Illyria, Duke Orsino is in love with Olivia, who does not love him. A young woman, Violeta, arrives in Illyria, washed away by the sea after a shipwreck. She has a twin brother, Sebastião, who she believes drowned in the shipwreck. Violeta disguises herself as a man, changes her name to Cesario and finds work as a messenger for Orsino. Violeta's job is to send love messages from Orsino to Olivia. Olivia, falls in love with Cesario (Violeta), thinking she is a man. Violeta falls in love with Orsino, but cannot reveal her love for him, as Orsino thinks she is Cesario, a man. Thus, a storm of love is created. |

|                    |   |
|--------------------|---|
| <b>Title</b>       | A Caminhada dos Elefantes   |
| <b>Creator</b>     | Inês Barahona e Miguel Fragata  |
| <b>Year</b>        | 2013  |
| <b>Country</b>     | Portugal  |
| <b>Description</b> | This show tells the story of a man and a herd of elephants. When the man dies, the elephants make a mysterious walk to his house, to pay him a last homage: he was not just any man, he was one of them.<br>"The Elephants' Walk" is about existence, life and death, and the journey we all have to make, one day, to say goodbye to someone. A show that reflects on the end, which is a mystery for all of us, children or adults.<br>"The Elephants' Walk" was preceded by extensive research work with about 200 children aged between 6 and 10, through meetings and workshops. The material collected served as inspiration and content for the show.<br>It is a show of enormous sensitivity, fruit of a long research with the age group for which it was intended. The theme of death is enunciated as a parable, deconstructed, with the motto let's not talk about death. |



|  |  |
|--|--|
|  | Dramaturgically the work unfolds in an interactive way with the spectator, where we are invited to be part of the proposed reflection. |
|--|--|

|                    |   |
|--------------------|---|
| <b>Title</b>       | O Gatilho da Felicidade   |
| <b>Creator</b>     | Ana Borralho & João Galante   |
| <b>Year</b>        | 2017  |
| <b>Country</b>     | Portugal  |
| <b>Description</b> | <p>The definition of happiness varies from person to person. It can even be considered a state. A state of well-being and satisfaction. How can we be happy? And above all, how can we be happy in an unhappy society? Ana Borralho and João Galante invite 12 young adults in each location. They are not actors and are between 18 and 23 years old. Sitting at the same table, with a gun in the middle, they play Russian roulette and answer a list of questions about their loves, their worries, their concerns, their families, their visions of the future.</p> <p>The words come out, the life paths overlap, drawing a multicultural and explosive cartography of the local youth.</p> <p>A deadly game in search of happiness.</p> <p>This work has toured several countries around the world, made with teenagers from each locality and in their native language. Within the experimental work of the artists, it is a show that maintains an enormous plastic sobriety and a roughness of the thematic proposals and scenic solutions, which gives it a strong identity.</p> |

|                    |   |
|--------------------|---|
| <b>Title</b>       | Moria   |
| <b>Creator</b>     | Una Hora Menos  |
| <b>Year</b>        | 2021  |
| <b>Country</b>     | Spain   |
| <b>Description</b> | <p>"Moria" is an immersive theatre proposal by the Canarian company Unahoramenos.</p> <p>This show tells the story of two refugees and their families who have no choice but to flee their respective countries. The dramaturgy is based on real testimonies filmed in the Moria camp under the supervision of Nicolás Castellano, a reporter specialising in forced migratory movements and human rights for the last 20 years.</p> <p>Zohra Amiryar had to survive four bombings and attacks in her country, Afghanistan, lose her brother in one of them and see how her four children ended up wounded, in the last of these attacks, to make the exodus to seek refuge for her family. And Douaa Alhavatem is a thirty-something Iraqi woman who had to leave her native Baghdad after her husband disappeared for fear that she and her three children might suffer the same fate. On the journey, she and her little ones gave up everything after 11 hours at sea with no sign of any rescue ship. Both lived a new stage of suffering in Moria that they never imagined they would suffer when they set foot on European soil, the place where they believed they would finally feel safe.</p> <p>A common space, where they are accompanied by countless unknown people, who are also going through the worst moment of their lives.</p> <p>A show of strong emotional impact, the audience sits inside the tent, on blankets on the floor, very close to the actresses, it is a show that uses</p> |



|  |  |
|--|--|
|  | much of the technology, most recent, in video capture and projection, in 360°. |
|--|--|

|                    |  |
|--------------------|--|
| <b>Title</b>       | Olo  |
| <b>Creator</b>     | Teatro de Ferro  |
| <b>Year</b>        | 2014   |
| <b>Country</b>     | Portugal   |
| <b>Description</b> | <p>This puppet play is animated by a question: is it possible to represent what happens when we lock ourselves alone in a rehearsal room, in a workshop with the aim of creating something new? What is discovered and what is invented in this chimera? What was already there? What does the black box record? Are we really alone when we are on stage solo? What good is what we live, read, dream, wish, fear? What good is what has already been done by others, in other times and in other places? It is also about the role of memory as a process of evocation that relates to life happening (a process of discovery) that this play tells us about.</p> <p>There is an enormous technical virtuosity of the manipulator, associated to the identity change of the company, which explores original plastic universes, in which all the most difficult materials to work with are called upon, from iron, to wood, to organic materials and live video.</p> |

|                    |  |
|--------------------|--|
| <b>Title</b>       | Ensaio Sobre a Cegueira  |
| <b>Creator</b>     | Teatro o Bando   |
| <b>Year</b>        | 2004   |
| <b>Country</b>     | Portugal   |
| <b>Description</b> | <p>In an anniversary year - O Bando celebrates its 30th anniversary, an unusual event in the Portuguese and European theatre scene - the company led by João Brites adds another stage to a journey marked by the questioning of universal values. The focus remains on the concern to go beyond the obvious in order to invest in a theatre "that looks inwards". Blindness, by José Saramago, Portuguese author of reference, could not reflect better this brand image. Taking the allegory "close your eyes to see better" as a starting point - which almost literally gave the name to the project Stop Seeing to See Better, the first time Saramago's work has been used - O Bando resumes the reflection on blindness, not of those deprived of the sense of sight, but of those who seeing, cannot see. The message seems clear: "When we all close our eyes and open them again, it will be possible to see, build and create a different, fairer, freer world." A show with a strong visual component. The characteristic plasticity of Teatro o Bando, supported by an original composition by the conductor Jorge Salgueiro, movement support by Luca Aprea, set design by Rui Francisco, and staging by João Brites. A large cast placed in a scenic device along the lines of Teatro o Bando's traditional stage machines (scenic devices that have their own dramaturgy).</p> |

## 2.4. Copyright management

Aiming to grant to the project, the use rights of the theatre and dance pieces selected and establishing the legal terms and conditions related to the existing Intellectual Properties (IP)

of these works, a copyright license agreement template was drafted and will be signed between the copyright holders of the works and the archive partners (FDA, FIT).

The copyright license is an official permission given by a copyright holder (the licensor) to the user of its copyrighted work (the licensee), by means of an agreement, allowing the licensee to use the copyrighted work. Considering relevant documentation of the European IP Helpdesk, regarding key ingredients of an IP license, a first version was drafted by TMP, including:

- Introduction explaining the context and the scope of the license.
- Part I. Definition of the parts and their legal data.
- Part II. Object of the license, definition of the work(s) involved and of the granted rights.
- Part III. Terms of the Agreement.

Overall, the agreement defines a non-exclusive license, in force for a period of five years and tacitly extended for successive five-year periods, that foresees the following uses:

- Internal reproduction for software development and testing.
- Reproduction for research communication in seminars, conferences, and similar events and for showcasing the project's results in user testing/design workshops, pilot projects and report documents, including for the European Commission.
- Use of the image and content of the work(s) for communication and dissemination activities.

The copyright holder remains the owner of the works, defines the rights statement and has the right to discontinue the collaboration.

On an initial phase of the IPR we identified the need to consider the copyright status of the selected video recordings that will be enhanced: FDA have consulted a lawyer specialized in copyright and related rights in order to prepare the process with the selected works; a dialogue was kept open with TMP to trigger a standard document for the consortium and intermediation process with our lawyer to review the document and suggest further improvements. The above included:

- Meetings between FDA, FIT and COL – in order to find the specifications needed to be included/stated in the agreement.
- Meetings with TMP for joint writing of the agreement.
- Meetings with the lawyer – in order to comply with the Portuguese legislation.

In writing the initial document with TMP, the artist companies of the consortium were also consulted about the agreement, to give their feedback (ICK, STO), and the artists with selected works have also been informed, during our contacts, that an agreement concerning copyright issues will need to be signed. At this stage, the final English version of the agreement has already been reviewed by the FDA's specialist lawyer and, at the moment, the technical translation into Portuguese is being drafted. The final document, to be signed by the parties, will be bilingual, with the Portuguese version prevailing over the English version, since these agreements must be under the purview of Portuguese legislation in the first place.

### 3. Data schema

#### 3.1. Introduction

The data is annotated with information that comes from both human annotators and algorithms that are developed by members of the consortium. Additionally, some annotations are uniquely pertinent to the use case of dance or of theatre, while others are common to both. Human-generated information mainly concerns bibliographical information and metadata, e.g., about when a performance was performed, who are the contributors in terms of artistic creation, performance, technical aspects and production, among others. This category of information is captured in an “annotation schema” that lends many characteristics from schema.org; this level of information is referred to by the consortium as “Level 1” schema.

“Level 2” annotations incorporate information about components or “annotations” that are common in dance and theatre and can be extracted from the video, audio or text (extracted from audio) of a performance recording. Annotations of these components refer to “quantifiable” aspects that can be located in time and/or space on stage or on screen and can either be a) extracted by **algorithmic analysis** on video, audio and text, b) **manually** annotated by humans, or c) annotated using a hybrid approach, i.e. a first automated annotation step, followed by human correction of verifications. The annotation for this level is computed on a time basis.

“Level 3” annotations are quite similar with level 2 but place a stronger emphasis on qualitative aspects of performance works that are of greater interest to scholars and artists.; these aspects are specialised for dance and theatre. They are usually computed on a time basis, e.g., the values of the various characteristics are evolving over time. They can be either annotated automatically or manually.

#### 3.2. Level 1 schema

Level 1 consists of entities such as videos, articles, persons, events, etc. For developing this schema, we started from the schema.org schema which is enriched with some of the Dublin Core Metadata properties. Afterwards we adapted the schema to fit PREMIERE’s needs. This was done in several meetings within the consortium between scholar/artistic/academic partners and technical partners. We ended up with a rather simple schema for the sake of simplicity and intuitiveness. We consider seven basic types of information:

- **Persons**: Correspond to individual persons (e.g., performers, dancer, directors, finders etc)
- **Organizations**: Can be any dance, theatre group or any other related organization such as a venue or a funding entity/organization.
- **Cast and credits**: It is a structure that consists of a group of persons and organizations that comprise the cast and credits of a production / performance.
- **Performances/videos**: It is the main object that represents a video archive of a specific performance or production.
- **Events**: Correspond to a performance or a series of a performance.
- **Articles/Books**: Correspond to articles or books that are related to a specific performance.
- **Other media**: Can be flyers, photos and other media material.

These data types are represented as a linked data type structure, as illustrated in Figure 1.

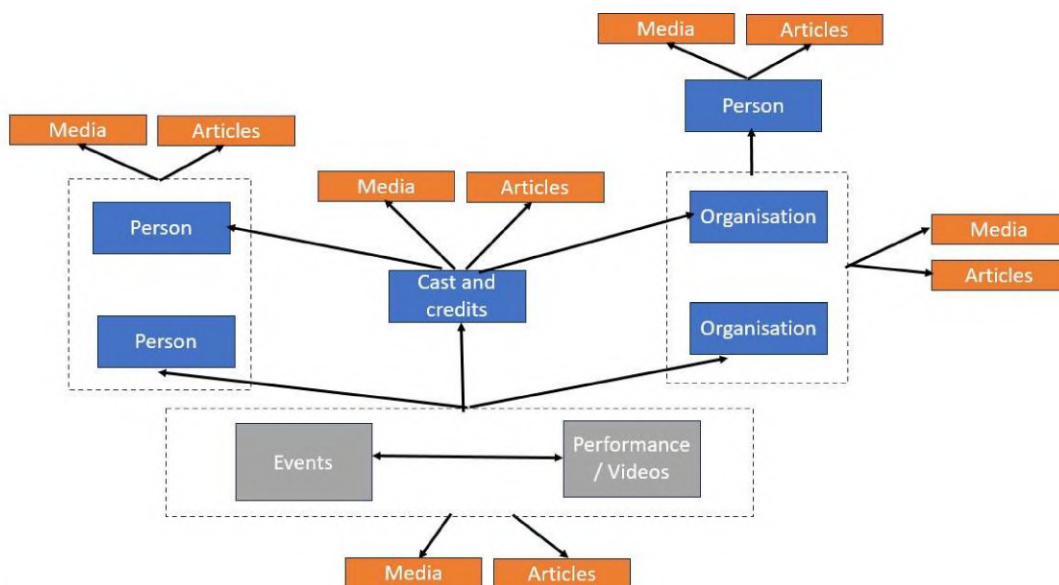


Figure 1 Main PREMIERE data types and their relations.

The main data structures are the **Event** and **Performance/Video** structures that represent either individual recordings of performances (Performance/Video) or a specific production (Event) that can be repeated (e.g. for a whole season). These two structures are associated with the structure **Cast and Credits**, other **Persons** and other **Organisations**. All these structures are further associated with other **Media** objects such as photos, posters, flyers and **Articles/Books** objects. In the following tables we will present each of the data types.

**Person:** Represents a person. Can be a performer, director or any casting member.

| Field Name                | Type         | Example  |
|---------------------------|--------------|--|
| Artistic Name             | Text         | André Lepecki  |
| Nationality               | Text         | Brazilian  |
| Gender                    | Text         | Male   |
| Birth date                | Date         | 1965-07-15   |
| Death date                | Date         | n.a.   |
| Family name               | Text         | Lepecki  |
| Given name                | Text         | André  |
| Affiliation               | Organisation | New York University – Department of Performing Studies |
| Contact point             | Text         | Atl1@nyu.edu   |
| Member of                 | Organisation | Pós d' Arte  |
| Associated media          | Media        | Any photo or other media associated with this person   |
| Associated Articles/Books | Article/Book | Any article, book or other associated with this person |
| Biography                 | Text         | Person's biography                                     |

**Organisation:** Represents an organization. Can be a dance or theatre group or a production company, or a funding organization.

| Field Name                | Type           | Example   |
|---------------------------|----------------|---|
| Name                      | Text           | O Rumor do Fumo   |
| Address                   | Text           | Espaço da Penha, Tv. do Calado, 26-B, 1170-070 Lisbon       |
| Founding date             | Date           | 1999  |
| Cessation date            | Date           | n.a.  |
| Founding location         | Text           | Lisbon  |
| Sponsor                   | Organisation   | DGArtes/Ministério da Cultura                               |
| Funding                   | Organisation   | Câmara Municipal de Lisboa                                  |
| Email                     | Text           | Info@orumodofumo.pt   |
| Website                   | Text           | www.orumodofumo.pt  |
| Founder                   | Person         | Vera Mantero  |
| Keywords                  | Text           | Choreographer, artistic research, independent dance company |
| Type                      | Text           | Cultural association  |
| Associated Media          | Media          | Any surrounding material such as photos, flyers etc         |
| Associated Articles/Books | Articles/Books | Any articles or books associated with this Event            |

**Cast and credits:** Represent the cast and credits of a production or of a specific performance. We have a unique structure for both dance and theatre. Thus, some of the fields may remain empty. For clarity, we also categorise the cast and credits into three categories, artistic, technical and credits.

| Field Name              | Type                  | Example                                    |
|-------------------------|-----------------------|--|
| <b>Artistic cast</b>    |                       |  |
| Performers              | Person                | Andy Deneys, Emio Greco                    |
| Director                | Person                | Pieter C. Scholten, Emio Greco             |
| Playwright/Scriptwriter | Person                | n.a.                                       |
| Choreographer           | Person                | Emio Greco, Pieter C. Scholten             |
| Music director          | Person                | n.a.                                       |
| Set designer            | Person                | Emio Greco, Pieter C. Scholten             |
| Costume designed        | Person                | Clifford Portier                           |
| Light designer          | Person                | Henk Danner                                |
| Sound designer          | Person                | Emio Greco, Pieter C. Scholten             |
| <b>Technical cast</b>   |                       |  |
| Producer                | Person / Organisation | Emio Greco   PC, Stichting Zwaanproducties |
| Assistants              | Person                | n.a.                                       |
| Audio technician        | Person                | n.a.                                       |
| Author of recording     | Person                | Erik Lint                                  |
| Stage manager           | Person                | Paul Beumer                                |
| <b>Credits</b>          |                       |  |
| Funding                 | Person / Organisation | Amsterdams Fonds voor de Kunsten - NL      |



|                         |                       |  |
|-------------------------|-----------------------|--|
|                         |                       | Springdance - NL<br>Kaaitheater - B<br>Tanz in August - D                                |
| <b>Organizer</b>        | Person / Organisation | Stichting Zwaanproducties  |
| <b>Sponsor</b>          | Person / Organisation | Kaaitheater (B), Springdance (NL),<br>Tanzwerkstatt Berlin (D), Klapstuk<br>Festival (B) |
| <b>Copyright holder</b> | Person / Organisation | EG   PC  |

**Performance/video:** This is the main element of the archives and it represent a specific video of a performance. It is also related to the Event item (see below), that represents a more general item in the hierarchy.

| Field Name                      | Type              | Example  |
|---------------------------------|-------------------|--|
| <b>Title</b>                    | Text              | Perhaps She Could Dance First and Think Afterwards   |
| <b>Cast and credits</b>         | Cast and credits  | Choreography/Dance/Costumes: Vera Mantero; Set Design: André Lepecki; Light Design: João Paulo Xavier; Production: Pós d' Arte; Assistant: Albino Moura; Funding: IPJ-Instituto Português da Juventude; Organizer: Klapstuk Festival '91; Sponsor: Europália '91 Portugal. |
| <b>About</b>                    | Text              | Solo performed in the frame of "Os Novos Portugueses" as part of Klapstuk Festival (1991)  |
| <b>Length</b>                   | Number            | 19M  |
| <b>Genre</b>                    | Text              | Contemporary dance   |
| <b>Language</b>                 | Text              | PT/EN  |
| <b>Keywords</b>                 | Text              | Vera Mantero, improvisation, Klapstuk Festival, Europália '91, European dance  |
| <b>Date</b>                     | Date              | 1991.10.25   |
| <b>Associated works</b>         | Performance/Event | "20 Dancers for the XX Century" by Boris Charmatz  |
| <b>Associated persons</b>       | Person            | Vera Mantero, André Lepecki  |
| <b>Associated organisations</b> | Organisation      | O Rumo do Fumo, Pós d' Arte  |
| <b>Associated Media</b>         | Media             | José Fabião (1991)   |
| <b>Articles</b>                 | Articles/Books    | Lepecki, A (1997). "Margins of the Present - A dialogical exploration of the work of Vera Mantero and Francisco Camacho", in M. J. Fazenda & Danças na Cidade (Eds) Present Movements (pp. 47-58). Lisbon: Cotovia.  |

**Event:** Represent events. This can be a series of a performances, or a specific performance (e.g. the premiere of a performance).

| Field Name | Type | Example |
|------------|------|---------|
|------------|------|---------|

|                                  |                   |  |
|----------------------------------|-------------------|--|
| <b>Title</b>                     | Text              | Perhaps She Could Dance First and Think Afterwards   |
| <b>About</b>                     | Text              | Solo performed in the frame of “Os Novos Portugueses” as part of Klapstuk Festival (1991) and still on Tour  |
| <b>Start / End date</b>          | Date              | 1991-10-25/...   |
| <b>Keywords</b>                  | Text              | Vera Mantero, improvisation, Klapstuk Festival, Europália '91, European dance  |
| <b>Location</b>                  | Text              | Soetezaal, Naamsestraat 96, Leuven (B)   |
| <b>Description</b>               | Text              | It is a solo that follows a structured improvisation. It has been presented internationally ever since, both in theatres and nonconventional spaces such as art galleries and gardens  |
| <b>Associated works</b>          | Performance/Video | Video recording of the premiere in Klapstuk Festival (1991-10-25)  |
| <b>Associated persons</b>        | Person            | Vera Mantero, André Lepecki, Boris Charmatz  |
| <b>Associated organisations</b>  | Organisation      | O Rumo do Fumo, Pós d' Arte  |
| <b>CASTING</b>                   | Cast and Credits  | Choreography/Dance/Costumes: Vera Mantero; Set Design: André Lepecki; Light Design: João Paulo Xavier; Production: Pós d' Arte; Assistant: Albino Moura; Funding: IPJ-Instituto Português da Juventude; Organizer: Klapstuk Festival '91; Sponsor: Europália '91 Portugal. |
| <b>Associated Media</b>          | Media             | Any surrounding material such as photos, flyers etc  |
| <b>Associated Articles/Books</b> | Articles/Books    | Any articles or books associated with this Event   |

**Articles/Books:** Represent scholar/review articles or books. These can be related to a performance, that can be published for instance in newspapers, culture magazines, or can be scientific publications of scientific / scholar books.

| Field Name              | Type   | Example  |
|-------------------------|--------|--|
| <b>Title</b>            | Text   | Structural aspects in Acusmatrix   |
| <b>Author(s)</b>        | Person | Pablo Palacio & Muriel Romero  |
| <b>Year of edition</b>  | Date   | 2009   |
| <b>File format</b>      | Text   | PDF  |
| <b>Type of document</b> | Text   | Scholar article  |
| <b>About</b>            | Text   | Reviews both the technologies and the theories that sustain 'Acusmatrix' and 'Catexis', the first two pieces of the trilogy in which “Stocos” is included. |

|                    |                       |   |
|--------------------|-----------------------|---|
| Number of pages    | Number                | 10  |
| Start / End page   | Number                | 53-63   |
| Editor / Publisher | Person / Organisation | CAIRON - Revista de Estudios de Danza / Universidad de Alcalá |

**Media object:** It represents surrounding media objects that accompany a performance/play. This can be any advertisement / communication material such as such as flyers, posters, drawings, photos of biographies, from rehearsals or from acting, or audio files of spoken interview, a commentary, or a song.

| Field Name         | Type                | Example   |
|--------------------|---------------------|---|
| Typology           | Text                | Poster  |
| About              | Text                | Poster of the Premiere in 2020-04-30 of "Fecundação e Alívio Neste Chão Irredutível Onde com Gozo Me Insurjo" |
| Author             | Person              | Hugo Santos   |
| File format        | Text                | pdf   |
| Date               | Date                | 2021  |
| Title              | Text                | n.a.  |
| Editor / Publisher | Person/Organisation | Paulo Costa/Nuisis Zobop  |
| Size               | Number              | 1024x512  |

### 3.3. Level 2 schema

The overall schema for Level 2 must accommodate different types of information about identified components or annotations, depending on the type of information that each annotation describes. For example, audio-related algorithms can identify that music starts at some specific time instance, or video-related algorithms can identify that the lighting conditions on stage have changed at a specific time instance; in both examples, there is no concept of "position on stage" for the identified events. In other examples, however, "position on stage" information may be meaningful, e.g., in the case where a video algorithm detects motion of a dancer or actor who moves from one place to another on stage. Additionally, the position of a dance or actor, object or event on stage, may be described in different ways that are meaningful for different applications. For instance, the position of an object on stage would be better to be described by a 2-dimensional or 3-dimensional bounding box, in the cases of video and 3D virtual theatre respectively. Similarly, the body of a dancer or actor would be better described by demarcating the exact shape of the part of body that is visible to the viewer (either on video or on a 3D virtual theatre setting).

Annotations on Level 2 are employed to both provide information to the user and allow advanced interactions. For example, in Use Case 1, where users are browsing the archives looking for a specific speech sentiment by a specific actor, Level 2 annotations will allow the identification of such segments. Regarding interaction, bounding boxes or dancer/actor outline demarcation will allow users to "click" on a dance/actor and retrieve information about them and details about their performance, e.g., in use case 2, in a live viewing setting

Level 2 schema comprises a hierarchy of "classes" that incorporate information on different levels of detail and with different descriptive qualities, depending on the kind of information



that is necessary to be described. The information is in turn described on a hierarchy of categories that provide different levels of detail about what an annotation describes. The goal for constructing this hierarchy was to standardize the information that the algorithms can extract and what is expected to be visualized to the user/viewer. This standardization allows parallel development across teams that work on different modalities and different parts of the outcome (e.g., algorithm development and information visualisation).

#### I. Vision

- a. Light conditions
  - i. Stage lights changed.
  - ii. Spotlight changed.
- b. Object
  - i. Object added.
  - ii. Object removed.
  - iii. Object moved.
  - iv. Object position
  - v. Object bounding box.
- c. Human
  - i. Human entered.
  - ii. Human exited.
  - iii. Human moved.
  - iv. Human action.
  - v. Dressing change.
  - vi. Human position.
  - vii. Human shape.
  - viii. Human bounding box.

#### II. Sound

- a. Music
  - i. Presence of music.
  - ii. Presence of singing voice.
  - iii. Presence of specific musical instrument.
  - iv. Tempo estimation.
  - v. Music style estimation.
  - vi. Music emotion estimation.
- b. Voice
  - i. Speech presence.
  - ii. Speech emotion / sentiment identification.
  - iii. Speaker identification.
- c. Audio Events
  - i. Event free-text description.
  - ii. Identification of event source.

#### III. Text

- a. Sentence recognition
  - i. Entity identification.
  - ii. Speaker of sentence identification.
- b. Subtitles
  - i. Subtitle content.
  - ii. Subtitle sentiment identification.
  - iii. Subtitle speaker identification.

Level 2 annotations are JSON-compatible. Next tables describe the “classes” that comprise Level 2 annotations, starting from the most basic type, a “TimedAnnotation”, which includes

only information about the category (based on the hierarchy provided above), content (JSON-compatible that provides details that are pertinent to the described event), start and end time and any other information that would be desirable to accompany an event (in JSON-compatible format). The hierarchy of classes is built upon the basic “TimedAnnotation” type; information that is inherited from super-class is shown in grey-coloured fonts.

The most basic type of events

```

TimedAnnotation: {
category: {
    top_level: STRING,
    middle_level: STRING,
    bottom_level: STRING
},
content: {
    CONTENT_JSON # should be defined for each annotation subcategory
},
start_time: INT (ms),
end_time: INT (ms),
other_info: {
    OTHER_INFO_JSON # should be defined for each annotation subcategory
}
}

```

Subclassing TimedAnnotation

```

TimedScreenPositionedAnnotation: {
category: {
    top_level: STRING,
    middle_level: STRING,
    bottom_level: STRING
},
content: {
    CONTENT_JSON
}
start_time: INT (ms),
end_time: INT (ms),
position_on_screen: {
    x: FLOAT,
    y: FLOAT
},
other_info: {
    OTHER_INFO_JSON
}
}

```

Subclassing TimedAnnotation

```

TimedStagePositionedAnnotation: {

```

```
category:{
    top_level: STRING,
    middle_level: STRING,
    bottom_level: STRING
},
content:{
    CONTENT_JSON
}
start_time: INT (ms),
end_time: INT (ms),
position_on_stage:{
    x: FLOAT,
    y: FLOAT,
    z: FLOAT
},
other_info:{
    OTHER_INFO_JSON
}
}
```

Subclassing TimedAnnotation

---

```
Timed2DBoundedBoxAnnotation:{
category:{
    top_level: STRING,
    middle_level: STRING,
    bottom_level: STRING
},
content:{
    CONTENT_JSON
}
start_time: INT (ms),
end_time: INT (ms),
box_position:{
    x_top_left: FLOAT,
    y_top_left: FLOAT,
    width: FLOAT,
    height: FLOAT,
},
other_info:{
    OTHER_INFO_JSON
}
}
```

Subclassing TimedAnnotation

---

```
Timed3DBoundedBoxAnnotation:{
category:{
    top_level: STRING,
```

```
        middle_level: STRING,
        bottom_level: STRING
    },
    content: {
        CONTENT_JSON
    }
    start_time: INT (ms),
    end_time: INT (ms),
    box_position: {
        x_front_top_left: FLOAT,
        y_front_top_left: FLOAT,
        z_front_top_left: FLOAT,
        width: FLOAT,
        height: FLOAT,
        depth: FLOAT,
    },
    other_info: {
        OTHER_INFO_JSON
    }
}
Subclassing TimedAnnotation
-----
Timed2DSkeletonAnnotation: {
category: {
    top_level: STRING,
    middle_level: STRING,
    bottom_level: STRING
},
content: {
    CONTENT_JSON
}
start_time: INT (ms),
end_time: INT (ms),
skeleton2D: {
    SKELETON_2D_CONTENT # COCO format
},
other_info: {
    OTHER_INFO_JSON
}
}
Subclassing TimedAnnotation
-----
Timed3DSkeletonAnnotation: {
category: {
    top_level: STRING,
    middle_level: STRING,
    bottom_level: STRING
```

```

},
content: {
  CONTENT_JSON
}
start_time: INT (ms),
end_time: INT (ms),
skeleton3D: {
  SKELETON_3D_CONTENT # COCO format
},
other_info: {
  OTHER_INFO_JSON
}
}

```

Subclassing TimedAnnotation

---

```

Timed2DShape: {
category: {
  top_level: STRING,
  middle_level: STRING,
  bottom_level: STRING
},
content: {
  CONTENT_JSON
}
start_time: INT (ms),
end_time: INT (ms),
shape2D: {
  SHAPE_2D_CONTENT # custom compressed form of grayscale PNG mask-image
},
other_info: {
  OTHER_INFO_JSON
}
}

```

### 3.4. Level 3 schema

“Level 3” annotations are quite similar with level 2, with the difference that they are specialised for either dance or theatre. As in the case of level 2 annotations, they are usually considered on a time basis and can be either computed by algorithms or manually annotated.

#### 3.4.1. Theatre schema

The Theater schema is, above all, related to the form of the performers' actions on stage. The voice and the body are the actor's main tools on the scene and supplemented with the text of the play they are the starting point for these annotations. Mainly in post-dramatic and

physical theatre, the movement of the performers provides us with a set of complementary information to the rest of the content of the scene that we consider important to highlight. The use of repetition, movement and immobilization, the tension of the bodies, the positioning of the bodies in the scene, all of this is relevant for the spectator and, many times, contemplated in the dramaturgy of the performance. The fact that those actions are almost always interrelated with the Director's choice on the essence of the performance, constitutes one more reason to be incorporated to Level 3 Theatre Schema. All basic actions such as walking from A to B or turning your face in a certain way, manipulating objects and scene props and the way this is done, are also important.

In theatre, the expression of voice also plays a fundamental role. Here we can include things like the way certain words are said, expressing a specific emotion, the repetition, this time not of a movement, but of a word or phrase, the silences, the shout or whisper, the cadence of the phrases, etc.

One of the biggest challenges of this work is the subjectivity of what we see on stage and the possibility of individual interpretation by the spectator, which, when defining these actions and associating them with emotions, for example, is lost.

In order to get around the point raised in the last paragraph, regarding the subjectivity of the scenes, we approach the creation of the Theater schema categories in a more concrete way and try to use less emotions, feelings or intentions in the description of the actions, unless they are clearly assumed on the scene or in the script's stage directions (that's why these are included in the scheme). This perspective, although it may initially seem limiting, allows us to maintain the subjectivity of the plays but, at the same time, make them more accessible to those who cannot see or hear them, for example, and add important and less obvious information that was also taken into account by the director/ playwright.

Below we provide an attempt to organize a taxonomy that will be further developed and refined after the upcoming workshops of PREMIERE partners with theatre experts such as directors, actors, scholars etc.

- Perspective
  - Internal Perspective (e.g., of a Performer)
    - Thinking (when a performer acts, which thought drives his/her actions)
    - Imagination (this refers to imaginary worlds or relationships that cross-cut the role shaping by both the director and the performer)
    - Intentions ( when a performer moves and speaks in a certain way, what is their intention)
    - Feelings (which specific feelings are expressed through the performer's movement and/or speech)
  - External Perspective (e.g., the Audience)
    - Actual Movements (What the audience actually sees in terms of movements on stage)
    - Positions on stage (Visually identifiable relations between performers, spatial arrangements or coordinated movements that are externally visible)
    - Actual voices, sounds and emotions (this refers to what the audience actually hears in terms of text and expressed emotions by the performers)
- Level of detail
  - Parts or Groups (How many performers act on stage)

- Qualities in Duets (activities or relations between two performers)
- Larger Groups (activities or relations between more than two performers)
- Scale
  - Small Movements (movements executed over minor distances in space and time)
  - Large Movements (movements executed over major distances in space and time)
  - Local Activities (single activities that are isolated in space and time)
  - Distributed Activities (single activities that traverse space and time or multiple activities that are distributed across space and time)
- Prosody: We will examine the possibility of including elements of prosody in the level 3 annotation schema for theatre which relate to the rhythm and melody of the voice including intonation, stress and pauses. This will be linked to the expression of feelings as it is closely coupled in the context of a theatrical performance. The goal is to move deeper in level 2 schema related aspects such as voice, sentence recognition and audio events.

### 3.4.2. Dance schema

The Dance schema mainly consists of movement qualities. These qualities are usually present in a time basis. During the first ten months of PREMIERE, a great effort has been dedicated by the dance specialist members of the consortium (ICK, STO, FD) to finding a common annotation schema for dance. This effort had the difficulty of the non-existence of a globally accepted dance movement/qualities vocabulary, or in cases that there is one (e.g, the Laban glossary), it is not complete in relation to contemporary developments in dance and needs additional specifications.

Although we aimed on working towards finding a common and globally accepted dance schema, this very optimistic goal requires much more time and effort that may fall out of the scope of PREMIERE. For the needs of the project, we decided to move on by defining a fragmented collection of dance annotation sub-schemas. This fragmentation reflects the current trends in the dance scientific community of a non-agreement, but on the other hand is descriptive enough for the sake of PREMIERE. We settled on three main fragments of ontologies. The first one consists of a general hierarchical structure of broader categories. The other two consist of specific movement qualities ontologies, a universal one based on Laban analysis, and a signature-specific one based on the research made by STO, ICK and FD.

#### 3.4.2.1. Dance qualities categories

The art of dance includes a wide array of artistic practices that work with bodily movement and choreographic decision making about the how, where and when of performing movements in front of an audience on a stage or other locations. In the contemporary dance landscape, there is no consensus but rather an extensive variety of different approaches to the definition of dance and choreography and to their ontological and possibly epistemological status as contemporary art forms and their cultural significance for society at large. Given the living moment bound appearance of dance and despite the widespread conception of dance as being impossible to document, many contemporary choreographers have developed approaches to record, analyse and share choreographic ideas, concepts and logics in co-creative dialogue with scholars from other disciplines and with developers of

innovative technologies. These approaches to document and transmit dance have focused on forms of logic that are corporeal, intrinsic to movement and movement creation and highly individualized, diverging from artist to artist<sup>1</sup>. These forms of movement logic are embodied and communicated through different movement qualities.

In co-creative dialogue with evolving new technologies to compute movement qualities as ways to preserve and transmit movement and movement knowledge, the definition of quantifiable descriptors of movement qualities is indispensable. In our approach to defining categories for dance qualities, we distinguish between descriptors that can be considered as “universal” (such as for example “body”, “effort”, “shape” and “space” according to LMA (Laban Movement Analysis) (see 3.4.2.2.) approach to and vision about dance of diverse choreographers, here exemplified by the bodily and verbal vocabularies of STOCOS and ICK (see 3.4.2.3). Evidently, the level of quantifiable measurability decreases significantly on the level of “signature specific” movement qualities given their subjective nature, articulated from the inner perspective of choreographers and dancers. This implies the necessity of additional manual annotations but delivers at the same time semantic enrichment of data (i.e., deepening the understanding of embodied knowledge) that can be of interest to artists, scholars, dance and theatre professionals as well as to larger audiences. Considering the general lack of signature specific vocabularies articulated from the inner, experiential perspectives of creators and performers and their significance for the epistemological potential of the performing arts, we consider the collection of “signature specific” descriptors as an asset for the PREMIERE project to be potentially further developed in other contexts. Negotiating the edges of measurability and standardization of movement qualities will be included in the ongoing research of PREMIERE, constituting a challenge for the further development of the technologies as well as to the development of annotation methodologies that can accommodate “universal” as well as “signature specific” movement qualities. We will explore further what kind of already existing conceptual framework for annotating and computing movement qualities<sup>2</sup> can be integrated into our proposal for a dance data annotation schema. Bridging the gap between the necessarily subjective and varying perspectives on movement qualities of dance experts and the collection of ground-truth datasets appears to be one major challenge which we hope to be able to address by building on previous research efforts done in this direction<sup>3</sup>.

Below we approximate a taxonomy as a departure point for a skeleton for a dance data annotation schema to further elaborate in dialogue with the consortium partners. This taxonomy seeks to give organic and logical continuity to Level 2 and is a first draft based on

---

<sup>1</sup> Scott deLahunta and Florian Jenett. “Making Digital Choreographic Objects Interrelate: A Focus on Coding Practices.” *Performing the Digital: Performance Studies and Performances in Digital Cultures*, edited by Martina Leeker et al., Transcript Verlag, 2017, pp. 63.

<sup>2</sup> For example, Caroline Larboulette and Sylvie Gibet, “A review of computable expressive descriptors of human motion. In Proceedings of the 2nd International Workshop on Movement and Computing (MOCO '15). Association for Computing Machinery, New York, NY, USA, 2015, 21–28. <https://doi.org/10.1145/2790994.2790999> and Antonio Camurri, Gualtiero Volpe, Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Nicola Ferrari, and Corrado Canepa, The Dancer in the Eye: Towards a Multi-Layered Computational Framework of Qualities in Movement. In Proceedings of the 3rd International Symposium on Movement and Computing (MOCO '16). Association for Computing Machinery, New York, NY, USA, Article 6, 2016, 1–7.

<sup>3</sup> Caroline Larboulette and Sylvie Gibet, “A review of computable expressive descriptors of human motion. In Proceedings of the 2nd International Workshop on Movement and Computing (MOCO '15). Association for Computing Machinery, New York, NY, USA, 2015, 21–28.



ongoing dialogues between ARC, STO, ICK, and CYENS. It is a first intuitive line up that will be developed further to the extent that the objectives and scope of PREMIERE project allows.

- I. Human activity (activities conducted by one or several human performers)
  - a. Human Action (discrete and functional activity such as opening a door)
  - b. Human Movement Intention (the inner disposition towards movement relating to its purpose and meaning)
  - c. Human Movement quality (the qualitative and expressive characteristics of body movements independent of its functional characteristics)
  - d. Group Interaction (relational activities among multiple performers)
- II. Human shape
  - a. Posture (the static overall configuration of a performer's body)
  - b. Alignment (correspondences between the sub-postures of body parts within or across a single body, multiple bodies, or between bodies and configurations in the environment)
  - c. Expansion (the compactness or extendedness of a human body, the increase of volume of the body i.e., 3D distance between the joints)
  - d. Contour (the silhouette or outline of a human body)
  - e. Kinesphere (movement space, the directionality and orientation of a performer's body in relation to the surrounding space that is reachable through extended limbs in all directions and on all levels)
- III. Human perspective
  - a. Internal Perspective (e.g., as seen from the subjective perspective of a performer)
    - i. Thinking (processing of information through the body-mind, conscious and deliberate choice making in movement)
    - ii. Imagination (related to utopian and visionary bodies, transcending corporeal limitations, moving through imagined worlds, experiencing imagined sensations)
    - iii. Intentions (related to the purpose and meaning of movement, inner disposition towards movement and action)
    - iv. Affect (the emotional state of the performer as experienced through the moving body, personal and impersonal)
    - v. Proprioception (the inner sensation and awareness of the position and movement of the body)
    - vi. Relationality/Resonance (to other performers, to the audience, to the space)
  - b. External Perspective (e.g., the Public)
    - i. Potential Movements (preparation, suspense, and shift in attention towards a future movement)
    - ii. Actual Movements (executed movements that become externally visible)
    - iii. Actual Poses (postures that are assumed and become externally visible)
    - iv. Positions on Stage (spatial positions that are taken and become externally visible)
    - v. Visually identifiable relations between performers (spatial arrangements or coordinated movements among performers that are externally visible)
- IV. Level of Detail
  - a. Parts or Groups

- i. Body Parts (activities or shapes involving only individual body parts)
  - ii. Full Bodies (activities or shapes involving the entire body)
  - iii. Duets (activities or relations between two performers)
  - iv. Larger Groups (activities or relations between more than two performers)
- b. Scale
  - i. Small Movements (movements executed over minor distances in space and time)
  - ii. Large Movements (movements executed over major distances in space and time)
  - iii. Local Activities (single activities that are isolated in space and time)
  - iv. Distributed Activities (single activities that traverse space and time or multiple activities that are distributed across space and time)

### 3.4.2.2. *Universals*

For *universal* dance qualities we work with the conceptual framework of LMA (Laban Movement Analysis). LMA semantically categorizes structural, dynamic and geometric properties of human motion in four main components: Body, Space, Shape and Effort. The Body and Space components refer to how the body moves, either within the body or in relation to the surrounding space. The Shape component refers to the morphology of the body engaged in motion. The Effort descriptor refers to qualitative aspects of movement that are categorized according to intention, dynamics and energy. LMA is widely used for movement indexing, segmentation and computer animation. It has also been explored in relation to affective expressivity yet there still lacks consistency in terms of quantifiable sets of motion measures that can be worked with as standardized data set for computing movement and for movement research across diverging artistic/choreographic movement language signatures<sup>4</sup>.

#### LMA (Laban Movement Analysis Categories):

- I. Body
  - a. Movement Initiation
  - b. Connection between different body parts
  - c. Movement Sequence
  - d. Neuromuscular Patterns
- II. Effort
  - a. Space (direct/indirect)
  - b. Weight (strong/light)
  - c. Time (sudden/sustained)
  - d. Flow (bound/free)
- III. Shape
  - a. Shape forms
  - b. Modes of Shape Change
  - c. Shape qualities
  - d. Shape flow support
- IV. Space
  - a. Kinesphere

---

<sup>4</sup> Caroline Larboulette and Sylvie Gibet, "A review of computable expressive descriptors of human motion," in Proceedings of the 2nd International Workshop on Movement and Computing (MOCO '15). Association for Computing Machinery, New York, NY, USA, 2015, 21–28. <https://doi.org/10.1145/2790994.2790998>

- b. Spatial Intention
- c. Geometry

### 3.4.2.3. Signature specific

The term *signature specific* stands for both physical and verbal movement vocabularies and qualities developed by choreographers that reflect a specific physical, mental, aesthetic, philosophical and ontological disposition towards dance and choreography. They define idiosyncratic coordinates for working with body and movement in time and in space. In PREMIERE we work with two signature specific ontologies originating from STO and ICK respectively. These two sub-ontologies provide a specialised and distinct point of reference for the dance qualities. The reason why we consider these sub-ontologies, is that we have a deep understanding of these vocabularies and can record examples for these qualities and thus analyse specific characteristics of these movements with an algorithmic manner, e.g. **T3.2 - Human body pose and motion analysis**. Annotating 3D Motion Capture recordings with these *signature specific* vocabularies can serve as examples and case studies for enriching the semantic content of performing art artefacts with insider knowledge developed from the experiential perspective of the artists and performers and thus making this knowledge accessible to other artists, educators and scholars as well as to larger audiences.

#### ICK movement vocabulary:

The ICK ABCdaire is a living archive with words, definitions, descriptions, categories, metaphors and typologies based on more than twenty years of notation and documentation research into the dance language of Emilio Greco and Pieter C. Scholten. Instead of choosing a symbol-based dance notation, the choice has been for finding appropriate words to describe the inner/outer landscapes of physical-mental states that are characteristic to their specific vision on dance: “The body has to be clear, and the words have to be right.” The ICK ABCdaire consists of ca. 200 alphabetically arranged terms that are divided into different categories: basic, workshop, Pre-Choreographic Elements, statements, mechanism. The ICK ABCdaire is used to re-stage repertoire, to research new movements, to transfer expert knowledge to next generations of dancers and to inspire dancers and choreographers in training and at ICK Artist Space with a specific artistic logic derived from the dancing body as protagonist and ground truth. The ICK ABCdaire can also be used for the annotation of creative processes. It has been the starting point for interdisciplinary dialogues with creative coders, philosophers, neuroscientists, linguists, art and performance theorists, visual artists and composers.<sup>5</sup>

In the category “workshop”, basic principles of the dance method [Double Skin/Double Mind](#) are described to support knowledge transfer on the level of movement intentionality which influences movement quality. For instance, [Armpit/Shoulder Breathing](#) is a principle that is related to the area of the shoulders and to the breath. Working with this movement principle causes a specific movement quality to emerge. From the inner, creator and performer perspective, this exercise is for example described as: “by breathing in and out in the region of the armpit you expose, you enlighten and manifest a hidden part of the body. The

<sup>5</sup> See for example Scott deLahunta, *(Capturing Intention) Documentation, Analysis and Notation Research Based on the Work of Emilio Greco* | PC, Amsterdam: Emilio Greco | PC and Amsterdam Hogeschool voor de Kunsten, 2007; “Harmonic Dissonance,” ICKamsterdam.com <https://www.ickamsterdam.com/nl/academy/producties/harmonic-dissonance-100>; “Annotation and Documentation: Expressing the Unsayable”, ICKamsterdam.com <https://www.ickamsterdam.com/en/academy/peers-researchers/annotation-and-documentation-47> .

shoulders are not considered to be a border or a frame of the body, but an open gate that allows to receive new inputs and diverse information, such as waves of light, energy and temperature.” The purpose of these kind of descriptions is multilayered: a) learning to verbally describe movement intentions that form the basis for movement qualities creates a bridge between the body and consciousness, allowing the dancer to become more aware of the non-verbal, intentional and qualitative dimension of corporeal languaging; b) the verbal description can serve as an educational and didactic tool to transmit dance qualities to other dancers; c) the verbal description can be used to annotate recordings of the movements and to deliver insider knowledge to other artists and general audiences. We consider PREMIERE as an ideal context to further explore the realizability of these layers as well as to discover new layers of meaning and potential application in relation to society at large.



Figure 2 - Emio Greco ICK Double Skin/Double Mind movement method, “Shoulder Breathing”.

### **STO movement vocabulary:**

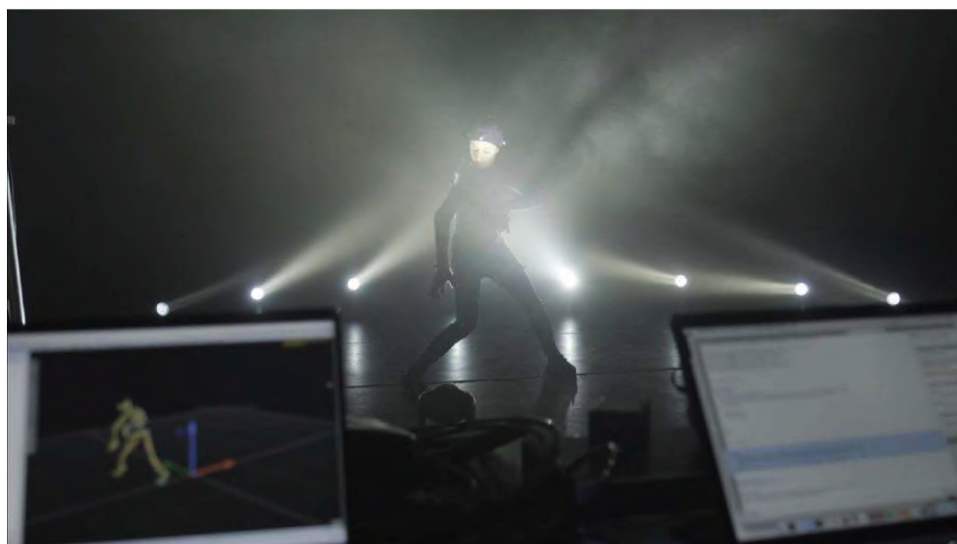
In the last fifteen years, Muriel Romero, choreographer of Instituto Stocos has developed a movement vocabulary evolving from the research activities and artistic productions of the company. Being the core activity of Instituto Stocos as a dance and technology company the cross fertilisation between art and science, the approaches to movement creation sometimes employ generative approaches to choreography and are also influenced by abstractions taken from other disciplines, sometimes translating these phenomena into movement strategies. Some of these borrowed abstractions that contributed to shape Muriel Romero’s vocabulary are actually computer models of natural phenomena that have a choreographic nature such as computer swarm simulations, stochastic processes such as *Brownian* movement or artificial neural networks.

The movement qualities she works with form part of a vocabulary that she is continually adjusting and extending based on her own preferences, the particular pre-dispositions of the dancers she works with, and the requirements of new dance productions. This subset of movement qualities has been organised into five groups, named Dynamics, Body, Shape,

Space, Time. The category "Dynamics" contains movement qualities that deal with the temporal evolution of movement, focusing on changes in velocity and directionality. Examples are the movement qualities Fluidity, Particles, Staccato, and Thrusting. The category "Body" contains movement qualities that deal with the juxtaposition of movements and poses within a single body or across multiple bodies. Examples are the movement qualities Breathing, Relations, Conversation, Polytopia, Contrast, and Paralyze. The category "Shape" contains movement qualities that deal with the transformation of movements or poses and the alteration or preservation of their recognisability during a transformation. Examples are the movement qualities Change of Plane, Transformation, and Repetition Transformation. The category "Space" contains movement qualities that deal with how movement traverses space or is located at positions within space. Examples are the movement qualities Levitation, Spatial Recovery, Volume, Out of Balance. The category "Time" contains movement qualities that deal with the temporal organisation of movement or the absence of movement. Examples are the movement qualities Silence, Continuity, Rigidity, Slowness, Accumulation. A more complete subset of these movement qualities has been documented through motion capture and can be found here<sup>6</sup>:

An indicative example of this vocabulary is *Polytopia*, which dissociates the body into multiple independent body parts, each performing different movement qualities and rhythms. From the performer's perspective, it is a multitude of different sensation depending on the quality that each part of the body moves. Another example is *Breathing*, which corresponds to a quality which is like the body has air inside, while from the performer's perspective it gives the sensation that the *Breathing* moves the whole body. A third one *Particles* Body parts float while constantly changing direction. This quality is inspired by *Brownian* Movement.

At the same time these qualities become the basis of cross modal approaches to artistic creation and have been translated into artificial simulations that embody these movement aspects, becoming virtual entities that are translated into robotic light movement choreographies or sonic entities that are spatially choreographed. Examples of this are documented in this Motion Bank link<sup>7</sup>



<sup>6</sup> <https://pureportal.coventry.ac.uk/en/datasets/e2-create-instituto-stocos-dataset>

<sup>7</sup> <https://embodied-machine.motionbank.org/#!/>



Figure 3 – Instituto Stocos Motion Capture Session.

## 4. Data processing

### 4.1. Audio and text related data

This section concerns the information that is indexed with audio-related tools, from the audio that is retrieved from an archived video (in the case of use case 1). In use cases that require real-time processing, the same techniques are applied to buffered data. Data processing of audio extracts features that are related to music, voice and audio events, as described in detail in Section 3.3.

The audio modality also provides information about text-related data. Text is extracted from speech audio using Automatic Speech Recognition models. To increase the accuracy of speech-to-text transcription, Voice Activity Detection is first applied for removing long pauses of speech or the presence of music. Given a clear transcription, the features described for the text modality in Section 3.3 are extracted. Entities that are identified in the text modality, are combined with metadata information and are used to retrieve information from online resources (e.g., Wikipedia). An overview of the audio and text data processing is given as follows:

- **The audio track** is retrieved from the video stream.
- **Voice, music and audio event activity detection** is performed.
  - **Music** features are extracted, as analysed in Section 3.3.
  - **Audio events** are identified and characterised. For each audio event, starting and ending timestamps are estimated along with a category tag, among around 500 categories.
  - **Voice** processed as follows:
    - Speech is segmented and characterised according to **sentiment and emotion**.
    - It is transcribed to **text** using Automatic Speech Recognition (ASR).
- **The transcribed text** is then processed as follows:
  - According to the transcription timestamps, **subtitles** are generated.
  - The **speaker** of each subtitle part is identified.
  - **Emotion** based on text and the context is attributed to each subtitle.
  - **Entities** are recognised and named from text and the context. Metadata that accompany the video file are also analysed to relate identified entities with entities that are described in the metadata file.
  - **Active URLs** to Wikipedia resources are automatically generated from identified entities and metadata (e.g., participating dancers, actors, identified locations, venue information, etc).

### 4.2. Video related data

Data processing of video extracts features that are related to frame (image) and multiple frames events, as described in detail in Section 3.3. This video needs to be split into sequences which break down the video into meaningful units, allowing for efficient organization, indexing, and analysis of video data. It is then possible to break down these

scenes into image sequences (frames) to automatically extract low level features that are then used to determine high level features.

An overview of the video data processing is given as follows:

- **Light conditions:**
  - Are initially analyzed on a frame basis and based on the following low-level features:
    - light color estimation.
    - light source detection.
  - Based on a sequence of frames these low-level features allow to detect the following high-level features and the corresponding data:
    - Stage lights changed.
    - Spotlight changed.
- **Objects**
  - Based on objects' estimation we can extract the following low-level features:
    - Object detection which refers to the process of automatically identifying and localizing specific objects. This generates a mask (set of pixels) for each object detected.
    - Object position which refers to the spatial coordinates or location of a specific object within each frame of a video sequence.
    - Depth maps, which refer to the estimation of the distance or depth information of objects in the scene captured by a single camera. It involves inferring the 3D structure of the scene from the 2D visual input.
  - These three features allow us to determine on a set of frames the following high-level features and the corresponding data:
    - Object added.
    - Object removed.
    - Object moved.
    - Object 2D/3D position.
    - Object 2D/3D bounding box.
- **Humans**
  - Human poses are analysed on a frame basis and based on the following low-level features:
    - Human body segmentation identifies and separates each pixel that makes up the different bodies in a frame. This generates a mask (set of pixels) for each body detected.
    - Human 2D/3D pose (with face and hands) estimation refers to the process of inferring the spatial configuration and movement of a person. It involves detecting and tracking key body joints, facial landmarks, and hand keypoints and shapes to reconstruct the 2D or 3D pose information.
    - Human 2D/3D position tracking refers to the process of continuously estimating and tracking the spatial coordinates or location of a person in a frame sequence.
  - These features allow us to determine on a set of frames the following high-level features and the corresponding data:
    - Human entered, human exited, human moved.
    - Human action, dressing change.

- Human 2D/3D position, human 3D shape, human 2D/3D bounding box.
- The posture, the alignment of joints, the angles of joints, the rotations of joints, the distance between the joints, the size and the skin contour data can easily be extracted from the human 3D shape.

## 5. Data annotation

### 5.1. Annotation software

For manually or semi-automatically annotating the data, we will deploy one or more open source available tools. These should be integrated in some way with the CMS, so that these annotations tools can be accessed by the annotators. Table 1 provides a list of existing annotation tools as candidates.

Table 1. List of considered data annotation tools.

| Tool         | Licence                | Input data  | Comments   |
|--------------|------------------------|---|--|
| Piecemaker   | Open source            | Video / MoCAP   | Only free text.  |
| ELAN         | Open source            | Annotation tool for audio and video recordings  | Annotation can be made on multiple layers called 'tiers'.  |
| FrameTrail   | Open source, Open code | Create, Annotate & Remix Interactive Videos   | <a href="https://frametrail.org/">https://frametrail.org/</a><br>It's a web based version. It can be useful for Theatre. |
| TagLab       | Open Source            | Semantic Segmentation. Has AI models that you can refine.                                 | <a href="https://taglab.isti.cnr.it/">https://taglab.isti.cnr.it/</a>  |
| Vidat        | Open source            | Supports video annotation, focus on machine learning for computer vision, runs in Browser | <a href="https://github.com/anucvml/vidat">GitHub - anucvml/vidat: Video Annotation Tool</a>                             |
| Label Studio |                        | Supports video/audio/text annotation  | <a href="https://labelstud.io/">https://labelstud.io/</a>  |
| LabelCloud   |                        | 3D cloud-point annotation   | <a href="https://github.com/ch-sa/labelCloud">https://github.com/ch-sa/labelCloud</a>                                    |

For the specific use case, we selected Label Studio as the baseline. Label Studio is a versatile and powerful data annotation software designed to streamline the process of creating high-quality labelled datasets for machine learning projects. Developed by Heartex, Label Studio has gained popularity due to its user-friendly interface, scalability, and extensive annotation capabilities.

At its core, Label Studio provides a collaborative platform that enables teams to efficiently label various types of data, such as images, text, audio, and video. It supports a wide range of annotation tasks, including object detection, image segmentation, named entity recognition, sentiment analysis, and more. This flexibility makes it suitable for diverse industries, from computer vision and natural language processing to audio processing and beyond.



The software's key features include the ability to customize annotation interfaces to suit specific project requirements, making it easy to create tailored labelling workflows. Its support for automatic pre-labelling and active learning further optimizes the labelling process, reducing manual effort and increasing overall annotation efficiency.

Label Studio promotes collaboration by allowing teams to work simultaneously on a project, track progress, and review annotations through an intuitive web-based interface. Additionally, it offers integration with popular machine learning frameworks like TensorFlow, PyTorch, and spaCy, facilitating smooth integration of annotated data into model training pipelines.

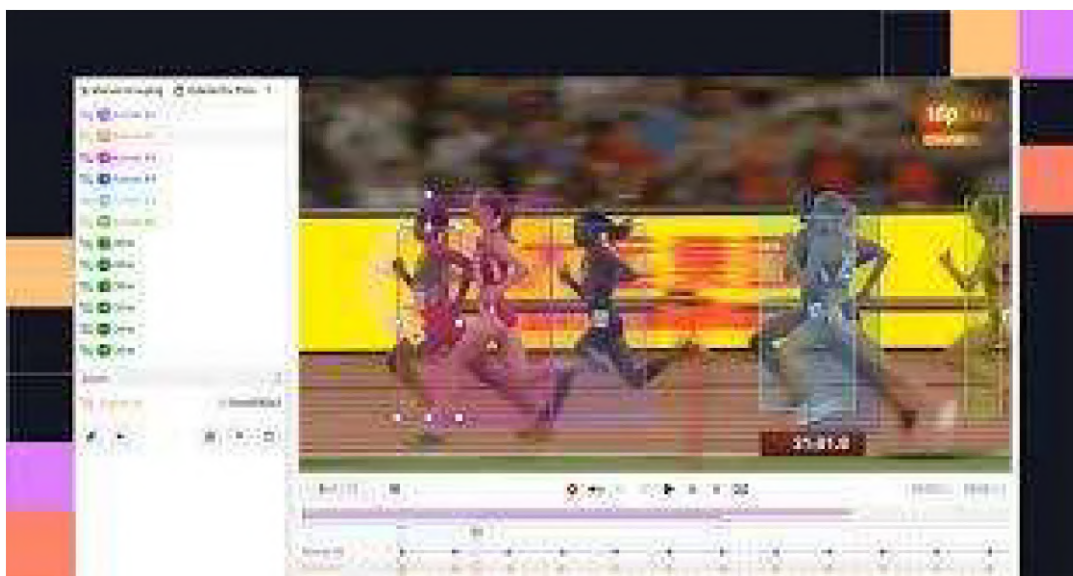


Figure 4 – Video annotation using Label Studio

## 5.2. Audio annotation

Label Studio enables the visualization of audio as a waveform and allows for further interaction with the user-annotator. It has all the basic features of a media player (play/stop buttons, an interactive progress bar) and additional functionalities such as timeline display, control of playback speed, x- and y-axis zooming.

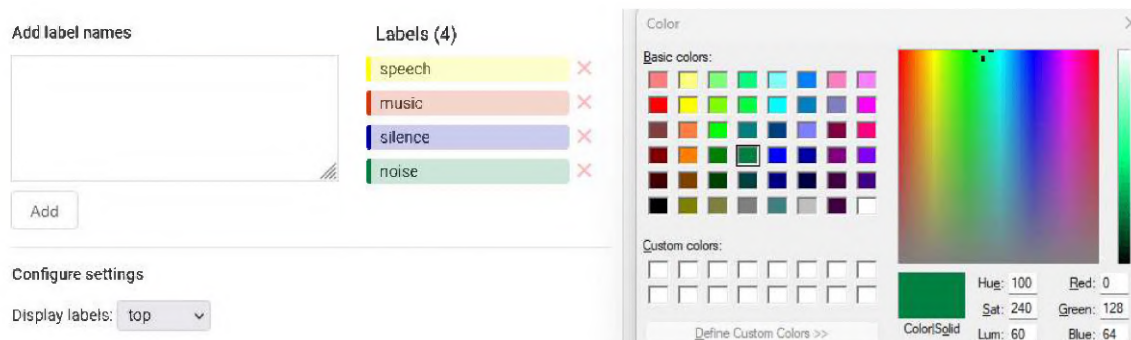
Concerning Label Studio's distinct annotation features for audio, the user can choose and customize the interface that best matches the requirements annotation project at hand.

To indicate the occurrence of audio events, one can draw waveform horizontal regions on the axis of time using the mouse. The time regions can be resized or shifted at any moment and the corresponding audio segment can be played separately with the standard 'play' button if the region is clicked beforehand.

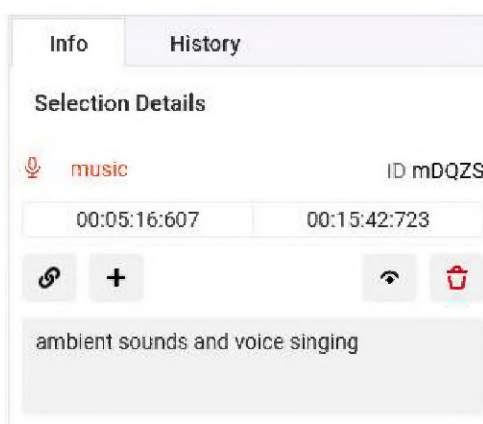


A set of labels is defined when configuring the labelling interface. Any region can be attached to a label of the annotator's choice. Each label is distinguished by a specific colour chosen

manually from a colour palette, and every region acquires the colour of its matching label with reduced opacity.



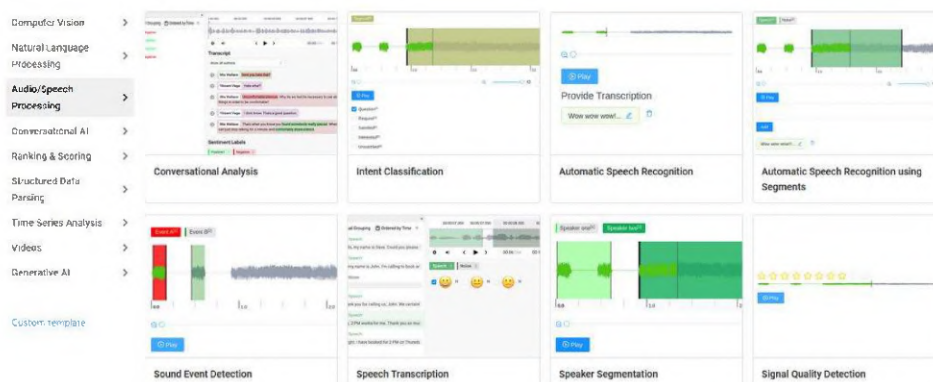
By clicking on a region, an interactive window appears on the right of the screen, enabling its deletion or hiding, and additionally the inclusion of “meta” information in the form of text. The definition of relations between regions is available, too.



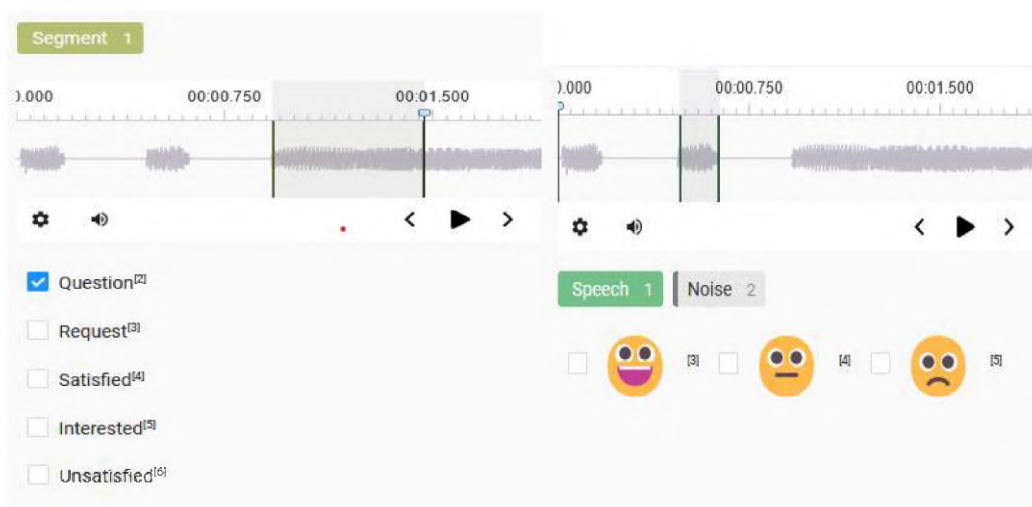
The interface can also be customized by intervening directly in the HTML code. This is the code for the interface defined above:

```
<View>
  <Header value="Sound Events Anndotation"/>
  <Labels name="tricks" toName="audio" choice="multiple">
    <Label value="speech" background="#ffff00"/>
    <Label value="music" background="#D4380D"/>
    <Label value="silence" background="#0000a0"/>
    <Label value="noise" background="#008040"/></Labels>
  <Audio name="audio" value="$audio_url" sync="video" speed="false"/>
</View>
```

Browsing interface templates provided by Label Studio can help choose the right functionalities. Annotators can choose to use a standard template as a starting point, further customizable using HTML code.



Available are also UI features for tasks such as intent classification, emotion recognition, speech transcription, and more.



Once created, the annotations can be submitted (i.e., saved) for each “Task” (see, the name that Label Studio used for the media file that requires annotation as part of a “Project”). The annotations can be displayed and exported in JSON format. Below is an example for emotion recognition with one region annotated:

```

"result": [
  {
    "id": "b4WMK",
    "meta": {
      "text": [
        "male speaking"
      ]
    },
    "type": "labels",
    "value": {
      "end": 4.816179513899323,
      "start": 1.5336437558226912,
      "labels": [
        "Speech"
      ],
      "channel": 0
    }
  }
]

```

```

    },
    "origin": "manual",
    "to_name": "audio",
    "from_name": "label",
    "original_length": 35.811927
  },
  {
    "id": "b4WMK",
    "meta": {
      "text": [
        "male speaking"
      ]
    },
    "type": "choices",
    "value": {
      "end": 4.816179513899323,
      "start": 1.5336437558226912,
      "channel": 0,
      "choices": [
        "Negative"
      ]
    },
    "origin": "manual",
    "to_name": "audio",
    "from_name": "sentiment",
    "original_length": 35.811927
  }
],
...

```

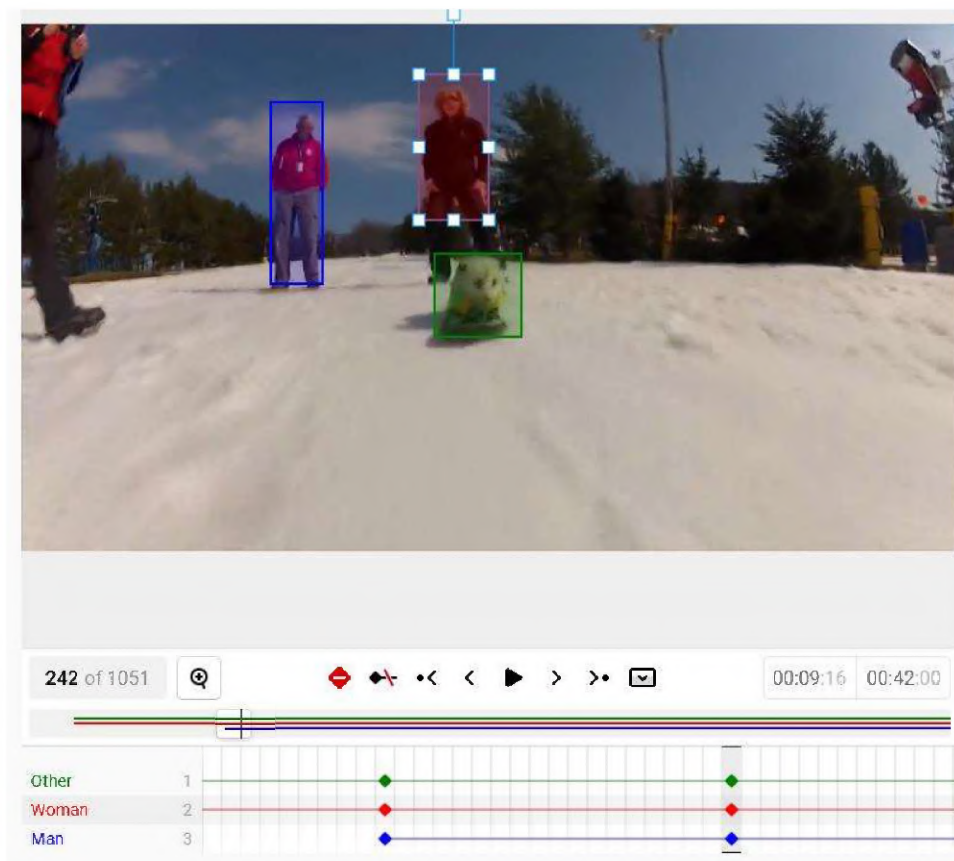
Label Studio ML backend can be used for automatic pre-labelling. To provide automatic pre-annotations for the task above, a JSON of the same structure should be returned by the ML backend that embeds a custom or pre-trained ML model of our choice. More details about the deployment of this SDK can be found online<sup>8</sup>.

### 5.3. Video annotation

As for the audio modality, Label Studio enables video handling by providing a media player. Probably the most important tool for video annotation is object tracking using bounding boxes that are defined in the same way as for the audio modality described above. They are created using the mouse and can be resized and moved in various positions on the 2D plane.

While such annotations can be reconfigured frame by frame, Label Studio enables a looser annotation technique that relies on making changes only when a significant alteration of an object's position is spotted. The annotation changes are marked with dots on a frame-level zoomed progress bar to enable better supervision and control.

<sup>8</sup> <https://labelstud.io/guide/ml.html>



Below we can see a JSON segment that corresponds to a bounding box at a current time:

```
...
  "result": [
    {
      "id": "DLV5mIdzYl",
      "type": "videorectangle",
      "value": {
        "labels": [
          "woman"
        ],
        "duration": 944.853333,
        "sequence": [
          {
            "x": 61.203125000000009,
            "y": 36.500000000000005,
            "time": 0.041666666666666664,
            "frame": 1,
            "width": 6.65624999999999165,
            "height": 13.8333333333333297,
            "enabled": true,
            "rotation": 0
          }
        ],
        "framesCount": 22677
      }
    }
  ]
}
```



```
    },  
    "origin": "manual",  
    "to_name": "video",  
    "from_name": "box"  
  },  
  ...
```

Moreover, Label Studio permits the inclusion of both audio and video interfaces in the same window (see, one above the other). This makes it possible for a parallel and multi-modal informed annotation procedure.

## 6. Recommendations for the Content Management System

### 6.1. Recommendation for the CMS backend

We propose a series of recommendations for the CMS backend based on the contents of this deliverable. These should be taken into account for the development of the CMS, in conjunction with the user requirements presented in Deliverable D2.1 - User Requirements v1.

- Data indexing and representation:
  - All items of the CMS should be represented by an item belonging to the ontology described in Section 3.
  - All items should have a unique identifier
  - The CMS should contain all the information between relations of items (e.g., Person A belongs to Organisation B Person C is the director of the play D).
- Data storing and access
  - Different data items should be potentially stored to different locations that are accessible through a URI. This allows for a scalable and distributed data storage workload.
  - Similarly, data and the extracted metadata (e.g., a video, along with the extracted features such as performers position).
  - The CMS should be able to handle permissions and access content layers
- Compatibility and Interoperability
  - Data formats stored in the CMS should be compatible with the data formats used in the annotation tool (Section 5)
  - Data stored in the CMS should be compatible and linked to the various algorithms from WP3 and WP4.


### 6.2. Recommendation for the CMS front-end

The main recommendation for the front-end is that the data should be visualised in a uniform, user friendly and coherent way. Moreover, these interfaces should be in line with back-end recommendations mentioned before. For example, they should allow to correct to some extent and add/remove new annotations from all three levels of the data annotation schema. It is recommendable, that for each of the main data types of the Level 1 schema, we have a dedicated template page to visualise this data. Below we propose some example templates. These templates should be defined in detailed in the 2<sup>nd</sup> version of the deliverable User

Requirements and the D6.2. Underlined fields correspond to links to other items represented in the CMS.

**Person**

## André Lepecki




**Bio**

André Lepecki works and researches at the intersection of critical dance studies, curatorial practice, performance theory, contemporary dance and visual ...

|               |   |
|---------------|---|
| Artistic Name | André Lepecki   |
| Nationality   | Brazilian   |
| Gender        | Male  |
| Birth date    | 1965-07-15  |
| Death date    | n.a.  |
| Family name   | Lepecki   |
| Given name    | André   |
| Affiliation   | <u>New York University – Department of Performing Studies</u> |
| Contact point | At11@nyu.edu  |
| Member of     | <u>Pós d' Arte</u>  |

**Photos**



**Articles**

Article 1 - .....  
 Article 2 - .....  
 Article 3 - .....  
 Article 4 - .....

**Cast and Credits**

## Cast and Credits

**Artistic**

|                  |                                       |
|------------------|---------------------------------------|
| Performers       | <u>Andy Design, Emio Greco</u>        |
| Director         | <u>Pieter C. Scholten, Emio Greco</u> |
| Scriptwriter     | n.a.                                  |
| Choreographer    | <u>Emio Greco, Pieter C. Scholten</u> |
| Music director   | n.a.                                  |
| Set designer     | <u>Emio Greco, Pieter C. Scholten</u> |
| Costume designer | <u>Clifford Portier</u>               |
| Light designer   | <u>Henk Danner</u>                    |
| Sound designer   | <u>Emio Greco, Pieter C. Scholten</u> |


**Technical**

|                     |  |
|---------------------|--|
| Producer            | <u>Emio Greco, Stichting Zwaanproducties</u> |
| Assistants          | n.a.   |
| Audio technician    | n.a.   |
| Author of recording | <u>Erik Lint</u>                             |
| Stage manager       | <u>Paul Beumer</u>                           |


**Credits**

|                  |  |
|------------------|--|
| Funding          | <u>Amsterdams Fonds voor de Kunsten Springdance Kaaatheater Tanz</u>     |
| Organizer        | <u>Stichting Zwaanproducties</u>   |
| Sponsor          | <u>Kaaatheater, Springdance, Tanzwerkstatt Berlin, Klapstuk Festival</u> |
| Copyright holder | EG   PC  |

**Performances**



**Photos**




**Articles**

Article 1 - .....  
 Article 2 - .....  
 Article 3 - .....  
 Article 4 - .....

**Performance**



## Noite de Reis




Year: 1998  
 Creator: [Ricardo Pais](#)  
 Country: Portugal

**About**

"Noite de Reis" is a comedy about love. In the kingdom of Ilyria, Duke Orsino is in love with Olivia, who does not love him. A young woman, Viola, arrives in Ilyria, washed away by the sea after a shipwreck. She has a twin brother, Sebastião, who she believes drowned in the shipwreck. Viola disguises herself as a man, changes her name to Cesário and finds work as a messenger for Orsino. Viola's job is to send love messages from Orsino to Olivia. Olivia falls in love with Cesário (Viola), thinking she is a man [...]

**Media**



**Cast and Credits**  
[Click to open](#)

**Articles**

[Article 1](#) - .....

[Article 2](#) - .....

[Article 3](#) - .....

## 7. Conclusions

This deliverable we described the data indexing process for the 1<sup>st</sup> PREMIERE use case T6.2 - **Performing Arts Archives browsing enhanced with AR/VR technologies**. The data considered in this use case was selected using several criteria so that they match PREMIERE objectives and consists of 11 theatre and dance pieces coming mainly from FIT and FD archives. These pieces are expected to be enriched by more pieces until the end the project. We used a relatively simple but descriptive ontology to describe this data in three levels of annotation granularity, that represent different annotation details. Starting by bibliographic indexing based on an ontology based on Schema.org, to the within the media annotation schema (level 2), until a more specific annotation schema incorporating domain knowledge from dance and theatre (level 3).

These data will be annotated using algorithms from WP3 and WP4, semi-automated annotations coming from Label Studio, as well as manual annotations from experts once that interface of the CMS is finalised. These data will be available through the use case 1 pilot to all PREMIERE user groups and stakeholders.