

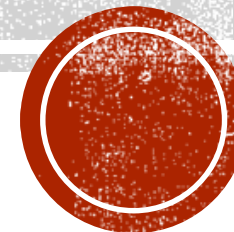
CultMEDIA

CultMEDIA “Machine learning-based services for harvesting multimedia documents to support low-cost video post-production and cross-media storytelling

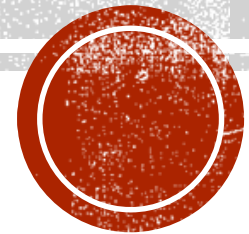
Italian Cluster TICHE Technologies for Cultural Heritage

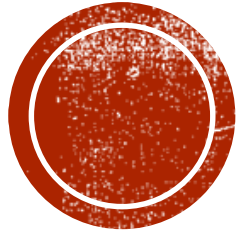


Ministero dell'Istruzione, dell'Università e della Ricerca



AI for DH Humanity





	Unit	People
UNISOB	Università Suor Orsola Benincasa CIRA Scienza Nuova	Roberto Montanari, Roberta Presta, Barbara Bagli, Alfredo Apicella
UNIMORE	Università di Modena e Reggio Emilia- Softtech-ICT (AImageLab)	Rita Cucchiara, Lorenzo Baraldi, Marcella Cornia,..
UNIFI	Università di Firenze- MICC	Alberto Del Bimbo, Lorenzo Seidenari,..
CNR	Consiglio Nazionale delle Ricerche IBAM (IT Lab, LAIM), ISTI (Visusal Computing), ISTI (NEMIS)	Daniela Malfitana, Francesco Gabellone, Roberto Scopigno,
Pdi, Pikkart	Progetti di Impresa-Pikkart srl	Marco Manfredi,
ETT	ETT spa	Giovanni Verreschi, ..

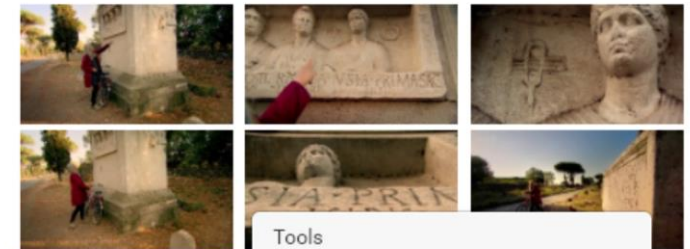
Partnership

VIDEO

- **One MODALITY of access : the VIDEO**
 - *video is becoming the principal component for knowledge communication and sharing*
- **MULTIMEDIA data :**
 - Video, Images,
 - Digitized Documents,
 - 3D Clouds, Computer Graphics products
 - ... Sounds, Text, Speech..
- **Goals**
- **High Quality & Low Cost Multimedia Production**
 - for re-using existing material
 - for integrating multimedia data in cross-media storytelling
 - To new business models in multimedia production
 - To boost diffusion of Cultural Heritage knowledge

Scene 7 copy

Shots



Speaker transcript

Tools

Caption: A woman in a red dress is standing in a field.

[Seek to this point](#)

This lady's Usia Prima, a priestess of the Egyptian goddess Isis, and there's her little sacred rattle. She's almost looking at you. I feel like saying, "Pleased to meet you, Prima."



VIDEO



the Youtube users watch more than 3 billion hours in each month

from the economical point of view ,
1 minute of video corresponds to
1.8 million of words.
(Insivia, 2016)



the 60% of Web users stop watching a video after 2 minutes

the 20% clicks away after 10 seconds.
(Visible Measure,)

The CultMEDIA project aims at facilitating the developing of audio-visual and transmedia storytelling by optimizing costs and complexity of cultural media production.



GOAL



- *A disruptive improvement in the processes and services related to the cultural heritage content production,*
- **a) handling the creation of multimedia video** (i.e. containing text, audio, images, clips, 3D animations) **and new transmedia storytelling** (i.e. exploiting different media platforms, from video to web, to augmented reality, to improve user experience), **thanks to the creation of a service platform and open-source software tools;**
- **b) providing large cost savings** through the extended use of **machine learning and artificial intelligence solutions** for the reuse of existing multimedia material and its integration in new CH productions.



GOAL



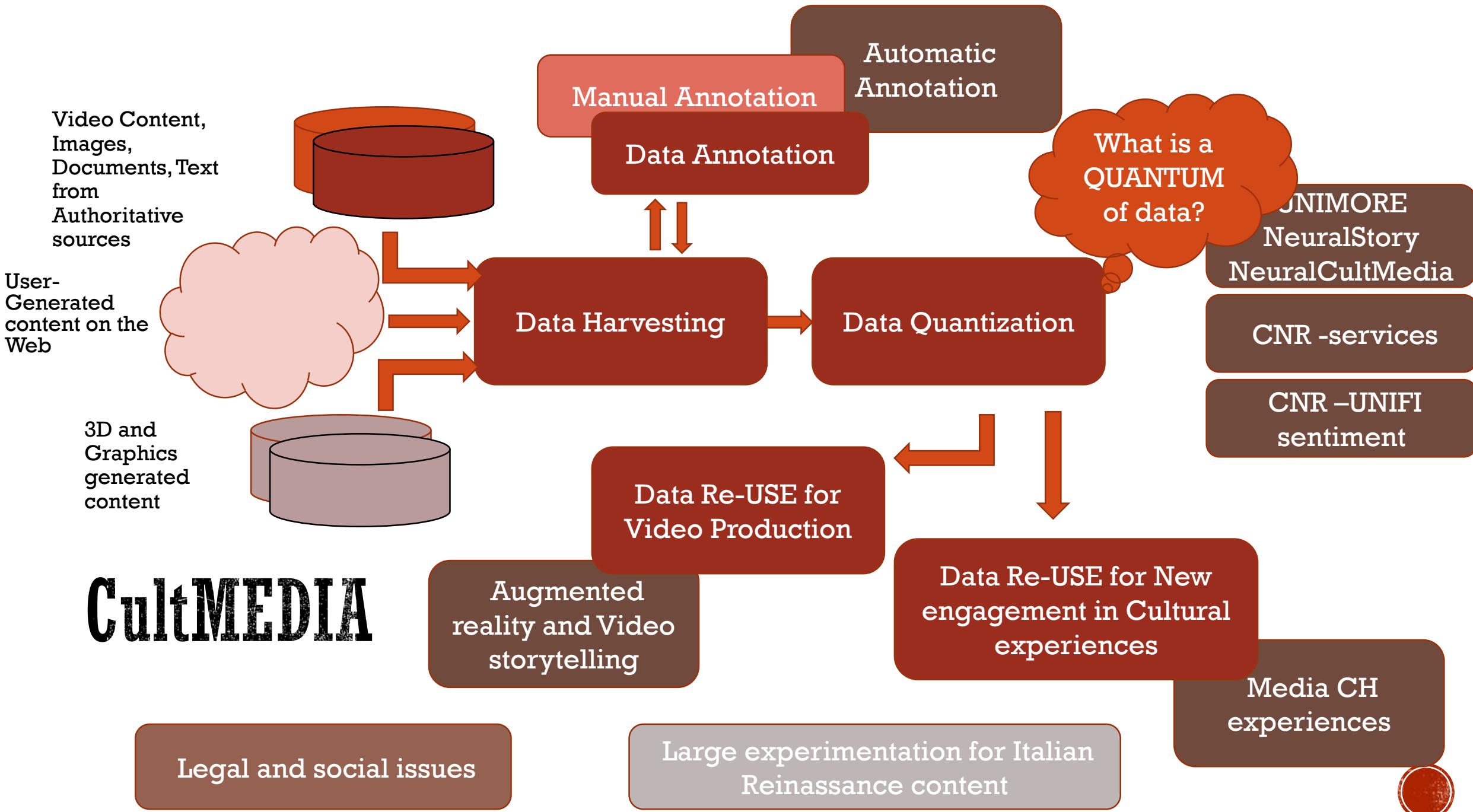
- The project final objective is the design and development of an innovative platform of services for producing multimedia and transmedia storytelling on cultural heritage.
- **software tools and services for:**
 - a) semi-automatic harvesting of reusable visual material for new productions
 - b) the support for combining heterogeneous content such as 3D graphics, text, audio and video, to create new storytelling on CH coherent with the rights of cultural products;
 - c) the production and post-production of low-cost video, capable to join content and user experience.
- *The platform will be based on advanced machine learning techniques and artificial intelligence for the automatic extraction of knowledge from video content (semantics, component scenes, emotional moods, saliency).*
- **a highly significant cross-regional use case** with materials
 - from national public and private museums of the Italian Renaissance,
 - from user-generated content
 - from the web.



TICHE CLUSTER

- CultMEDIA
- CHERIE





Video Content, Images, Documents, Text from Authoritative sources

User-Generated content on the Web

3D and Graphics generated content

CultMEDIA

Manual Annotation

Automatic Annotation

Data Annotation

Data Harvesting

Data Quantization

Data Re-USE for Video Production

Data Re-USE for New engagement in Cultural experiences

Augmented reality and Video storytelling

Media CH experiences

Legal and social issues

Large experimentation for Italian Renaissance content

What is a QUANTUM of data?

UNIMORE
NeuralStory
NeuralCultMedia

CNR -services

CNR -UNIFI sentiment



Technical and scientific limitations	Project ways to overtake limitations in reaching the objectives
<i>Videos are often too long and only clips are available for reuse, and data are available mostly in unsuitable formats, not segmented in semantically homogeneous scenes.</i>	Guidelines for reuse-by-design and new services for automatic scene segmentation using all multimedia information available. (OR 1,3,4,5)
<i>3D graphic animations of virtual reality are almost never reused in video and cultural production</i>	Instruments for supporting automatic video clip creation from 3D animation. (OR1)
<i>Media data are scarcely semantically annotated: and video are not described at all: often there are not captions for shots, and clips.</i>	Automatic semantics extraction will be provided for designing new services of automatic <i>video captioning</i> with natural language based on deep learning and advanced computer vision, useful for clip selection and content compositions. (OR3)
<i>No annotations about the mood and emotional sentiment are available.</i>	New services of automatic <i>emotional semantics extraction</i> will be provided to support new forms of video compositions based on sentiments and moods. (OR3)
<i>No information about experiences of users.</i>	The <i>user-experience analysis</i> both to support the platform design and to assess the efficacy of the produced contents will be exploited. (OR2,5)

- **Keypoints..**



OR1: THE PLATFORM

CNR LEADER

- **A1.1 - Specifications of methodologies for video production endorsing the *reuse by design* approach** [CNR, UNIFI, UNISOB, ETT]
- **A1.2 – Design of the platform component for the creation/harvesting of 3D content** [CNR, UNISOB, ETT, Pdi]
- **A1.3 – Design of the platform component for the semi-automatic production of video pills from 3D models** [CNR, UNISOB, Pdi]
- **A1.4 – Review of storytelling methodologies and guidelines for the design of an authoring system** [UNISOB, UNIFI, UNIMORE, ETT, CNR]
- **A1.5 – Review of legislation and best practices concerning copyright management of visual media content in the CH domain** [UNISOB, CNR, UNIFI]



OR2: USER EXPERIENCE

- UNISOB LEADER
- **A2.1 - User experience evaluation methodology** [UNISOB, UNIMORE, UNIFI]
- **A2.2 - User experience tests** [UNISOB, UNIMORE, UNIFI].
- DATA COLLECTION

-TO CORRECT THE PROJECT?
For including all parters??



OR3: AUTOMATIC VIDEO UNDERSTANDING FOR RE-USE

- Leader UNIMORE
- **A3.1 – Design of solutions for automatic scene segmentation of long video and searching and retrieving re-usable video content [UNIMORE, UNIFI, PdI, ETT]**
- **A3.2 - Design of solutions for automatic video understanding for tagging and video captioning [UNIMORE]**
- **A3.3 - Design of solutions for automatic emotion and mood understating in video [UNIFI]**
- **A3.4 Design of solutions for text sentiment analysis and cross-media searching [CNR]**



OR4: SYSTEM INTEGRATION & GUI

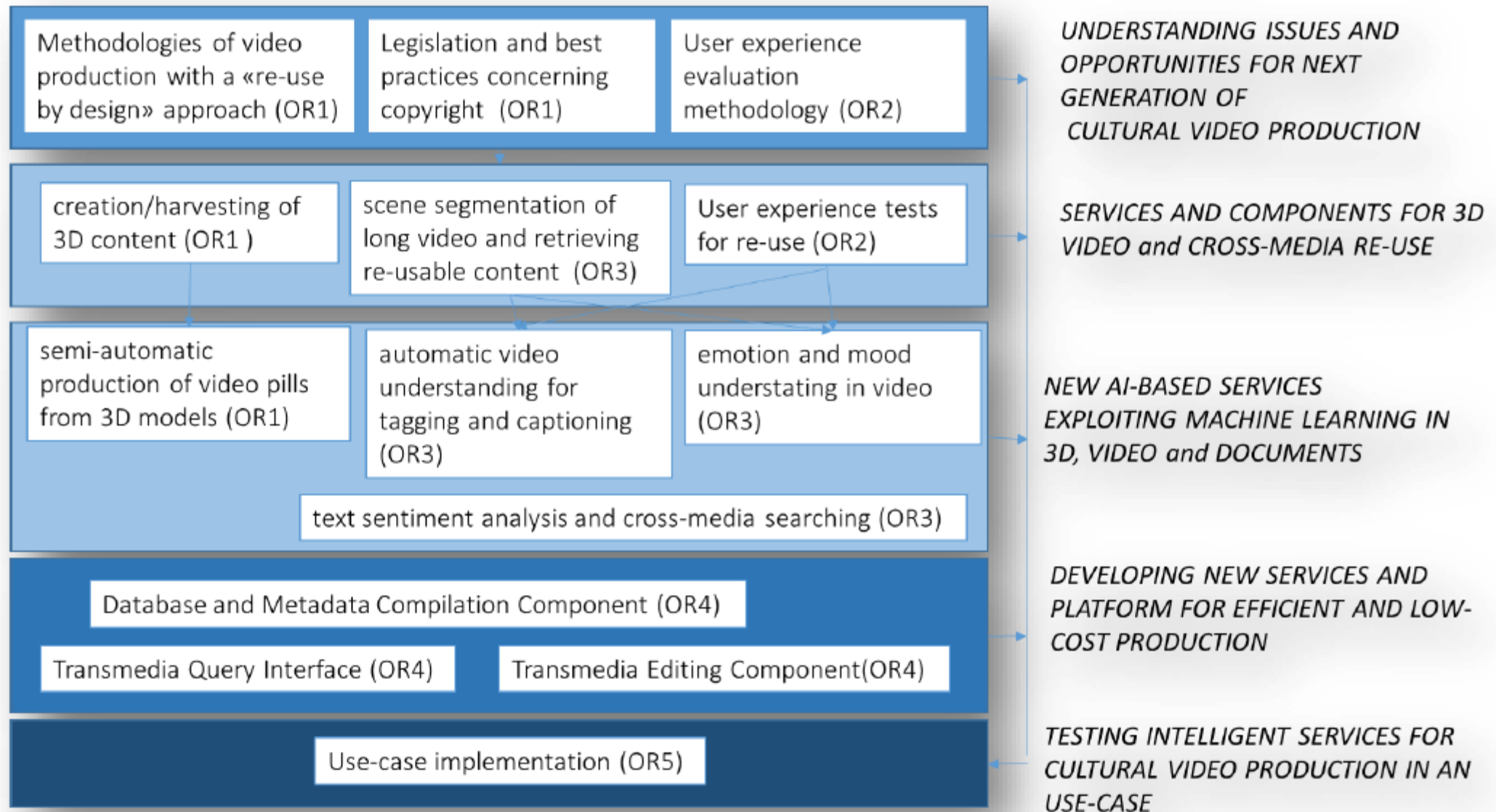
- Leader ETT
- **A.4.1 Database and Metadata Compilation Component [ETT, PdI]**
- **A. 4.2 Transmedia Query Interface [ETT, PdI, UNIFI]**
- **A.4.3 Transmedia Editing Component [ETT, PdI]**
- A.4.4 Development of the Graphical User Interface [ETT, PdI]



OR5: VIDEO PRODUCTION WITH TRANS-MEDIA RE-USE: A CASE STUDY

- Leader ?
- **A5.1 – Use-case requirements definition** [CNR, UNISOB, UNIMORE, UNIFI, ETT]
- **A5.2 – Use-case implementation** [UNIFI, UNISOB, UNIMORE, CNR, PDI, ETT]
- **A5.3- Use-case exploitation and transferability** [CNR, PDI, ETT]





WHAT WE WILL DO REALLY

- 1) PLATFORM FOR THE RE-USE OF VIDEO , DOCUEMNTS AND 3D CONTENT
- 2) Scientific research in Video captioning and annotation, emotion and sentiment analysis, user experience, 3D modeling.....
- 3) Use case for Reinassance Families
 - Farnese Family
 - Este Family
 - ..
- What else? (by the way the web pages..)



MY QUESTION- MY DREAM

- Can we create a **(Renaissance) art expert digital twin?**

What we need

- **A (new ?) list/taxonomy of artistic (Reniassance) concepts**
- **The biggest Dataset of images/video/3D elements with correspondent ANNOTATION**
- **New DL architectures from learning art concepts, (and art moods??)**
- **New models for captioning and connection with NLP**
- **New models for re-using this knowledge**
- **Some “ simulators”- computer graphics**



ANOTHER QUESTION

- Which kind of data?
 - Architecture?
 - Furniture?
 - Paintings?
 - Digitalized Documents?
-
- And the **TRUE art expert** partners?
 - Gallerie Estensi (Dr. Martina Bagnoli)
 - Galleria Farnese di Capodimonte?



A FINAL QUESTION

- And the final engagement?
- What can we do as experimentation?
- How can we measure the engagement?
- Where?
- When?
- What?



UNIMORE AImage Lab

- About 20 people
- Four/five staff people

Research in

- Computer vision
- Deep learning
- Multimedia
- HCI for Automotive



UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA



 **Facebook AI Research (FAIR)**

RedVision Lab

Aimage Lab UNIMORE and Ferrari spa



 **SCAI**
SuperComputing Applications and Innovation

Panasonic



THE COMPUTER VISION FOUNDATION

A NON-PROFIT ORGANIZATION THAT FOSTERS AND SUPPORTS RESEARCH IN ALL ASPECTS OF COMPUTER VISION



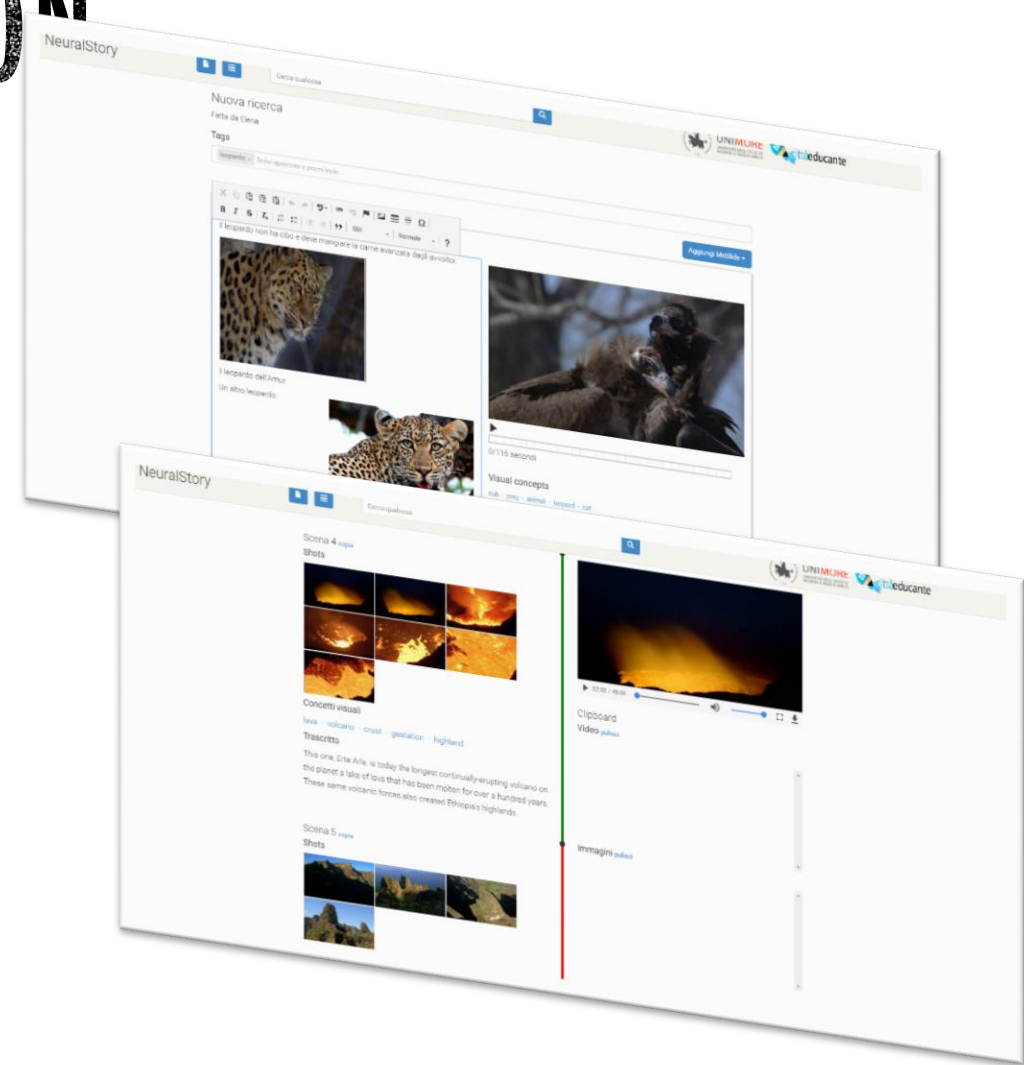
INTERACTION FOR EDUCATION

■ Neuralstory

1. Browsing and searching by summarization
2. Querying by text
3. Selecting from thumbnail

And then.. INTERACTING

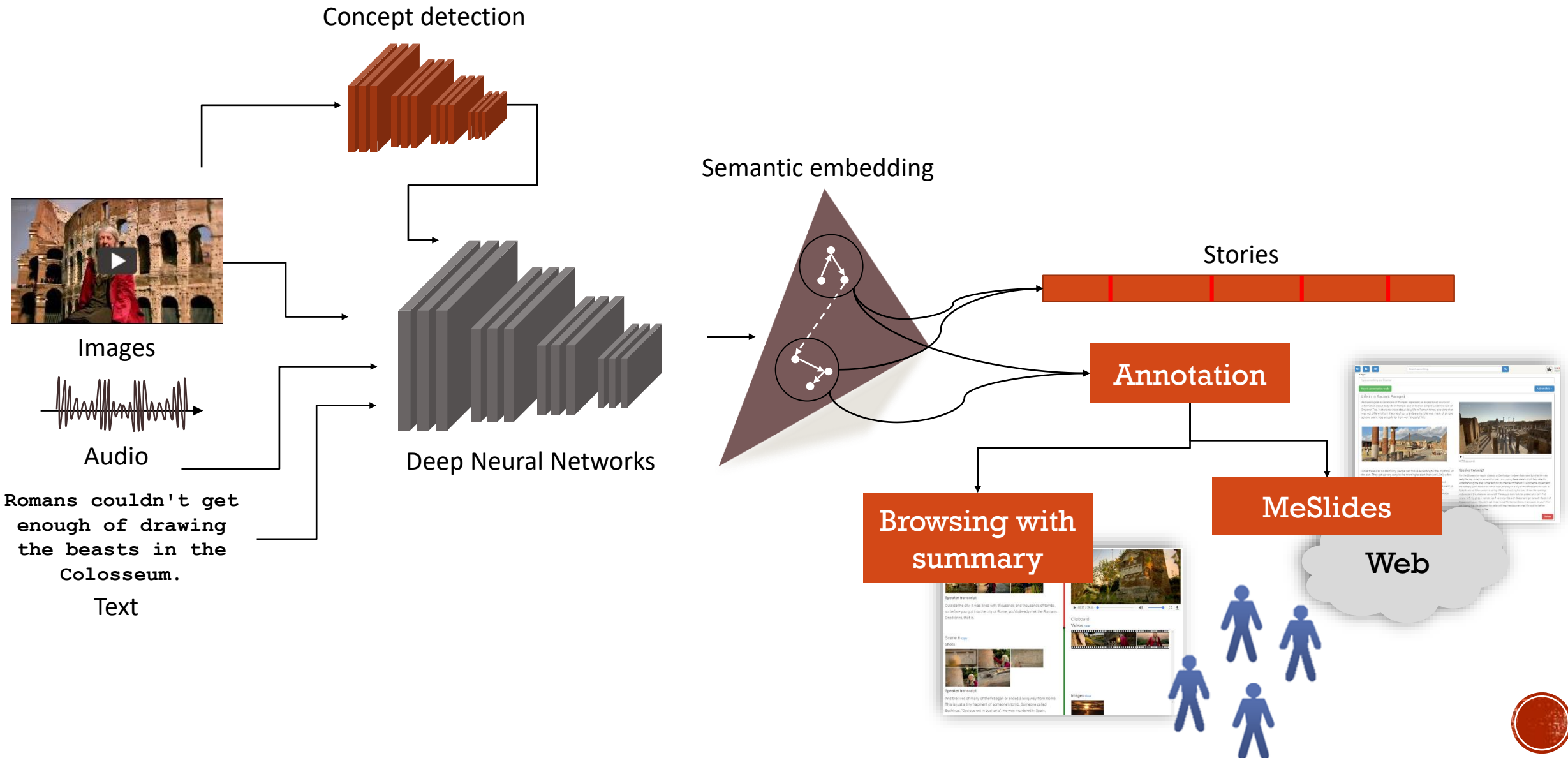
- Select text
- Select images
- Select video shots or stories



and then .. RE- USING



NEURALSTORY



From Pole to Pole

The first episode illustrates a journey around the globe and reveals the effect of gradual climatic change and seasonal transitions en route. During Antarctica's winter, emperor penguins endure four months of darkness, with no food, in temperatures of -70°C (-94°F). Meanwhile, as spring arrives in the Arctic, polar bear cubs take their first steps into a world of rapidly thawing ice. In northern Canada, 3 million caribou complete an overland migration of 3,200 kilometres (2,000 mi), longer than that of any animal, and are hunted by wolves during its journey. The forests of eastern Russia are home to the Amur leopard; with a population of just 40 individuals in the wild, it is now the world's rarest cat. This is primarily because of the destruction of its habitat, and Attenborough states that it "symbolises the fragility of our natural heritage". However, in the tropics, the jungle that covers 3% of the planet's surface supports 50% of its species. Other species shown include New Guinea's birds of paradise, African hunting dogs in their efficient pursuit of impala, elephants in Africa migrating towards the waters of the Okavango Delta, a seasonal bloom of life in the otherwise arid Kalahari Desert, and 300,000 migrating Baikal teal, containing the world's entire population of the species in one flock. The Planet Earth Diaries segment shows how the wild dog hunt was filmed unobtrusively with the aid of the Heligimbal, a powerful, gyro-stabilised camera mounted beneath a helicopter.

Scene 1 copy

Shots



Scene 2 copy

Shots



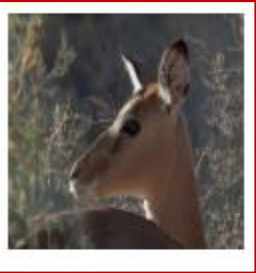
Visual concepts



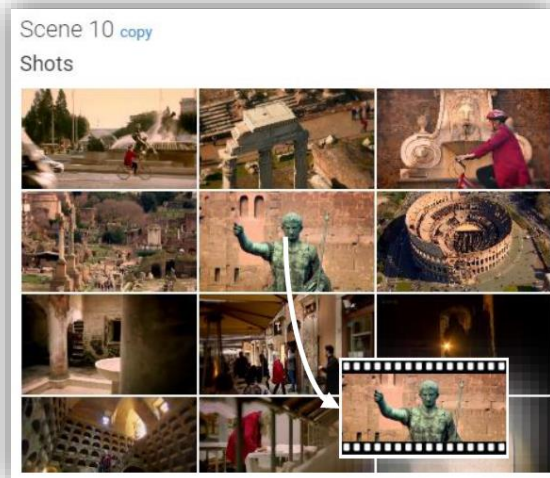
Clipboard

Videos clear

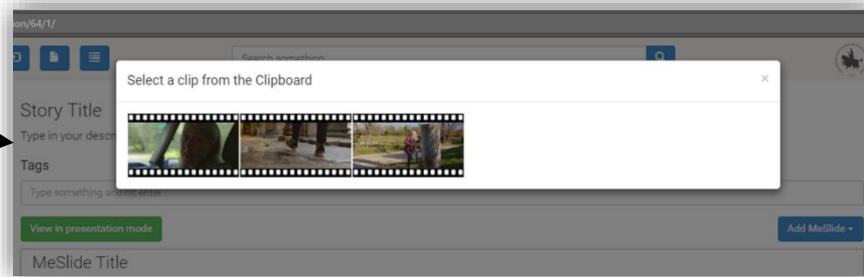




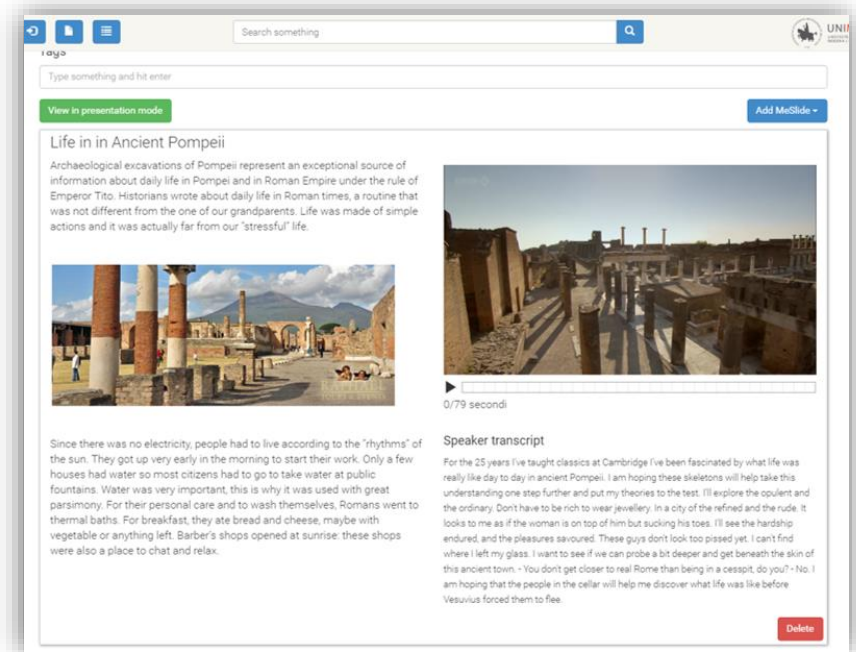
NEURALSTORY Meslide



Images, shots and scenes can be picked during watching



Selected clips can be used to create new multimodal slides



Which can be enriched with text and images

