

Received 14 April 2023, accepted 28 April 2023, date of publication 3 May 2023, date of current version 10 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3272641

## RESEARCH ARTICLE

# Agents Based Cyber-Physical Diffused Museums Over Web Interoperability Standards

DARIO BRANCO<sup>1</sup>, ALBA AMATO<sup>2</sup>, SALVATORE VENTICINQUE<sup>1</sup>, AND ROCCO AVERSA<sup>1</sup>

<sup>1</sup>Department of Engineering, University of Campania "Luigi Vanvitelli," 81031 Aversa, Italy

<sup>2</sup>Department of Political Science, University of Campania "Luigi Vanvitelli," 81031 Aversa, Italy

Corresponding author: Salvatore Venticinque (salvatore.venticinque@unicampania.it)

This work was supported by the Cleopatra Project funded by the University of Campania "Luigi Vanvitelli" through the VALERE 2019 Research Program.

**ABSTRACT** The availability of interoperability standards allows using a uniform model for integrating technologies and designing collaborating applications with a shared understanding of the exchanged information. In this perspective, we propose here the design and development of an agent-based platform that uses an interoperability framework to build a *diffused museum* in a cyber-physical system, where intelligent agents deliver cultural content through innovative interaction models. In particular, agents' awareness, natural language understanding capability, and interaction ability are extended through the monitoring of user's interaction with the IIF (International Image Interoperability Framework) standard and the integration with IIF compliant technologies. The scalability of the proposed solution is evaluated stressing the platform with simulated workloads.

**INDEX TERMS** Conversational agents, cultural heritage, web interoperability, web intelligence, context awareness, digital repository, semantic web.

## I. INTRODUCTION

Providing quality metadata and effective presentation of multi-media content remains one of the key challenges in promoting information access and use in digital environments, despite the high availability of information through the web. Research efforts in digital Humanities, Semantic Web, Linked Open Data, and other digital movements, continue to investigate innovative solutions, to define and promote interoperable standards [8]. Web services offer technology-neutral and well-defined interfaces for clients, but they are just a way of exposing system functionalities over the Internet. The semantic web can help to describe services, but the lack of one common ontology or the alignment of proprietary ones could be a problem [5]. For this reason, automatic retrieval of relevant contents and their effective delivery are critical issues as it needs to guarantee the correct understanding and an adequate level of engagement of users. Web intelligence techniques and technologies for Internet searching, information retrieval, content analysis, and link analysis have

been widely used in this context [16]. On the other hand, the availability of interoperability standards allows using a uniform model for integrating technologies and designing collaborating applications with a shared understanding of the exchanged information. Cultural Heritage represents an application domain in which organization, management, understanding, delivery, and content presentation are relevant. In this context, public institutions and private entities have launched large campaigns for digitization of cultural artifacts leading to the creation of massive digital libraries, whose design and development have been based on a set of widely accepted standards.

However, these large digital libraries are generally difficult for the general public to access by not offering intuitive mechanisms for navigating their contents. [14]. Moreover, whenever the relevant contents have been retrieved, in many cases correct understanding of the different cultural meanings and the exploration of related contents according to a specific itinerary, which is of interest for the user, become very difficult rather than unfeasible without a guide.

Intelligent conversational agents, exploiting interoperable interfaces and standard representations, provide a solution to

The associate editor coordinating the review of this manuscript and approving it for publication was Giuseppe Desolda<sup>1</sup>.

facilitate access to contents and services by a natural language interaction. However, a virtual storyteller to be effective as a human guide must be able to follow the user's visit, being increasingly aware of her knowledge and her interest. It is imperative that this awareness is built both by learning from the conversation and observing how the user interacts with the environment.

The International Image Interoperability Framework (IIIF) addresses interoperability problems, related to the delivery and presentation of multi-media contents, providing a technology layer for the uniform presentation of media contents and digital archives [13]. However, it lacks mechanisms for interactive storytelling that support the correct understanding and coherent fruition of delivered content, which are widely investigated in the field of smart-museums [28].

In this paper, an original approach to build adaptive storytellers for the presentation of digital repositories is investigated. The proposed approach:

- extends the natural language understanding (NLU) capability of storyteller agents making them aware of the user interaction with the digital contents through interoperable APIs. The same interoperable APIs are exploited to improve agent's argumentation ability.
- models both the user's profile and digital contents in terms of topics of interest (defined by concepts of an ontology) to address the problem of drawing personalized cultural itineraries recommending the most relevant contents.
- formulates an algorithm to select the optimal recommendations based on the maximization of contents relevance and on optimization of global criteria.
- investigates a methodology that uses synthetic workloads to verify the compliance with the desired service levels of a deployment configuration under stressing conditions.

The proposed approach can be generalized to a class of applications that support both conversational and cyber interactions and for which the context awareness can be built only by an information fusion, combining the learning outcomes obtained from the analysis of the two sources of data.

In Section II we introduce and discuss related work. In Section III we describe how, in the Cleopatra platform, intelligent software agents implement adaptive storytellers, which are able to retrieve, understand and use data and meta-data from digital libraries. Context aware mechanisms are presented in Section IV. The algorithm for selecting the most relevant contents to be recommended by the adaptive storyteller is described in Section V. Section VI presents the software architecture of the Cleopatra platform. A methodology for the dimensioning and evaluation of the deployment configuration is discussed in Section VII. Finally conclusion is due.

## II. RELATED WORK

The development of virtual or augmented cultural experiences in the context of digital humanities has introduced

new requirements for delivering media contents and related meta-data through the web. One among others is the utilization of interoperability standards for delivering high-quality, attributed digital objects online at scale [10]. The availability of interoperability standards allows for removing the virtual silos that cultural institutions have created to provide images on the Web [23] and to leverage consistency, flexibility, and interoperability. Besides, developing shared APIs have also been cost savings [21].

The International Image Interoperability Framework<sup>1</sup> (IIIF) has been defined by a community of academic and national libraries, research institutions, museums, archives, nonprofits, and commercial organizations that are committed to interoperable image delivery on the web [25].

IIIF was designed to access and share contents by promoting the building of a global and interoperable framework by which image-based resources could be easily shared and reused across institutions using any combination of image servers or client viewing software [6]. IIIF standardizes Application Programming Interfaces (APIs) at different levels. The Presentation API provides structural and presentation information. It adds hierarchical layers to the images so that API clients can display them in a comprehensible way, using collections, sequences, and page layers for example. Image API provides a standard image delivery API. Content search API allows users to make search queries on the image meta-data and text transcriptions. Audio and Video Content API delivers time-based media (audio, video); Authentication API manages access rights to the four previously described APIs. It is optional, as most of the available content is in the public domain.

Based on such a defined standard API for image and metadata delivery, best-in-class image delivery and comparison software have been developed. A sustainable community model for interoperable image repositories has recently grown. Some examples of compliant servers are Loris, IIP-ImageServer, Cantaloupe, ContentDM, Djakota, and SIPI. Web clients include OpenSeaDragon, Leaflet-IIIF, Diva.js, IIIFViewer, Universal Viewer (UV), and Mirador. They provide a world-class user experience in viewing, comparing, manipulating, and annotating images. Such a framework, and the related technologies, offer uniform and rich access to image-based resources hosted around the world and support interoperability between digital repositories and the related meta-data.

In recent years an increasing number of digital archives have been published using the IIIF APIs, which leveraged the development of innovative solutions in the field of Digital Humanities. In [4] an architecture for the automatic processing of historical documents is described. Those documents are owned by different institutions but processed and presented thanks to the IIIF framework. Authors implemented this architecture and processed a large collection of books with a page classifier trained on an annotated sample.

<sup>1</sup><http://iiif.io>

The result is freely distributed and can be viewed with any IIIF-compatible viewer. In [2] authors describe an experience of porting a collection of semantically annotated images from a legacy content management system to a new technological stack made of independent open-source software to support the IIIF interoperability standard.

The interoperability tools made available by the IIIF standard encourage the development of intelligent software agents that are capable for presenting multimedia content and investigating new models of cyber-physical collaboration [7].

In the proposed paper IIIF is used as an opportunity to apply, across distributed digital repository, an innovative approach and AI techniques to improve the context awareness and the decision making for the adaptation.

Even if several digital repositories have adopted such standards, they still do not integrate intelligent techniques that can enhance content delivery and support fruition, such as automatic or interactive storytelling. Interactive storytelling is a form of digital entertainment where authors, the public, and virtual agents participate in a collaborative experience [20].

Digital storytelling is the use of digital media platforms and interactivity for narrative purposes, either for fictional or non-fiction stories. It has applications in museums, teaching, tourism, and many more [17]. There are several works on forms of interactions for interactive storytelling in the literature. They cover the subject from traditional GUI interfaces to more complex interaction mechanisms, such as speech recognition, body gestures combined with speech, hand-drawn sketches, and physiological inputs [20].

Conversational agents have been widely used in studies exploring their potential for engaging with and sharing information about cultural heritage. These studies demonstrate the effectiveness of conversational agents in providing a natural and intuitive way for users to access and interact with cultural heritage information, making it more accessible and engaging for a wider audience. In [14], authors discuss the challenges associated with searching digital cultural heritage collections, such as Europeana, which has vast amounts of digital objects but lacks efficient search engines. The paper presents a solution that utilizes conversational agents and provides a user-friendly interface for exploring Europeana's content, allowing users to search using natural language and refine results to support specific queries. However, agents are not context-aware and do not implement learning or storytelling capabilities.

The approach presented in these papers use stateless services (expand the query using an ontology, recommend related content using related data) are based only on the user query (in natural language or keyword based) and do not learn from the user's interaction with the digital contents but use static profiles. Nevertheless, in our approach adaptive storytelling allows for the valorization of the presentation of a digital repository and NLP improves the interaction and allows us to dynamically build a storyboard.

To select most relevant contents semantic techniques have been exploited both for annotate data with concepts of the domain knowledge and to dynamically profile the users. Semantic Web technology and ontology modeling methods are applied in [28] to construct advanced digital services, supporting the study and evolution of museum collections. An ontological model is used to collect and interlink information in the form of a semantic network and the information ranking is based on attribute and connection structure in the semantic network. Cultural trip planning service is described in [12]. This service is based on a software platform for creating smart spaces where part of a smart space includes "agents" and an information "hub." Each agent is an autonomous knowledge processor (KP), which is a software module running on some device. Cultural trip planning service provides the ability to process personalized search requests and make routes for selected points with a time plan.

The evolution and continuing advancement in research and development of technologies over the past decade influenced the significant increases in the interest in complex computer systems development. A domain that drew a lot of interest is the area of Cyber-Physical Systems (CPS). In particular, [19] focuses on demonstrating the development of a Cyber-Physical-Social Eco-System (CPSeS) capable of seamlessly blending the real world with virtual social spaces by intertwining several technologies, including real and artificial agents and elements capable of dynamically interacting, reflecting, and influencing each other with the interactions engendered by humans and their behavior. To demonstrate the potential of the proposed system, a Virtual Museum case study methodology has been implemented. The use of Virtual Reality through the use of smartphone devices as a means of a Cyber-Physical-Social system to support, improve and enhance the visitors' experience is also discussed in [18] where authors present a prototype that has been developed to investigate the potentials of such technology to support and enhance museum experience, aiming at introducing the concept of a Cyber-Physical-Social system that supports social, interactive and immersive experience to visitors. A relevant issue in these work is related to performance and scalability of the proposed solutions. It is an important tool to make informed decisions about optimizing the system performance and scalability analysis because it allows you to test a system under different load conditions. Simulated workloads can help to evaluate the capacity of a deployment configuration and to determine how it will handle high-demand periods of user activity or future utilization growth. Simulating different user activity levels, it is possible to understand how the system performs and to identify any bottlenecks or other issues that may affect its performance. For this purpose, a distributed load simulation software, has been used successfully in various application areas. For example, in [15] Tsung,<sup>2</sup> an open-source multi-protocol distributed load testing tool,

<sup>2</sup><https://github.com/processone/tsung>

is used to simulate two cloud-based solutions for connected cars with different functionality such as Eco-Driving, data collection, and weather forecast. In [3], authors compare three popular load stress testing tools: Jmeter, Tsung, and Autop-erf. Their analysis shows that Tsung is a feature-rich tool with excellent performance capabilities. It is one of the best options available for simulating network load and assessing the performance and scalability of a system. Throughout the paper, the authors provide detailed comparisons of the features and performance of each tool, highlighting the strengths and limitations of each.

### III. THE CLEOPATRA PLATFORM

The Cleopatra software platform allows for building a dynamic P2P overlay of users and software agents. It represents at the same time a technological testbed for evaluating the effectiveness of the proposed solution and an outcome of related research activities. Users visit virtual or physically distributed cultural assets and exploit advanced services on their devices or through other terminals available at the physical site. Intelligent software agents are in charge of the ubiquitous delivery of cultural information, accessing distributed digital repositories, and exploiting other IoT data. On the left side of Figure 1, at the lower level, we can see blue-physical points of interest, which are geographically distributed, and red-virtual locations, which can have a geographical attribute or not (virtual access to a physical museum or multi-media information about a physical location). At the middle level, we have the meta-representation of locations by collections of IIF manifests. An IIF manifest is used as a representation of either a concrete artifact or an abstract cultural content by a sequence of media elements. It can be the virtual representation of a building, a monument, a masterpiece, a natural landscape, the room of a museum, or any other self-consistent information, through a sequence of semantically related multimedia objects (pictures, video, draws, . . .) with their meta-data.

On the right side of Figure 1 is shown a web presentation of a list of IIF manifests (at the bottom), through an IIF viewer and the specific picture of one selected manifest with the related metadata.

In Figure 2, a snapshot of a prototype implementation that uses the Mirador viewer to integrate a web chat to communicate with the Cleopatra NLP (Natural Language Processing) agent is shown. The user can navigate the images related to the archaeological remains of the ancient Appia [22], and can ask for any related information.

The Cleopatra software agent is able to plan a cultural itinerary answer the users' questions, and complement interactive storytelling with the presentation of semantically related multimedia.

In general, the Cleopatra platform can be used both on-site and remotely being that the technology on which it is based can be exploited either by setting up kiosks with monitors, microphones, and other input equipment or remotely using users' terminals. A museum, for example, might decide to

adopt the Cleopatra platform on its website to make its content usable remotely to users visiting the website. There are no major differences between the two implementations, except for the different handling of the user access mechanism to the platform.

### IV. CONTEXT AWARE MECHANISMS

A seamless integration of the web chat with the viewer is not enough to provide a good experience. The agent's awareness would be limited to the content of the messages exchanged with a user. For example, a simple question like "What is this picture?" cannot be answered as the agent does not know what canvas is shown by the Mirador viewer. Similarly, if the user were asked to specify the argument of his question, for example, "the Conocchia funerary monument," then throughout the conversation, the agent would assume that the topic has not changed unless the user makes an explicit reference to a different matter.

We mean that agents should be able to learn from the conversation with the user, but also from the interaction between the user and the environment.

#### A. ANALYSING THE CONVERSATION

The Cleopatra platform integrates different mechanisms, which can be exploited to increase the context awareness of intelligent agents, implemented by state-of-the-art technologies of sentiment analysis. In fact, as it is stated in [26], it is relevant for the agent to be sensitive and understand human emotions from the conversation. For this reason, the user's reaction to the presentation of cultural content, as well as their opinions on social interaction, has been considered a relevant source of information to learn about the impact of content delivered by Cleopatra. In this context sentiment analysis is used to process users' comments in order to:

- increase context awareness evaluating user's satisfaction to information or a recommendation or a general topic [9],
- increase context awareness evaluating user's disbelief or belief about a topic [11],
- detect misinformation or disinformation in multi-users conversations.<sup>3</sup>

The user's reaction to the agent's message can be used to understand if that phrase raises a positive, negative, or neutral emotion, or if it arouses astonishment or disbelief. In that case, the agent becomes aware of the user's emotions and can choose to provide additional information or to argue about the same topic.

In the second case, a popular reaction from many users to a certain expression can be exploited to identify the source of misinformation or malicious messages, such as the wrong cost of the ticket or the recommendation of scam exhibitions. Such mechanisms can be activated by invoking custom actions when a specific user's intent is detected in a defined conversation flow, can be integrated into the NLP

<sup>3</sup><https://github.com/alescardicchio/COVID19-FakeNewsDetection>

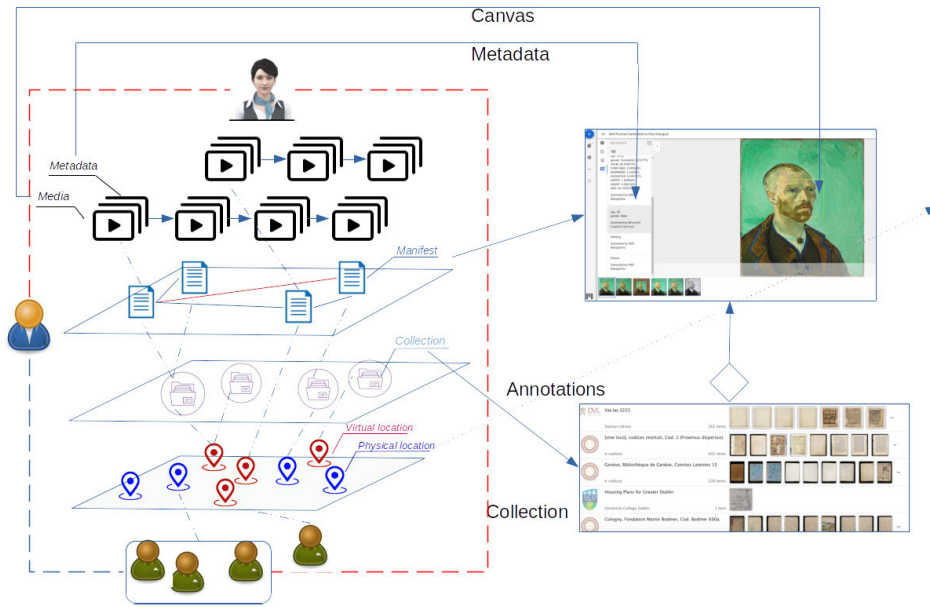


FIGURE 1. The cleopatra platform.

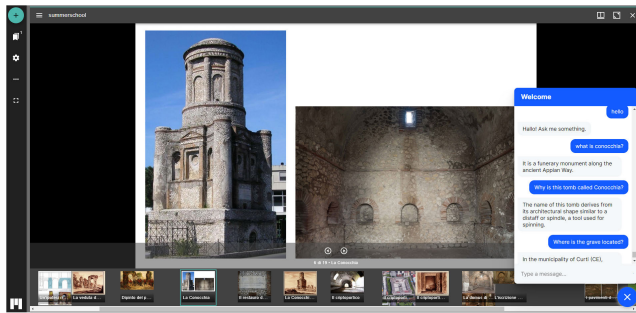


FIGURE 2. Web-chat support for user-agent communication.

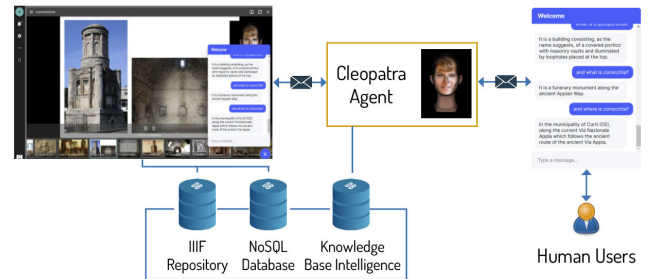


FIGURE 3. NLP augmented interaction.

processing pipeline, or can be performed offline on tracked conversations.

The latter two mechanisms are made available by Cleopatra through trained models, however, they can be trained for the specific use case if a related dataset is available. Currently, the outcome of the analysis is used to monitor the effectiveness of the conversations and to alert human moderators for the offline revision of the agent’s knowledge base.

**B. USERS INTERACTION IIIF SESSION**

Beyond the improvement of the NLP agent’s capability, advanced learning mechanisms allow to turn the conversation into personalized interactive storytelling. Moreover, building collective knowledge from such kind of information offers an additional opportunity to improve the user’s personal or social experience. Relevant examples are the coordination of visitors, in the case of physical sites, the identification of most preferred contents, and the identification and enhancement of secondary cultural contents. In Figure 3 the Cleopatra agent

can have a conversation with the user either through a simple messaging app or by supporting the visualization of media content by the viewer. In the first case, the NLP capability of the Agent is enough to provide an effective interaction because the topic of conversation will be made explicit in the messages. It means that users will never refer to an implicit foreground image. Nevertheless, it will be asked by the agent to send a picture if the user refers to something he is looking at.

On the other hand, when the interaction occurs through the viewer such a limitation is overcome by making aware the agent about the user’s interaction with the viewer, notifying the IIIF events.

The list of information that contributes to defining the IIIF context are:

- the selected manifest;
- the selected canvas shown to the user;
- the zoom level;
- any selected annotation.

The value of this information allows the agent to be aware of what the user is focusing on and to improve the understanding

of the user’s request. The context can be changed by the agent itself, for example showing a different image, or its detail, to better illustrate a response. For this reason, the information about who has changed the canvas, or zoomed or selected the annotation is an additional element that needs to complete the context.

Moreover, the history of occurred events allows one to evaluate how much time the user has spent on each selected content, the visiting order of the contents, and the details of each content that captured his attention.

Digital IIIF contents are eventually semantically annotated. Semantic wrapping of delivered images is supported as well by the IIIF standard. Cleopatra use cases exploit the IIIF annotation using concepts of a domain ontology and providing the support for indexing and discovering relevant contents by a SPARQL query [2]. Moreover, the annotations of enjoyed contents are stored in the user’s profile to make the application context aware with respect to the user’s preferences.

An underline infrastructure will allow for the detection of IIIF events to be communicated to a remote service that handles the IIIF session and the enforcement of application controls asked by the agent to the viewer itself.

**C. IMAGE RECOGNITION**

In live exhibitions, IIIF sessions are complemented with additional information which helps to characterize the context. They include GPS position, for archaeological areas and image recognition for indoor showcases and museums. Image recognition technologies have been integrated into the Cleopatra platform to provide two different services. Face recognition is used to identify and localize the user through physical kiosks equipped with monitors and cameras. When the camera identifies a registered user, an idle kiosk localizes the user and eventually sends him a message to recommend content, to ask about her need, or simply for salutation.

Only the result of image processing, which are anonymized models, are persistently stored. Nevertheless, when they are related to personal information of registered users they become sensitive information and, in real deployment, they must be handled according to the European GDPR regulation.<sup>4</sup>

In this case, the user identity, stored in a persistent database, allows for retrieving the information related to his profile and the current session if it is active. If the user is unknown, a new session is started and the user is invited to use the platform on her smartphone and eventually register.

QR-code scan or image recognition is also available to retrieve and visualize the digital version and related information of a painting or any other artifact. An example is the magnification of a detail accessing the IIIF interface of the digital twin or showing the obverse side of a coin.

The two mechanisms are complemented by voice-to-text and text-to-voice mechanisms which allow for a faster and

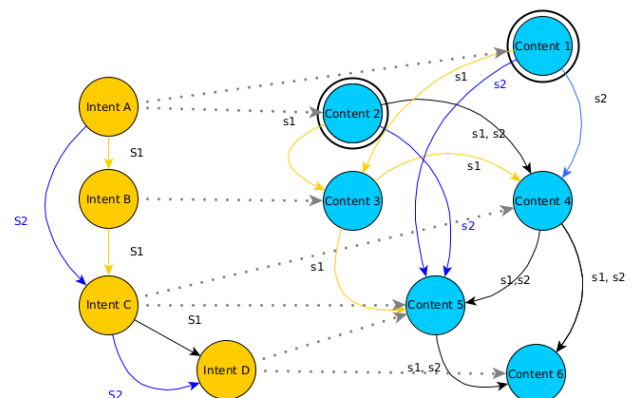
more comfortable interaction by those users who cannot or prefer not to write messages or want to listen to responses and storytelling. They work also wherever new requirements could emerge. An example is the prevention of shared use of physical devices installed in museums or any other live exhibition. In this case, a deployment configuration has been developed where a voice assistant allows for a touchless interaction with the agent at the physical kiosk equipped with an ambient microphone, monitor, and eventually a camera. In this case, the voice assistant represents an external component of the front-end agent that sends messages to the virtual agents whose responses are returned to the front-end instance running into the browser on a fixed installation.

**V. CONTEXT AWARE STORYTELLING**

A Cleopatra agent is able to deliver context-aware storytelling through a personalized recommendation of the cultural content that is relevant to the ongoing conversation. In this section, the algorithm to adapt the storyboard and the mathematical formulation of the decision problem about the selection of the best content to be recommended to the users are presented.

**A. THE STORYBOARD COMPOSITION**

In Figure 4, the yellow nodes represent the user’s intentions understood by the agents in two different conversations  $s1$  and  $s2$ . The user’s intention is detected by the NLU capability of the NLP agent on each received message. They could correspond to any user’s question about a topic of interest. The two short conversation flows  $s1$  and  $s2$  are composed respectively of four and three interactions. Usually, NLP agents are trained with examples of conversations to better understand the user’s intention, even if at run time the actual conversation can be different from any sequences of the training set. The dashed arrows represent annotations that link intentions to some relevant cultural content, which may be presented to the users. We assume here that only one relevant content can be recommended by the agent, after that an intention has been detected. If no other criteria are defined the agent can tell a story for  $s1$  and  $s2$  starting from  $content1$  or  $content2$ ,



**FIGURE 4.** Drawing as storyboard according to the ongoing conversation.

<sup>4</sup><https://gdpr-info.eu>

presenting respectively four and three contents. We mean that the decision problem of drawing the storyboard consists of finding the best paths in the graph to the right side of Figure 4 for all users. In the example, a storyboard for  $s_1$  can be drawn traversing the graph from *content1* or *content2* in four hops, following the yellow or the black edges. A storyboard for  $s_2$  is drawn traversing the graph in three hops, following the blue or the black edges.

The decision problem can be formulated considering multiple parameters (e.g., the user’s interest, position, and the time available for the visit) and may aim at improving a selfish or a social utility. The main goals of a Proactive Agent are defined as fostering emergent behaviors by a group of visitors, which improve their cultural experience or allow them to leverage an effective utilization of available facilities. Some examples are the exploitation of unmissable artworks and the path of a complete and coherent itinerary through the available spots. With these aims, the agents can plan actions to leverage the achievement of their objectives. Examples are prompting users to take action within a predefined time or informing them about relevant artworks or findings in order to increase curiosity, engaging discussion with users to make them curious about the visit, or recommending to change location within a predefined time.

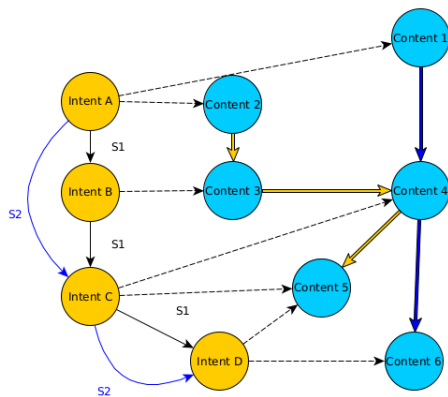


FIGURE 5. Optimal storyboards adapted to the user’s conversation.

In Figure 5 a solution of the problem is represented by the two yellow and the blue marked arrow. They have different lengths and share one content, which is presented on the third interaction with  $s_1$  and the second interaction with  $s_2$ . In the next sections, an implementation of such an adaptive storyteller is proposed. The planning of the storyboard according to the ongoing conversation is reduced to a decision problem about the content to be recommended to each user on his next intention. Personalized recommendations are calculated asynchronously for each user trying to improve both user satisfaction and social usefulness. In fact the recommendation system aims at promoting the less popular content but preferring the ones which are relevant to the user’s interest.

**B. IMPROVING THE GLOBAL UTILITY**

As a global utility, agents aim to estimate the most popular content and direct the users’ attention toward the less attractive artworks, promote lesser-known content, and decongest the most visited places in case of a live experience. The user’s satisfaction is improved by recommending those content that results in more relevant respect to her interest, learned from the list of viewed contents. The input for this activity is the evolution of all conversations, tracked by the *Rasa* NLP software and the IIF sessions. The following measures will be taken into account, in order to evaluate the popularity of artworks.

- $q_{i,j}$  represents the number of questions asked by the user  $V_i$  about artwork  $O_j$ .
- $t_{i,j}$  represents the time spent by the user  $V_i$  enjoying the artwork  $O_j$ . Tracking the continuous interaction related to the same content.

Based on these parameters, we define the *temporal approval rating* of the  $i$ -th user with respect to the  $j$ -th artwork:

$$st_{i,j} = \frac{t_{i,j}}{\sum_j t_{i,j}}$$

which expresses the ratio between the time spent observing that artwork and the total time of her experience.

Similarly, the *conversational approval rating* of the  $i$ -th user concerning the  $j$ -th artwork is defined as:

$$sc_{i,j} = \frac{q_{i,j}}{\sum_j q_{i,j}}$$

which expresses the ratio between the number of user’s questions related to the artwork  $j$  and the total number of questions asked during her experience.

We introduce the *temporal* and *conversational* rating vectors  $\vec{st}$  and  $\vec{sc}$  defined as:

$$\vec{st} = [st_0, \dots, st_n]$$

$$\vec{sc} = [sc_0, \dots, sc_n]$$

where the generic element of the vector represents the sum of users’ ratings for the  $j$ -th artwork:

$$st_j = \sum_i st_{i,j} \quad \text{and} \quad sc_j = \sum_i sc_{i,j}$$

Finally, making a weighted sum of the two vectors obtained (it can be assumed, for example, that temporal liking has a greater impact than conversational liking or vice versa) the *global liking vector* is derived:

$$\vec{\zeta} = \alpha \cdot \vec{st} + \beta \cdot \vec{sc}$$

The  $\vec{\zeta}$  vector encapsulates the approval for each artwork and is the tool for quantifying and comparing the interest that each of them arouses in visitors. The common goal for the NLP agents is to minimize the variance among the components of  $\vec{\zeta}$  in order to increase the popularity of less appreciated

content. The objective function to be minimized is defined in Equation 1.

$$\sigma_{\zeta}^2 = \frac{1}{n} \sum_j (\zeta_j - \bar{\zeta})^2 \quad (1)$$

where  $n$  is the total number of artworks. The array of normalized deviations  $\hat{\zeta}_j$  represents the first decision criteria to be used by the agent to select the content to be recommended.

It is defined as the differences between the average value deviations normalized over the maximum value.

$$\hat{\zeta}_j = \delta_j * \frac{\zeta_j - \bar{\zeta}}{\max_j |\zeta_j - \bar{\zeta}|} \quad \text{where} \quad \begin{cases} \delta_j = 0, & \text{if } \zeta_j > \bar{\zeta} \\ \delta_j = 1, & \text{otherwise} \end{cases} \quad (2)$$

It will be the first input for the optimization algorithm, helping to influence the conversation and the sequence of exploited contents so that less popular artworks can be still included in the visit if they are relevant to the user's interest according to the criteria defined in the next section.

### C. MAXIMIZATION OF THE CONTENT RELEVANCE

To take into account the user's preferences, the context awareness of the services is exploited by evaluating the relevance of each annotated content to the user's profile  $\vec{p}$ . The annotation and the user's profile are represented using the Vector Space Model (VSM). VSM is a model for semantic representations of contents as vectors of items created by G. Salton [24]. In our case,  $\vec{a}_j$  describes the annotation of content  $j$ . It contains  $l_j$  concepts  $c_{i,j}$  of the domain ontology, each one occurring  $o_{k,j}$  into the annotation. The user profile contains  $p$  concepts of the domain ontology which are relevant to her own interests.

- $\vec{a}_j = \langle \{c_{1,j}, o_{1,j}\}, \{c_{2,j}, o_{2,j}\}, \dots, \{c_{l_j,j}, o_{l_j,j}\} \rangle \forall j = 1..N$
- $\vec{p} = \langle c_1, c_2, \dots, c_m \rangle$

We defined a score  $\phi(\vec{a}_j, \vec{p})$  to measure the relevance of an annotation  $\vec{a}_j$  to the profile  $\vec{p}$  in Equation 3.

$$\phi_j = \phi(\vec{a}_j, \vec{p}) = \frac{1}{l} \sum_{k=1}^l r_k$$

$$\text{where } r_k = o_{k,j} * \frac{1}{m} \sum_{i=1}^m \frac{1}{d_{k,i} + 1} \quad (3)$$

where  $r_k$  represents the relevance of the profile respect to the concept  $c_{k,j}$  in the annotation.

To compute  $r_k$  we defined the affinity between a concept  $c_{k,j}$  of the annotation and any other concept  $c_i$  of the profile dividing 1 by  $d_{k,i}$ , which is the minimum number of edges of the ontology that connect the nodes, plus 1. The affinity with each concept of the profile is multiplied by the occurrences  $o_{k,j}$  of  $c_{k,j}$  and the average value on  $m$  elements of  $\vec{a}_j$  is computed to obtain  $r_k$ .

The higher will be  $\phi_j$  the higher will be the interest of the user in the annotated content. In addition, the user's constraints, such as available time for the visit, the device's

capability to play certain types of content, etc., can be taken into account as it was presented in [1].

This problem can be reduced to a *discrete optimization problem* that consists in searching the optimal value (maximum or minimum) of a function  $f : \vec{x} \in \mathcal{Z}^n \rightarrow \mathcal{R}$ , and the solution  $\vec{x} = \{x_1, \dots, x_n\}$  in which the function's value is optimal.  $f(\vec{x})$  is said *cost function*, and its domain is generally defined by means of a set of  $t$  constraints on the points of the definition space. Examples of constraints could be the time available for the visit or the compliance with the user's personal device.

Constraints are generally expressed by a set of inequalities:

$$\sum_{i=1}^n b_{i,j} x_i \leq a_j \quad \forall j \in \{1, \dots, t\} \quad (4)$$

and they define the set of feasible values for the  $x_i$  variables (the *solutions space* of the problem). In our case:  $w_i \geq 0 \forall i = 1..N$ ,  $B = \{b_1, \dots, b_t\}$  where  $\phi_i$  represents the score and  $B$  a set of constraints  $b_i$ , each one composed of  $N$  integer. We have to compute

$$\max \sum_{k=1}^N \phi_i x_i$$

so that

$$\sum_{k=1}^N b_{i,j} x_i \leq b_{j,N+1}$$

with  $x_i \in \{0, 1\} \forall i = 1..N$ . The goal is to maximize the value delivered, that is to select the set of content to recommend, without violating constraints. The vector  $\vec{x}$  represents a possible solution: its components  $x_i$  are 1 or 0 depending on whether the object is included or not in the best set.

The vector  $\hat{\phi} = \frac{\vec{\phi}}{\max_j \phi_j}$  of relevant recommendations that contain normalized relevance of annotated contents will be used to make the final decision that takes into account how to contribute to improving global utility.

### D. REDUCTION TO THE SINGLE CRITERIA

The final criteria for a decision on the personal recommendation are computed by a weighted sum of the two criteria which are related to the user intent. We give here the same relevance to the personal interests and to the global utility to compute the resulting vector  $\vec{q}$ , defined as

$$\vec{q} = 0.5 * \hat{\phi} + 0.5 * |\hat{\zeta}|$$

where  $\hat{\phi}$  must be computed for each user. The elements of  $\vec{q}$  are ordered from the highest value to the lowest one and are proposed in this order to the user.

The activity diagram of the optimization algorithm is given in Figure 6. The combined optimization represented in the central swim lane is triggered for each user on the profile update. In the swim lane on the left side, the affinity between the user's profile and the annotated contents is evaluated and normalized, selecting those contents which can be exploited



according to time and other constraints. In parallel, the current values or the updated global deviations are read. The maximization of the global utility is modeled swim-lane at the right side of Figure 6. It occurs once for all users periodically. From the tracked interactions, the normalized deviations are updated minimizing the variance  $\sigma_c^2$ , and the solution is combined with the relevance of each content to the user profile. The two criteria are combined and the recommendations are proposed to the user from the highest value.

## VI. SOFTWARE ARCHITECTURE

The UML component diagram of the software platform is shown in Figure 7. The user's front-end consists of a standard IIIF viewer and a front-end agent integrated as a javascript plug-in. The front-end agent implements an IIIF event listener and dispatcher. It detects IIIF events triggered by the user's interaction and notifies them to write into a No-SQL database which will later be called Events-DB. The front-end agent is also in charge of enforcing controls, received from a back-end agent intended to change the IIIF session proactively. Supported examples are the visualization of a specific canvas or the zoom on image detail.

The interaction between the viewer and the digital library follows the IIIF standard and is independent of the interaction protocol between the front-end agent and the Cleopatra platform. Any other data about the application context collected on the user side, such as sensor data, user's profile, client device, etc, are stored and accessible in the same way. Some examples are facial recognition of the user near an artwork, the geographical position, scanning a QRCode of artwork, etc. Using a document-oriented non-relational database allows storing data with a weak schema that can change depending on the type of event to be stored. The IIIF session is in charge to asynchronously notifying the NLP agents about updates of the IIIF session, but every collected data is made available to other components of the Cleopatra platform through direct access to the Events-DB. NLP agents receive updates from the session handler about the current status of the user's session to improve their natural language understanding. Nevertheless, they can access the Events-DB to execute custom actions.

Mirador<sup>5</sup> is the IIIF viewer extended in the prototype implementation of the presented architecture. The front-end agent is a NodeJS plug-in that writes IIIF events into the Mongo DB through the use of a Servier API bridge<sup>6</sup> and receives controls from remote agents. The RASA<sup>7</sup> framework has been used as NLP technology to develop the conversational agent.

An IIIF event is detected by the front-end agent when the user activates any functionalities of the viewer that require the IIIF API, either if it uses cached data or invokes a remote service. The front-end agents detect the event and notify the

Events-DB. The SessionHandler, which listens to the occurrence of new notifications, updates related parameters of the affected conversation.

## VII. SYSTEM DEPLOYMENT AND EVALUATION

In this section, we describe the application of a methodology for the evaluation of the deployment configuration that allows estimating in advance KPI performance indexes in operating conditions. Experimental results will allow also to evaluate the overhead of the storyteller implementation and the degradation of service levels when the workload scales up. The proposed methodology is based on the execution of a reference workflow that is preliminary recorded in a test session and is replicated for multiple users until the system is overloaded.

### A. WORKFLOW REPRESENTATION

In Figure 8 the reference workflow has been modeled using the Business Process Model Notation.

The BPMN process describes a short interaction between the user and the software platform. Different pools include the activities carried on by the software components of the system. The upper pools of Figure 8 represent the digital repository and the service that receives and stores the notified IIIF events. The lowest pool represents the remote NLP agent that integrates the IIIF session handler.

The pool in the middle models the user's actions. After the start event, on the first connection, the web page is loaded and introductory content is shown to the user. A new conversation is started by the NLP agent, that sends a salutation. In this stage, all the necessary software libraries are downloaded and the IIIF manifest is loaded and processed. Moreover, the first canvas is shown to the user.

In the second stage, the user continues her cultural experience by manually changing the canvas to explore a second content. She zooms on an image area to focus on a detail of interest. Such interaction triggers several IIIF requests to the image server to download the tiles of the selected images at the zoom level set by the user.

In the third stage, the user sends a chat message to the agent to get information about a specific topic. The agent understands the message according to the current session and replies with a chat response in natural language and controls the client application to change the current canvas. This canvas change triggers an exchange of IIIF messages again. Each time an IIIF event is triggered on the client side, the same event is stored to track remotely the evolution of the user's session starting the Events-DB populating sub-process.

In the last stage, idle time at the client side triggers a notification to the agent that selects and sends a recommendation to the user through a chat message in natural language.

### B. DEPLOYMENT CONFIGURATION

The software platform has been deployed on three different containers which respectively host the web and image server,

<sup>5</sup><https://projectmirador.org/>

<sup>6</sup><https://www.mongodb.com/>

<sup>7</sup><https://rasa.com>

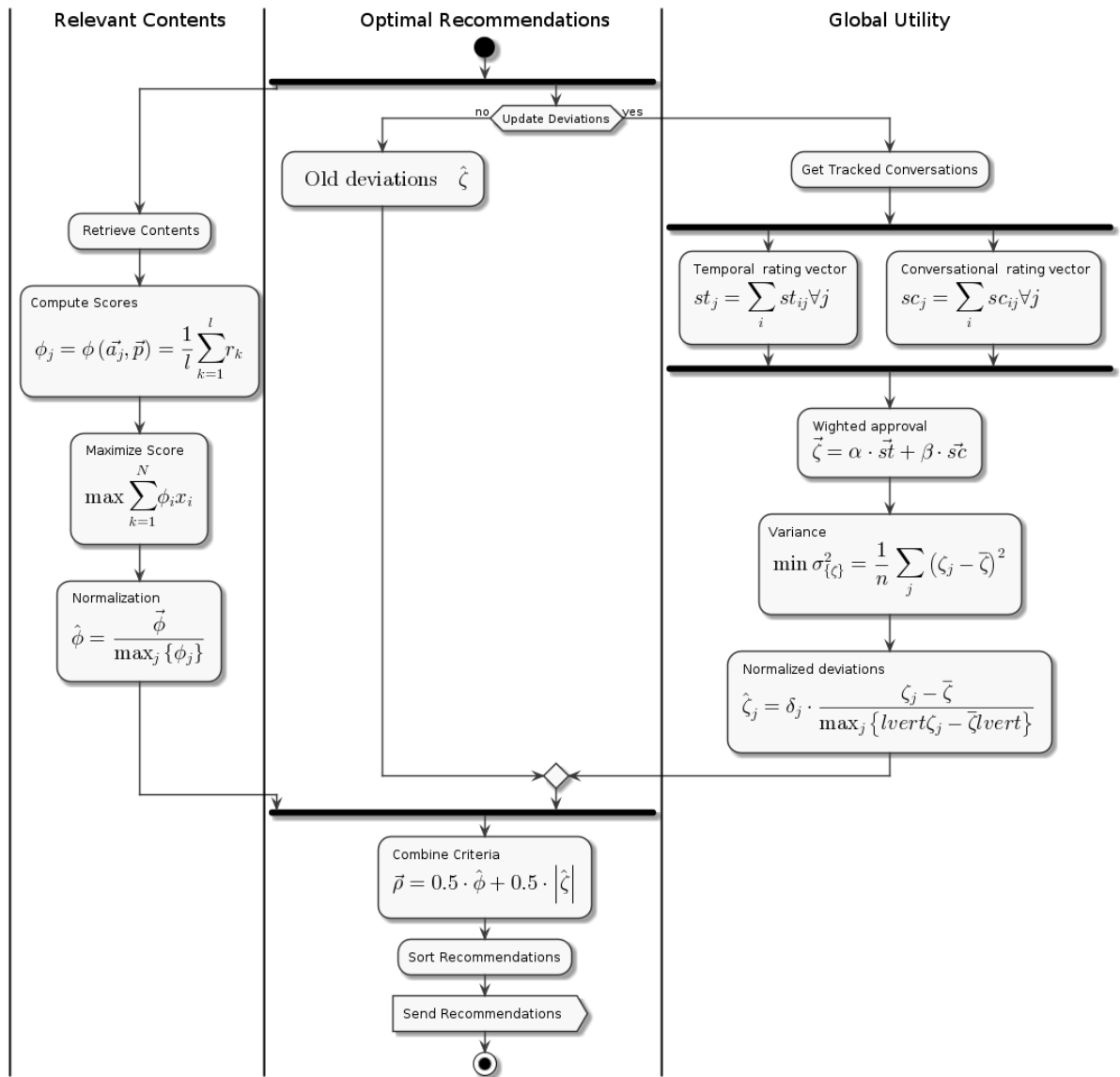


FIGURE 6. Optimization algorithm.

the Events-DB server that tracks the IIIF session, and the software agent platform.

They have been deployed on a workstation equipped with an Intel(R) Xeon(R) W-2245 with 8 cores and 16 threads at 3.90 GHz, 512KiB of L1 Cache, 8 MiB L2 Cache, 16 MiB L3 Cache, 4 memory banks of 32 GB of RAM for a total of 128 GB at 3200 MHz. The first container hosts the Apache web server and the IIPImage image server.<sup>8</sup> The second container runs the MongoDB service. Finally, the third container runs the IIIF Session handler, the Rasa platform that has been used to implement the NLP agent and the recommendation service. The direct interactions between the client and the different actors of the use case described in the previous section are shown in Figure 9. In particular, the purple and

the red interactions represent the communication overhead in the Cleopatra platform, handled by the javascript extension, concerning the legacy IIIF viewer.

The interactions between the client application and the back-end services have been recorded and replayed at a larger scale using the Tsung<sup>9</sup> stress tool. Tsung is an open-source tool for multi-protocol distributed load testing, allowing to stress HTTP, WebDAV, SOAP, PostgreSQL, MySQL, LDAP, MQTT, and Jabber/XMPP servers. It is developed in Erlang, an open-source language created by Ericsson for creating robust fault-tolerant distributed applications. In addition to the web request, the recorded trace includes waiting time between requests to model the user's behavior.

<sup>8</sup><https://iipimage.sourceforge.io/>

<sup>9</sup><https://github.com/processone/tsung>

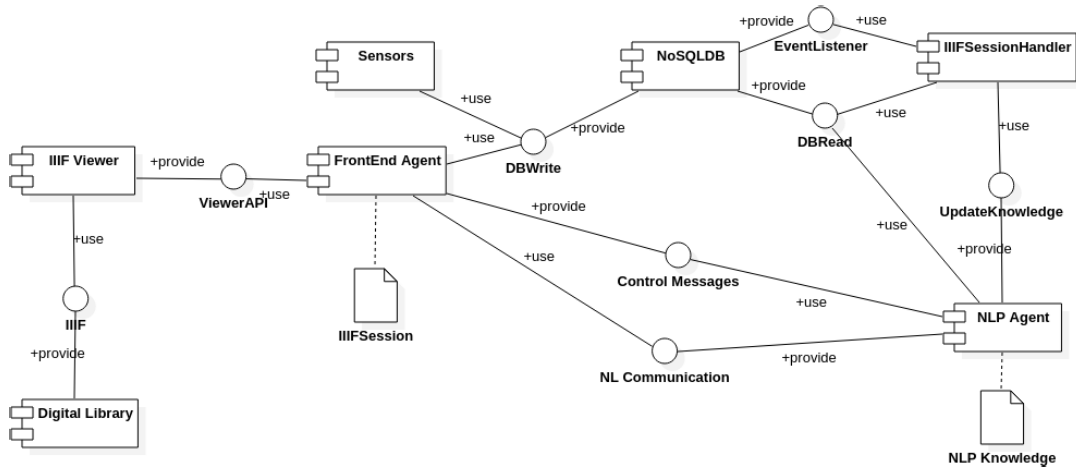


FIGURE 7. Component diagram of cleopatra platform.

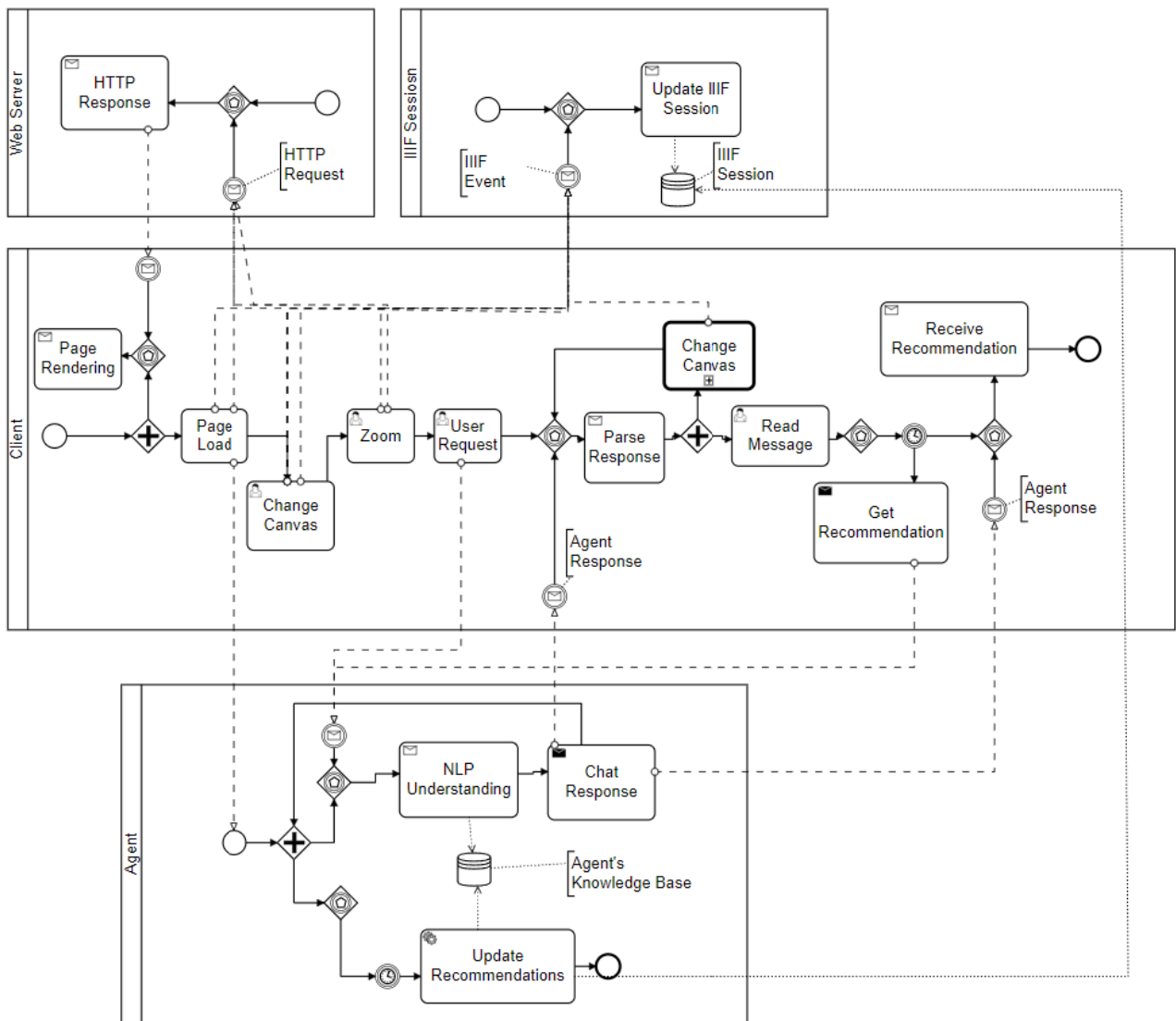


FIGURE 8. Use case scenario.

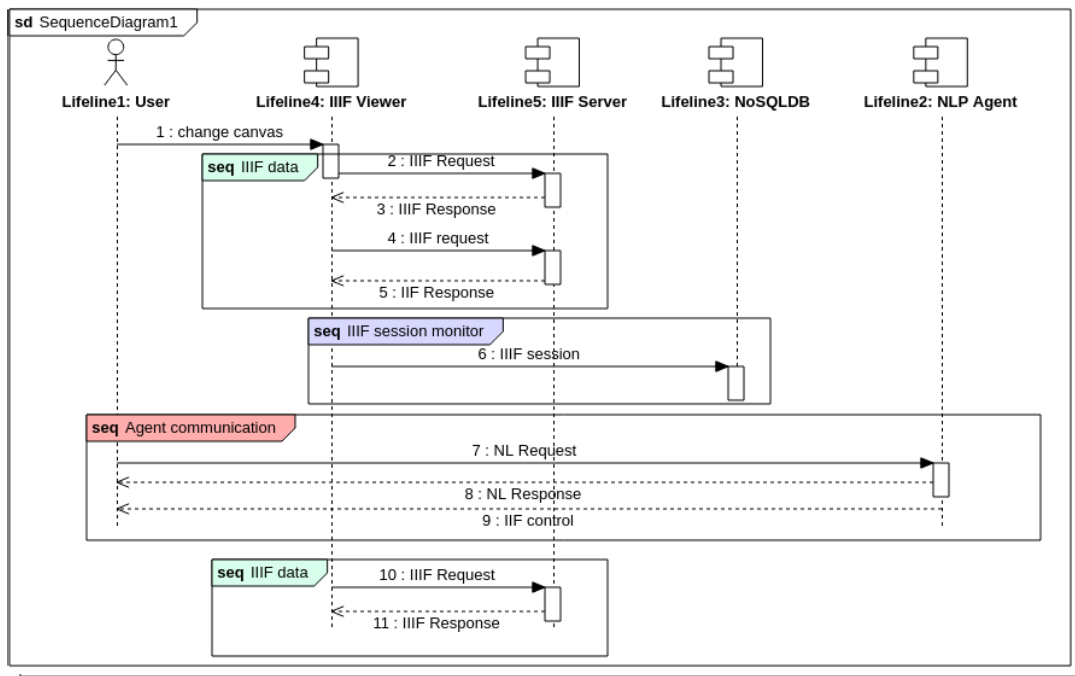


FIGURE 9. Communication type.

TABLE 1. Measured workload.

Requests	Mirador	Cleopatra
IIF	113	146
Mongo	0	22
Agent	0	5
Total	236	396

C. WORKLOAD CHARACTERIZATION

The workload of the test case generated by one user is shown in Table 1 in terms of the total number of web requests, either when the Cleopatra extension or when the standalone Mirador viewer is used. IIF requests consist of the download of the IIF manifest, the thumb images, and the tiles from the image server when the following events occur: *manifest download*, *change of canvas*, and *zoom*. Two requests are executed for the notification of each IIF event to the IIF Session Manager. In the Cleopatra test case, the additional IIF requests correspond to the change made to the viewer as an effect of the presentation of the agent’s response. Five interactions have been engaged between the client and the agent. One of five is the request sent by the client application after an idle time to get a recommendation. The remaining message corresponds to the download of the web page, required libraries, and CSS styles.

To compare the two different implementations, the performance evaluation of the Mirador experiment has been reproduced by replaying the Cleopatra trace but removing the additional Cleopatra requests. Performance statistics shown

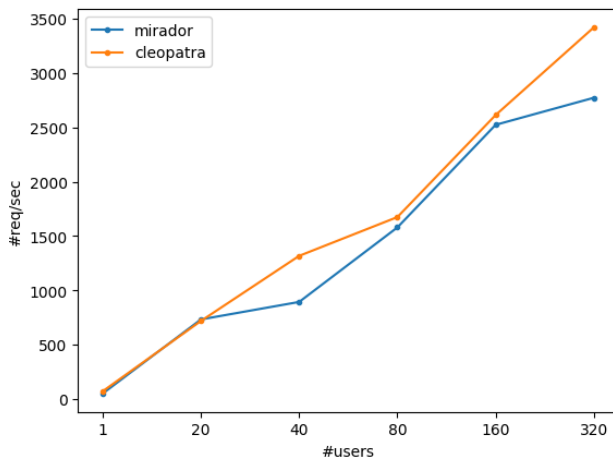
TABLE 2. Performance statistics.

Parameter	Cleopatra	Mirador
Request Mean Time	11.26 ms	11.24 ms
Page Mean Time	0.45 s	0.39 s
Bytes Received	4.76 Mb	2.05 Mb
Bytes Sent	64.58 Kb	40.57 Kb
Highest Received Rate	562.60 Kbits/sec	566.22 Kbits/sec
Highest Sent Rate	14.06 Kbits/sec	19.62 Kbits/sec

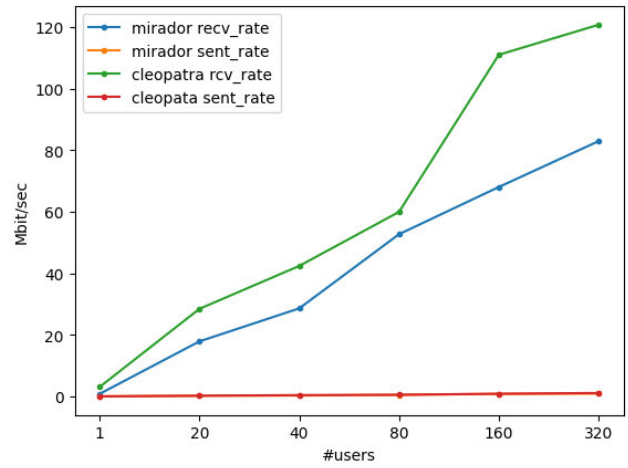
in Table 2 have been measured when only one user is connected, meaning the best level of service can be offered. In particular, *Page Mean Time* is the average time elapsed to execute a group of requests not separated by a waiting time. Higher values for the total byte transmitted and highest rate depend on the additional requests sent from the Cleopatra plug-in to the NLP agent and the IIF session manager.

D. PERFORMANCE ANALYSIS

To evaluate the scalability service we varied the number of users from 1 to 320 with an inter-arrival time of 0.2 seconds. Figure 10(a) shows how the maximum request rate and the throughput of exchanged data increase with the maximum number of users. The Cleopatra implementation exchanges a higher percentage of shorter requests, which motivates the slightly higher request rate. Moreover, a more significant amount of data is necessary to download the web page at the startup because of the additional required libraries. However, the maximum data transmission rate shown in Figure 10,



(a) Requests rate



(b) Size rate

FIGURE 10. Maximum throughput of requests and exchanged data.

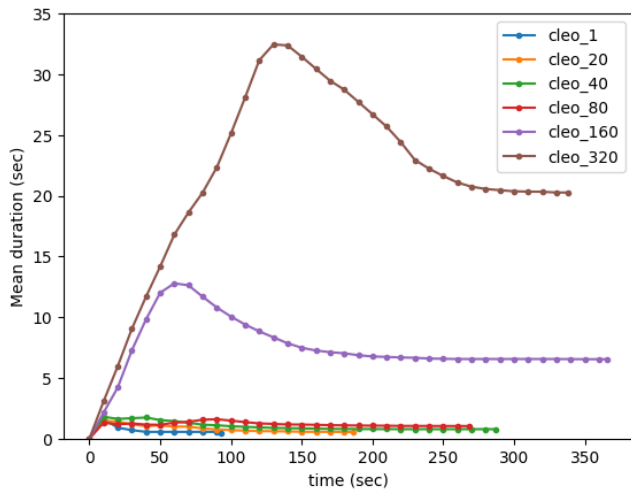


FIGURE 11. Mean response time for page download, measured from the beginning of the experiment.

as we will better observe in the following discussion, depends on the number of simultaneously connected users, which overloaded the server, rather than on the total amount of data exchanged in one session.

Figure 11 shows how the mean response time for a page download, computed from the beginning of the experiment, increases in the different experiments. We can observe that the mean time grows during the experiments with the number of users. When the workload decreases, due to the terminating sessions, also the mean response time starts to decrease.

The correlation between the number of simultaneously connected users and the mean page download time, measured in the last 10 seconds, is shown in Figure 12 for two specific experiments with 320 users, in both the case of the Cleopatra platform and the standalone Mirador viewer. It can be observed that the maximum number of simultaneously connected users occur almost simultaneously, both for Mirador

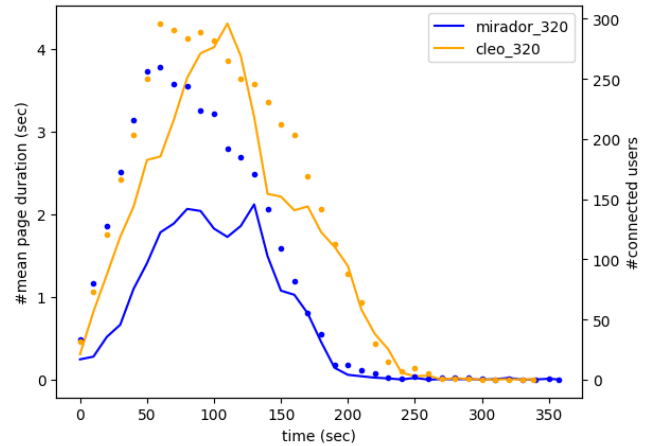
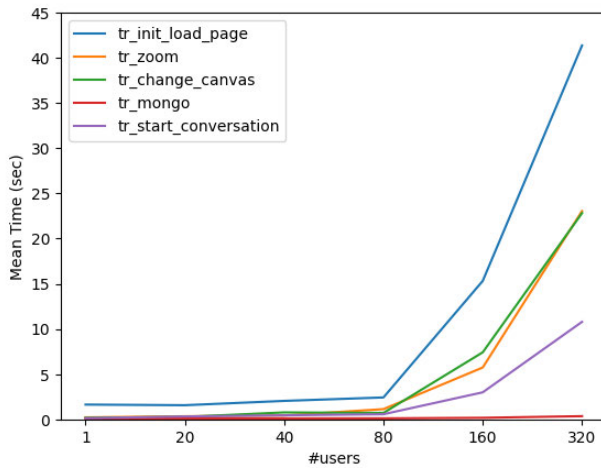


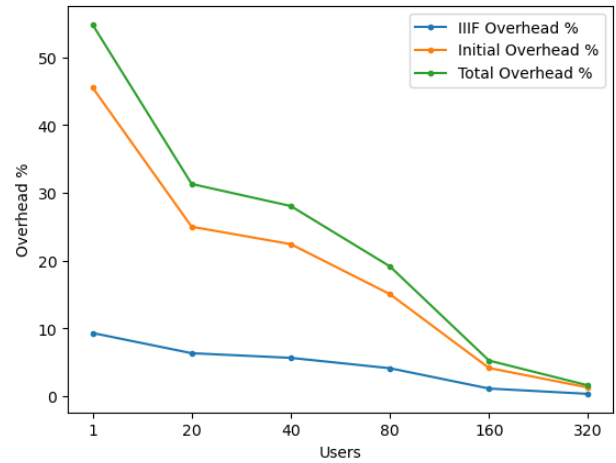
FIGURE 12. Mean page loading time, measured in the last 10 seconds, compared to the number of connected users.

and Cleopatra. Because of the higher values of connected users, and the longer persistence of such values, the mean page time measured in the last 10 seconds is longer in the case of the Cleopatra experiment. Such higher overload is more frequent in the Cleopatra experiments because of the slower init phase, which delays the users' sessions and fills the queues of the web server.

To better investigate how the different interactions within a session are affected by the increasing workload, requests have been grouped into transactions. An *transaction* is defined as a group of requests that satisfy a user's interaction with the viewer. The *tr\_init\_load\_page* corresponds to the download of the required HTML code and libraries for running the viewer and the Cleopatra client. The *tr\_canvas* transaction corresponds to the change of the image shown by the viewer and, as well *tr\_zoom*, the download from the image server of the tiles required for the visualization of image details. The transactions *tr\_mongo* and *tr\_start\_conversation* correspond



(a) Degradation of Cleopatra transactions.



(b) Cleopatra overhead over the IIIF workload.

**FIGURE 13. Degradation of performance.**

respectively to the requests sent to the Rest interface of the IIIF Session Handler and to the messages exchanged through the web socket of the NLP agent.

In Figure 13 we report the average response time of the different transactions which occur in a user's session.

How it can be observed in both Figure 13(a), the response time does not get worse until there are 160 users simultaneously connected. Because of the queued requests, the loading page and IIIF transactions overload the web and the image servers which need to deliver the related response. On the other hand, even if such limitation can be removed by changing some software modules and their configuration, we are interested in observing that the event notification does not affect global performance. Response time of *tr\_mongo*, for the events notification to the Session Handler, slightly increases linearly, but their response time is always negligible. In fact, either they are served by a different web service and the number of simultaneous transactions is low, because they are fast and executed asynchronously, or their arrival is slowed down by delayed requests.

Finally, in Figure 13(b) an analysis of the overhead introduced by Cleopatra over the standalone IIIF viewer is shown. In this analysis, two different contributions to the total overhead are distinguished. The first contribution represents the initial overhead due to the increased number of downloaded libraries, while the second contribution encompasses all the additional requests sent during the actual use of the Cleopatra application. The overhead is calculated as  $ov\% = \frac{MORT}{MTRT}$ , where *MORT* is the *Mean Overhead Requests Time*, which contains the sum of the average times of the requests that produce overhead in a session (in other words it is the sum of all requests that are not present in the Mirador baseline) and *MTRT* is the *Mean Total Requests Time* which contains the sum of the average times of all requests in a session. Both values are multiplied by the number of completed requests. The initial overhead is instead calculated by considering at

the numerator only those requests which are required during the loading phase of the web page, while the IIIF overhead is calculated by using at the numerator only the requests made during the actual use of the framework.

### E. KPIs EVALUATION

In order to verify that the application responds with satisfactory service levels and to estimate the number of users which can be served simultaneously, three widely used KPIs for measuring the performance indices of a web application have been selected [27]. In Table 3 the computed values and the reference thresholds are shown.

The *Speed Index (SI)* metric estimates how quickly the content is displayed during webpage loading. It is considered good if less than 3.4 seconds while it is considered moderate between 3.4 and 5.8. The computed value is more than acceptable for 160 simultaneous users and must be improved when reaching 320 simultaneous users.

The *effectiveness* indicates, as a percentage, the number of tasks successfully completed relative to the number of total tasks. In our case, we have identified a user's entire session as a task. The measured values indicated that the system cannot satisfy all requests of some sessions when it reaches the peak load.

The *Time Overhead* is computed as the ratio between the total time used to satisfy the HTTP requests which are performed either with or without the storyteller. We observe that is limited and decreases with the number of users.

From these results, we can conclude that the evaluated deployment configuration is not able to serve more than 100 simultaneous users when it reaches the load peak. However, the storyteller introduces a limited overhead that does not affect the service level, but it decreases with the number of users. Such consideration allows the implementation of a straightforward solution to scale up the system, e.g. through

TABLE 3. KPI evaluation.

KPI	160 users	320 users
SpeedIndex good:<3.4; moderate: <5.8	2.81s	6.45s
Effectiveness	97.5%	94%
Time Overhead	6%	2.3%

a horizontal scale-up of the web server instances, which is dynamically supported by any Cloud platform.

### VIII. CONCLUSION AND FUTURE WORKS

The spread of standards for the delivery of digital content through the web allows for the development of innovative services to enhance their presentation and their exploitation. In this paper, we proposed general mechanisms that allow for providing context awareness to intelligent agents which exploit the IIIF standard for assisting the fruition of a digital repository through interactive storytelling. We described how this knowledge has been exploited to integrate assistive technologies supporting the fruition of cultural content through the web. The knowledge about the IIIF session improved the NLP capability of a smart assistant, to use the IIIF APIs for coordinating the presentation of cultural contents and the presentation of the related media and to implement decision-making strategies that can take into account information about the global fruition of observed archives. Experimental results, performed by a prototype implementation of an agent bases platform developed by the Cleopatra project, have been used to demonstrate that the overhead introduced by this extension to an IIIF standard technology is negligible and affects the performance only when the system is overloaded by the utilization of the IIIF contents.

The proposed solution can improve the impact, in terms of presentation and communication of digital archives, but also of physical museums, augmenting the user's visit. Moreover, intelligent agents can foster the delivery of semantically related content from secondary sites, which are usually out of touristic itineraries, or digital libraries to augment the cultural value and to recommend alternative itineraries. In such scenarios, the dissemination of knowledge, based on the digital exchange of cultural content among museums and cultural sites could be also regulated by well-defined innovative business models.

Nevertheless, the users have been not involved in the evaluation loop yet, preliminary positive feedback has been collected and taken into account in the design and development process, presenting the intermediate prototype implementation at two summer schools in Archaeology on "The Via Appia in Campania: knowledge, conservation and valorization" in 2020 and in 2021.

In the future developments of this work, the framework will be implemented in a real museum scenario of small dimensions to test the system in a less controlled but still

limited environment to compare the results obtained thanks to the simulation mechanisms with those of the real scenario. This will allow the development of further considerations which will certainly be useful for the improvement of the system and its technological and methodological evolution.

### REFERENCES

- [1] A. Amato, B. D. Martino, M. Scialdone, and S. Venticinque, "Adaptive recommendation to dynamically changing profiles for delivery of ubiquitous services," *Int. J. Comput. Sci. Eng.*, vol. 13, no. 4, p. 322, 2016.
- [2] A. Amato, R. Aversa, D. Branco, S. Venticinque, G. Renda, and S. Mataluna, "Porting of semantically annotated and geo-located images to an interoperability framework," in *Proc. Comput. Intell. Secur. Inf. Syst. Conf.*, vol. 497, 2022, pp. 508–516.
- [3] V. Apte, T. V. S. Viswanath, D. Gawali, A. Kommireddy, and A. Gupta, "AutoPerf: Automated load testing and resource usage profiling of multi-tier internet applications," in *Proc. 8th ACM/SPEC Int. Conf. Perform. Eng.*, New York, NY, USA, Apr. 2017, pp. 115–126.
- [4] E. E. Boros, A. Toumi, E. Rouchet, B. Abadie, D. Stutzmann, and C. Kermorvant, "Automatic page classification in a large collection of manuscripts based on the international image interoperability framework," in *Proc. Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2019, pp. 756–762.
- [5] C. Brady, "Run-time discovery, selection, composition & invocation of web services using semantic descriptions," Ph.D. dissertation, Dept. Comput. Sci., Trinity College Dublin, Dublin, Ireland, Sep. 2004.
- [6] P. J. Emanuel, "Stitching together technology for the digital humanities with the international image interoperability framework (IIIF)," in *Digital Humanities, Libraries, and Partnerships*, R. Kear and K. Joranson, Eds. Oxfordshire, U.K.: Chandos Publishing, 2018, pp. 125–135.
- [7] N. Freire, G. Robson, J. B. Howard, H. Manguinhas, and A. Isaac, "Cultural heritage metadata aggregation using web technologies: IIIF, sitemaps and Schema.org," *Int. J. Digit. Libraries*, vol. 21, no. 1, pp. 19–30, Mar. 2020.
- [8] A. Hawkins, "Archives, linked data and the digital humanities: Increasing access to digitised and born-digital archives via the semantic web," *Archival Sci.*, vol. 22, no. 3, pp. 319–344, Sep. 2022.
- [9] C. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 8, May 2014, pp. 216–225.
- [10] IIIF Consortium. (2023). *International Image Interoperability Framework*. Accessed: Feb. 13, 2023. [Online]. Available: <https://iiif.io>
- [11] S. Jiang, M. Metzger, A. Flanagin, and C. Wilson, "Modeling and measuring expressed (dis)belief in (mis)information," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 14, May 2020, pp. 315–326.
- [12] K. A. Kulakov, O. B. Petrina, and A. A. Pavlova, "Smart service efficiency: Evaluation of cultural trip planning service," in *Proc. 19th Conf. Open Innov. Assoc. (FRUCT)*, 2016, pp. 106–112.
- [13] S. A. Long, "Review: International image interoperability framework (IIIF); Gallica; e-codices: Virtual manuscript library of Switzerland," *J. Amer. Musicolog. Soc.*, vol. 71, no. 2, pp. 561–572, Aug. 2018, doi: 10.1525/jams.2018.71.2.561.
- [14] O.-M. Machidon, A. Tavčar, M. Gams, and M. Duguleană, "CULTURALERICA: A conversational agent improving the exploration of European cultural heritage," *J. Cultural Heritage*, vol. 41, pp. 152–165, Jan. 2020.
- [15] A. C. Marosi, R. Lovas, A. Kisari, and E. Simonyi, "A novel IoT platform for the era of connected cars," in *Proc. IEEE Int. Conf. Future IoT Technol. (Future IoT)*, Jan. 2018, pp. 1–11.
- [16] K. Michalakis and G. Caridakis, "Context awareness in cultural heritage applications: A survey," *J. Comput. Cultural Heritage*, vol. 15, no. 2, pp. 1–31, Jun. 2022.
- [17] E. Muratović and I. Prazina, "A web-based service-oriented solution for a mobile digital storytelling application," in *Proc. 21st Central Eur. Seminar Comput. Graph.*, 2017, pp. 1–12.
- [18] L. Nisiotis, L. Alboul, and M. Beer, "Virtual museums as a new type of cyber-physical-social system," in *Augmented Reality, Virtual Reality, and Computer Graphics*. Cham, Switzerland: Springer, 2019, pp. 256–263.
- [19] L. Nisiotis, L. Alboul, and M. Beer, "A prototype that fuses virtual reality, robots, and social networks to create a new cyber-physical-social ecosystem for cultural heritage," *Sustainability*, vol. 12, no. 2, p. 645, Jan. 2020.

- [20] F. Peinado, "Interactive digital storytelling: Automatic direction of virtual environments," *Upgrade*, vol. 7, pp. 42–46, Jan. 2006.
- [21] J. A. Raemy, "The international image interoperability framework (IIIF): Raising awareness of the user benefits for scholarly editions," Ph.D. thesis, Dept. Inf. Sci., Haute école de gestion de Genève, Carouge, Switzerland, 2017.
- [22] G. Renda, S. Gigli, A. Amato, S. Venticinque, B. Martino, and F. R. Cappa, "Mobile devices for the visit of 'anfiteatro campano' in Santa Maria Capua Vetere," in *Progress in Cultural Heritage Preservation*, D. Fritsch, J. Leissner, R. Davies, F. Remondino, and R. Caffo, Ed. Berlin, Germany: Springer, 2012, pp. 281–290.
- [23] A. Salarelli, "International image interoperability framework (IIIF): A panoramic view," *JLIS.it*, vol. 8, no. 1, pp. 50–66, Jan. 2017. [Online]. Available: <https://www.jlis.it/index.php/jlis/article/view/385>
- [24] G. Salton and M. E. Lesk, "Computer evaluation of indexing and text processing," *J. ACM*, vol. 15, no. 1, pp. 8–36, 1968.
- [25] R. S. Snyderman and T. Cramer, "The international image interoperability framework (IIIF): A community & technology approach for web-based images," in *Proc. Archiving Conf.*, Los Angeles, CA, USA, 2015, pp. 1–10.
- [26] S. Varitimias, K. Kotis, A. Skamagis, A. Tzortzakakis, G. Tsekouras, and D. Spiliotopoulos, "Towards implementing an AI chatbot platform for museums," in *Proc. Int. Conf. Cultural Informat., Commun. Media Stud.*, Jan. 2020, pp. 1–4.
- [27] C. Xilogianni, F.-R. Doukas, I. C. Drivas, and D. Kouis, "Speed matters: What to prioritize in optimization for faster websites," *Analytics*, vol. 1, no. 2, pp. 175–192, Nov. 2022.
- [28] S. Yalovitsyna, V. Volokhova, and D. Korzun, *Smart Museum: Semantic Approach to Generation and Presenting Information of Museum Collections*. Hershey, PA, USA: IGI Global, Jan. 2020, pp. 236–255.



**ALBA AMATO** is currently a Researcher in computer science with the University of Campania "Luigi Vanvitelli." She has involved in research activities dealing with digital humanities. She is the author of more than 50 publications in international journals, books, and conferences, in collaboration with national research organizations and foreign academic institutions.



**SALVATORE VENTICINQUE** has been an Associate Professor with the University of Campania "Luigi Vanvitelli," since 2006. He was a lecturer of computer programming and computer architecture in regular academic courses. He has involved in research activities dealing with parallel and grid computing and mobile agents programming for distributed systems. He is the author of more than 100 publications in international journals, books, and conferences, in collaboration with national research organizations and foreign academic institutions. He has participated in research projects supported by international and national organizations.



**DARIO BRANCO** is currently pursuing the Ph.D. degree with the University of Campania "Luigi Vanvitelli." He is also a Research Fellow with the University of Campania "Luigi Vanvitelli." His scientific activity was documented by publications made for conferences and journals. He has participated in the Cleopatra National Research Project and the Greencharge European Research Project, in collaboration with foreign research institutes and industrial partners. His main research interests include the use of the agent paradigm applied to innovative solutions, intelligent distribution of electrical loads through simulation techniques, and conjunction with the use of machine learning techniques.



**ROCCO AVERSA** is a Full Professor of computer science with the University of Campania "Luigi Vanvitelli." He has participated in various research projects supported by national organizations and EU in collaboration with foreign academic institutions and industrial partners. His research interests include the use of the mobile agents paradigm in the distributed computing, the design of simulation tools for performance analysis of distributed systems, and the project of innovative middleware software to enhance cloud computing platforms. Such scientific activity is documented on scientific journals and international and national conference proceedings. He is an Associate Editor of *International Journal of Web Science*.

• • •

Open Access funding provided by 'Università degli Studi della Campania "Luigi Vanvitelli"' within the CRUI CARE Agreement