RECBOT: Virtual Museum navigation through a Chatbot assistant and personalized Recommendations

KONSTANTINOS TSITSEKLIS, National Technical University of Athens, Greece
GEORGIA STAVROPOULOU, National Technical University of Athens, Greece
ANASTASIOS ZAFEIROPOULOS, National Technical University of Athens, Greece
ATHINA THANOU, National Technical University of Athens, Greece
SYMEON PAPAVASSILIOU, National Technical University of Athens, Greece

The trend for digitalization of museums has been on the rise in recent years, as museums seek to make their collections and exhibitions more accessible to a wider audience. This has involved the use of technologies such as virtual and augmented reality, online exhibits, and digital archives. These digital initiatives have allowed museums to reach new audiences and provide immersive experiences that enhance visitors' engagement with the exhibits. Following this trend, in the current work, we propose a conversational agent that assists remote visitors in accessing a museum's collection. The proposed architecture includes a chatbot for user interaction that employs Natural Language Processing techniques for understanding the user's input. To increase visitor engagement, a hybrid recommender system is developed that combines content-based and collaborative-filtering components. The available data is modeled in the form of a Knowledge Graph, which allows for useful insights to be extracted from it.

CCS Concepts: • **Computing methodologies → Natural language processing**; *Machine learning*; • **Information systems → Recommender systems**.

Additional Key Words and Phrases: virtual tour, conversational agent, chatbot, recommender system, Natural Language Processing, online museum

ACM Reference Format:
Konstantinos Tsitseklis, Georgia Stavropoulou, Anastasios Zafeiropoulos, Athina Thanou, and Symeon Papavassiliou. 2023. RECBOT: Virtual Museum navigation through a Chatbot assistant and personalized Recommendations. In *PATCH '23: Workshop on Personalized Access to Cultural Heritage, June 26–29, 2023, Limassol, Cyprus.* ACM, New York, NY, USA, 15 pages. https://doi.org/XXXXXXX

1 INTRODUCTION

In recent times, the advancements in information and communication technology (ICT) have brought about new opportunities for remotely accessing cultural heritage information and services. The significance of this progress has been magnified by the COVID-19 pandemic and the ensuing quarantines, which have led to a surge in demand for virtual services, including those that allow access to cultural heritage sites [1, 15]. Many cultural heritage organizations have embraced this trend and now offer navigation services in both physical and virtual modes. To facilitate such services, conversational agents have been developed as one of the most common techniques to provide real-time guidance to end-users and enhance their experience while using online services. These agents not only grant access to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Association for Computing Machinery.
Manuscript submitted to ACM

the site's content but also offer recommendations and collect information about the user's preferences and actions, which can be utilized for further analysis and improvement of the provided services.

In this work, we introduce a conversational agent developed for the provision of online navigation services in the Museum of Paleontology and Geology in Athens. This agent is in the form of a chatbot that serves as an interface between the museum's system and its visitors. During the visitor's online navigation of the museum, the chatbot provides recommendations for exhibits and related multimedia information that the visitor can view. The agent takes advantage of Natural Language Processing (NLP) techniques that employ Named Entity Recognizer (NER) components trained on synthetically generated data to help interpret the user's input in either English or Greek. Additionally, we present a Recommender System (RS) that offers suggestions for relevant exhibits at each step of the user's navigation through the museum's collection. The RS uses a hybrid method combining content-based and collaborative filtering approaches. The conversational agent's functions are supported by a Knowledge Graph (KG) database that stores the relevant data.

The remainder of this paper is structured as follows. In Section 2 we briefly discuss some related works on chatbots, NER and recommender systems and highlight their relation to the proposed approach in this manuscript. In Section 3 the system's architecture is presented and the operation of each component is explained in detail. Additionally, in Section 4 we provide evaluation results on the NER and RS components. Finally, Section 5 concludes the paper and shortly refers to plans for future work.

2 RELATED WORK

2.1 Chatbots and Named Entity Recognition

The idea of using Chatbots as tools for enhancing the visitors' quality of experience (QoE), when visiting museums or other areas of cultural interest, has been around for a while. In 2004, Max, a conversational agent, was installed in the HNF museum in Germany [14], where visitors could make queries through a keyboard and the system would generate an appropriate answer based on predefined rules. Another example is CulturalERICA [17], which is interconnected with the Europeana database [10], and its conversational agent matches specific intents with the context of the conversation. CulturalERICA is developed based on the Google's DialogFlow [9] system. Chatbots have also been used for navigation in cities with rich cultural backgrounds, such as the proposed chatbot for Naples [23], which uses an encoder-decoder scheme to generate responses. The encoder maps the text in points of a vector space, and the decoder uses a Recurrent Neural Network (RNN) to generate a response. To increase visitor satisfaction in museums, three different conversational agents were designed for the Korean National Museum [19], with different appearances and language styles to cover various demographics, all using a knowledge database about Korean culture to answer user inputs.

Chatbots are taking advantage of NER techniques to identify and extract relevant information from user inputs, such as names, dates, locations, and other entities. By using NER, chatbots can provide more personalized and accurate responses to users' queries. In [13], a chatbot for stock information queries is developed focused on entity extraction. To this end, Rasa Natural Language Understanding (NLU) and Neural Network (NN) experimental comparison and evaluation is performed, revealing that Rasa NLU, using a conditional random field classifier (CRF), outperforms in accuracy the NN approach, with one-hidden layer and an additional representation layer, for a single experiment. On the other hand, NN behaves in a more robust way, regarding the entity classification from segmented words. The researchers in [21] explore NLP improving techniques for building chatbots through NER integration. Towards this goal, the Stanford CoreNLP framework is used to extract default named, numerical and temporal entities, with the capability of extension by adding regular expression annotator to recognise entities like email, URL, etc. Another NER model combined with intent classification is described in [2] for business domains, based on Neural Network Architecture and a knowledge base with manually annotated examples, used to train the NN.

By considering the available studies, in the proposed work in this manuscript we develop NER components, both in English and in Greek, in order to extract entities related to museum's exhibits from user's messages. To the best of our knowledge, pre-trained datasets in the paleontological domain focusing on such purposes do not exist, so we address this problem by automatically generating datasets in both languages to train the NER components. We enable the recognition of domain specific entity types, achieving high performance over text provided in English or Greek.

2.2 Recommendation Systems for Cultural Content

Nowadays, recommender systems (RS) are ever present, ranging from systems operating in Online Social Networks (OSNs) to systems that aid users in their visits to cultural heritage sites. The importance of RS is further highlighted by their constantly increasing market share, with forecasts estimating a 29.8% increase in the U.S. market share for the years 2021-2028 [20]. RS are useful tools that help users discover new content according to their tastes and preferences. The authors in [24] develop a virtual assistant capable of working in museums that are part of Google Arts & Culture project. They propose two modes of recommending exhibits to a visitor of a museum. Both of them are content based, with the first method recommending items similar to an implicitly extracted user profile vector, while the second one recommends exhibits based only on the exhibit-exhibit similarity. In [18] the semantics between exhibits are taken into account. In detail, by examining certain aspects of exhibits such as their age, region, artist, etc., influence relationships between the exhibits are derived. The RS utilizes these directed relationships to calculate two scores (aggregation and influence) and provide recommendations to the user. This approach is inspired by the HITS algorithm presented in [3]. In our system there is no need for any specific input from the user, contrary to [18], where the user must explicitly declare their area of interest.

In [22], a mobile recommendation system is created to assist users who are physically present at a cultural site. The system relies on Resource Description Framework (RDF) semantics and is based on an ontology. A profile is established for each user, which is updated using explicit information provided by the user as input, as well as implicit information such as their location and feedback on exhibits. To suggest objects to users, clustering methods are used to group objects based on the user's "perspective". To enhance the diversity of recommendations, objects belonging to multiple perspectives are recommended. The paper's key finding is that using ontologies to represent objects results in increased recommendation accuracy. In our work, we adopt and exploit this conclusion by using a knowledge graph to organize items and relevant information.

Focusing on augmented reality applications, the authors in [6], develop a RS to enhance the Quality of Experience (QoE) of a museum visitor. The system utilizes machine learning techniques to establish rules for suggesting exhibits to a user. The system can provide recommendations in various modes, such as suggesting nearby exhibits, recommending exhibits along a specific path or in a "surprise me" manner. In [25], the authors propose a RS for creating museum itineraries based on data collected from Internet-of-Things (IoT) devices (sensors) and user opinions in social media. To extract meaningful insights from the collected data, the authors employ semantic analysis, emphasizing the importance of incorporating semantics when designing RS for cultural heritage applications. Towards the goal of creating interesting itineraries for visitors of museums, the authors in [11] present the key components for a recommendation system that tracks museum visitors while respecting their anonymity, and uses facial feature analysis to suggest exhibits similar to their preferences. The aim is to create engaging itineraries for museum visitors by recommending exhibits based on their reactions and preferences. The researchers in [16] create a personalized tour RS for the Museum of Zoology in London using a content-based approach. Similar to the framework presented in this paper, the system assumes no prior knowledge about the visitor. After the user has viewed three exhibits, the system recommends exhibits that are similar to their features based on cosine similarity. The authors conducted a survey and gained useful insights, including the importance of addressing the curiosity of users regarding how the recommendations are generated. In our work, we tackle this issue by providing explainable recommendations.

3 SYSTEM DESCRIPTION

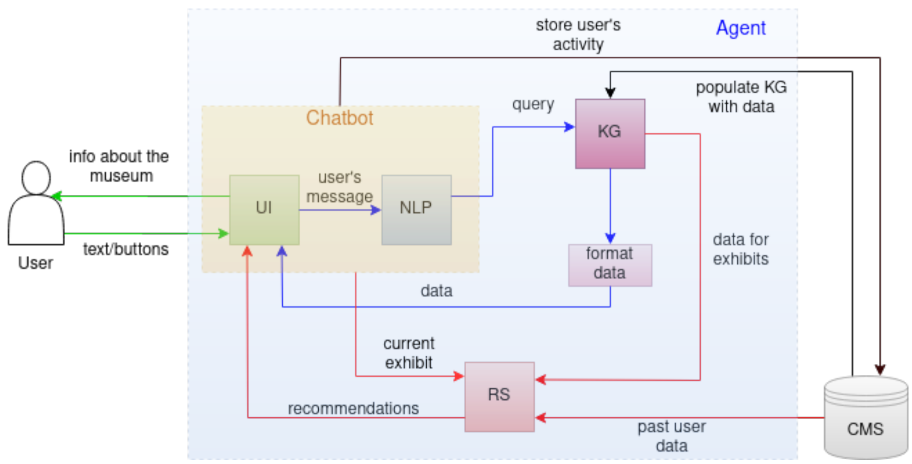

Fig. 1. Agent Architecture

In this Section, we describe the overall operation of the proposed framework. This framework consists of a conversational agent in the form of a chatbot that interacts with the user (museum's visitor). The visitor interacts with the agent's user interface (UI), included in the chatbot component, either by selecting one of the options presented in the form of buttons or through the provision of text. If the provided message is in a typed text format, it is firstly processed by the NLP component to extract knowledge about the requested item or its characteristics. In all cases, a suitable query is constructed and is directed to the KG that hosts the available information for the museum's content. The items that match the request are firstly transformed to the correct form and then provided to the user in a suitable and customized answer through the UI. The user's activity, concerning which exhibits have been seen so far, is stored in a separate Content Management System (CMS) to boost the recommendation system component.

The RS provides the user with recommendations for relevant exhibits as a short list of recommendations relevant to the exhibit that the user is currently viewing. In order for the RS to operate, several kinds of data are required. First of all, data regarding the current exhibit that the user is viewing are retrieved from the chatbot. Moreover, information concerning the museum's collection is necessary and is retrieved from the KG through suitably formed queries and finally, the activity of users that have visited the museum in the past is fetched through suitable queries to the CMS. The overall architecture of the agent can be seen in Fig. 1. In the following subsections, the operation of each component is explained in detail.

3.1 Available data and formation of Knowledge Graph

Despite the fact that the CMS that can be seen in Fig. 1 is a separate system from the conversational agent, it is nevertheless, a necessary one since it stores all the data required for its operation. This data specifically refers to the museum's exhibits (e.g., name, location discovered, body part, animal type, natural habitat of specimen, etc.), the thematic areas in which the collection is organized, the educational programs that act as guided tours for the corresponding thematic areas and finally, educational games mostly targeted to the younger audience.

From this pool of data, the most relevant to our application are periodically retrieved and stored in the form of a KG. A KG is a way of representing the knowledge we have about a specific area in a clear and structured way. The nodes of the graph denote entities or concepts and the edges represent the relationships among them. Knowledge graphs are widely used in recommender systems [7]. The structure of the KG enables the semantic reasoning between exhibits by making it easier to compare them. The schema of the KG including the basic properties of the nodes can be seen in Fig. 2.

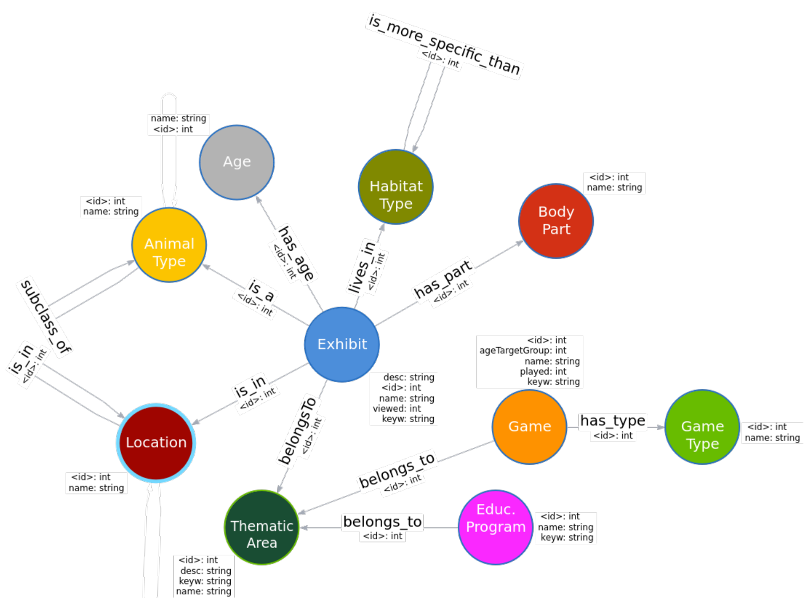

Fig. 2. Knowledge Graph Schema Visualization

The most central entity of our system are the exhibits of the museum. These nodes have properties denoting their name, a short description, their popularity and keywords provided by experts. Based on the content stored in the CMS, the exhibits are further detailed, in order to extract relationships and similarities among them. Thus, the new entities that are created are:

- *Animal type*: Denotes the type of animal of the exhibit. Animal types are hierarchically organized from the more specific to the more generic type. An example of this mini taxonomy is that a mosasaur is a reptile and a reptile is a vertebrate animal. By having this structure in the KG we can extract similarities, e.g., a turtle is a reptile. This means that there is now a path of length 2 in the KG connecting the entities turtle and mososaur.
- *Body Part*: The exhibit's corresponding body part.

- *Location*: The geographical location where the exhibit was discovered. The location entities have relationships among them denoting connections between more specific regions (e.g., Pikermi) to larger ones (e.g., Athens).
- *Habitat Type*: The habitat where the corresponding animal lived (e.g., lake, land, mixed, etc). Habitats follow a hierarchical organization logic similar to the *location* and *animal type* entities.
- *Age*: The paleontological period in which the exhibit belongs.
- *Thematic Area*: Entities representing a semantic classification of exhibits. Each exhibit is assigned to exactly one thematic area.
- *Games*: Educational games. Each game is related to a single thematic area and is associated with *Game Type* entities that reveal the kind of the game (e.g., puzzle, quiz, etc).
- *Educational Programs*: Each program is related to one thematic area.

3.1.1 Knowledge Graph Embeddings.

The term graph embedding refers to a mathematical process through which the nodes of a graph are translated in low dimension vectors, or, in other words, in points of a geometric space of low dimensionality. In the case of KG embeddings, both the entities (i.e., nodes) and the relationships (i.e., edges) of the graph are assigned coordinates. During the past years, various algorithms have been proposed for the task of KG embedding [8]. For the purposes of this work, we compare three different models, namely, TransE [5], DistMult [27] and ComplEx [26]. TransE is one of the first popular algorithms designed for KG embedding. For each vector corresponding to an element of a triad of the form (h, r, t), where h refers to the "head" entity, r to the relationship type and t to the "tail" entity, the algorithm tries to achieve $\overrightarrow{\mathrm{h}} + \overrightarrow{\mathrm{r}} \approx \overrightarrow{\mathrm{t}}$. TransE is intuitive but suffers from the fact that it cannot handle more complex types of relationships such as 1-N or N-1 relationships. DistMult is also a popular algorithm that employs the bilinear dot product to model interactions between entities. Contrary to TransE, DistMult can handle more complex relationships. Finally, ComplEx is an extension of DistMult that employs complex vectors for the representation of entities and relationships. Comparatively to both TransE and DistMult, ComplEx can represent more complex relationship types and thus models more accurately intricately structured KGs at the cost of greater computation time. In this work, the KG embeddings are employed in order to discover similar exhibits through simple computations.

3.2 Natural Language Processing

NLP techniques aim in enhancing the user and computer interaction and therefore are an integral tool in the conversational agents' development. One of the main goals of NLP integration in a chatbot's operation is the analysis and interpretation of the user's messages and the generation of suitable and explainable answers. In order for the agent to understand the user's input, techniques such as Intent Classification (i.e. association of the user's input to predetermined set of intents) and NER (i.e. classification of named entities mentioned in unstructured text into pre-defined categories) ought to be applied to the user's message.

We are utilizing the DIETClassifier, which is a multi-task transformer architecture developed by Rasa Software [4], to classify messages into specific intents that are related to the museum's services. This architecture is capable of performing both intent classification and Named Entity Recognition (NER). To train the model, we provided various examples of potential user messages corresponding to each intent, allowing the model to recognize the hidden intent presented in many different and heterogeneous ways, and generalize to unseen input data. One of the main advantages of the DIETClassifier is that it can effectively classify messages in multiple languages, as long as examples for each language are provided under the appropriate intent. This allows us to deploy the agent in both English and Greek.

The DIETClassifier can handle simple user inputs for NER, as long as the provided examples for intent classification contain easily recognizable annotated entities like games, educational programs, and thematic areas. However, more complex user messages that request for exhibits by their characteristics which can be classified into various types of entities require a tailor-made approach. In order to achieve this, a customized method was created to extract information from the user's input, allowing for relevant queries to be performed on the KG to find requested items based on their explicit characteristics or other related attributes. It should be noted that pre-trained NER models struggle to identify entities in specific domains like paleontology, resulting in poor and inadequate performance. Additionally, the lack of paleontological data, especially in the Greek language, further complicates the issue, as the accuracy of NLP models heavily relies on having large amounts of relevant and diverse data. To overcome these challenges, NER components trained on artificially generated data were developed specifically for our use cases and targeted towards both languages. The NER has been trained to recognise entities related to museum's exhibits in a user's message, after it has been classified into the "see specific exhibit" intent. This denotes that the user asks to see exhibits by giving some of their characteristics (e.g., their animal type, their habitat, their age, the geographic location where they were found and/or their displayed body part in the museum).

In order to train the NER component, examples of the way a visitor could pose a request to see an exhibit by providing some of its characteristics, have been fed to the model, in both languages. However, instead of providing the explicit name of the characteristic, placeholders have been used and lists of different values that could replace them, so that the lack of data could be tackled and lots of training data could be generated (e.g., I would like to see % s of % s from % s -> I would like to see antlers of deer from Crete, I would like to see skulls of bears from Pikermi, etc). With the help of spaCy [12], a software library for NLP, these words have been annotated as entities of specific type (Animal type, Body Part, Location, Age, Habitat type) and the generated data have been divided into training, test and validation sets. The test set that was used to evaluate the custom trained component consisted both of heterogeneous forms of sentences and lists of entity values that the model had not seen during training, and forms of data more familiar to the model. Comparisons and results of the model are demonstrated in detail in the Evaluation Section 4.

3.3 Chatbot

3.3.1 Overview. The visitor of the museum interacts with the conversational agent, through a UI. The chatbot supports four functionalities related to the virtual museum tour (main options). Initially, the user can select in which language (English or Greek) wants to navigate through the museum and then the main options are displayed to them (Fig. 3).

The user's interaction with the chatbot is via buttons or text by either selecting one of the choices displayed or typing a message. Buttons support the complete navigation of the user through the museum's functionalities.

Concerning the textual case, the agent classifies the user's message into the following intents:

- 360 museum tour
- Games: either through provision by the user of specific attributes (e.g., the thematic areas or related words about the games) or not (e.g., I want to play a game/ Puzzle game/ I would like to play a game about the mesozoic era).
- Exploration of exhibits: either through provision by the user of specific attributes (e.g., the thematic areas or related words about the exhibits) or not.
- Educational programs: either through provision by the user of specific attributes (e.g., the thematic areas or related words about the educational programs) or not.
- See specific exhibit: request exhibits with specific attributes.

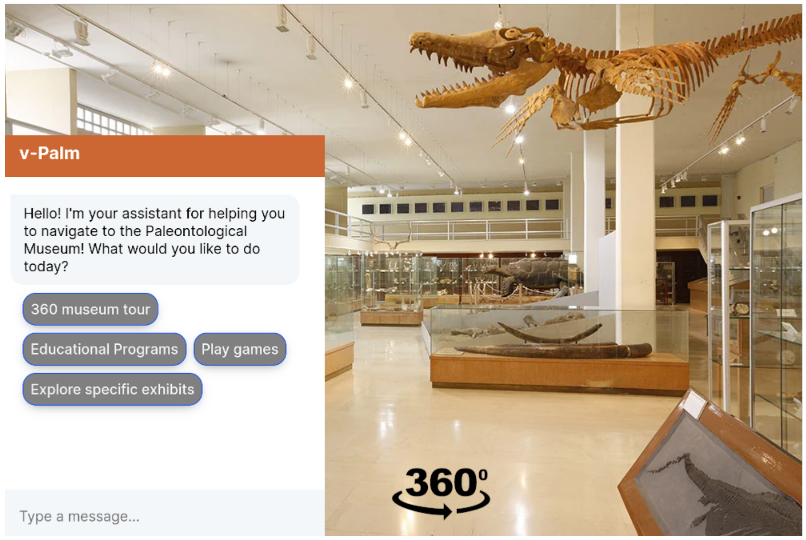

Fig. 3. Main Navigation Options.

- Provide a thematic area of interest: the agent requests clarification of what the user wants to do concerning this thematic area.

Additionally, the agent supports secondary functionalities (e.g., Greet, Say Goodbye, Ask for help, Change language).

The agent responds to the user's message by presenting the available options that match the user's request via buttons and the user can navigate further by clicking them, until they find the desirable item. In the case of exhibit request, relevant and personalised recommendations are displayed to the user, afterwards. In case the agent does not understand or cannot handle the request of the user, the main options are displayed, through their corresponding buttons, as a smooth fallback mechanism.

3.3.2 Message Processing and interaction with the Knowledge Graph. Once the user's message is classified into an intent (Section 3.2), it is processed by NLP techniques (Section 3.2). Entity recognition is firstly applied to the user's input by the DIETClassifier in some simple cases that the user does not request specific exhibits. In case no entities are extracted and since the user's message might incorporate keywords related to the desirable item, parts of the user's message (that possibly include keywords) are compared to names or keywords related to the museum's items, by performing queries to the KG (section 3.1). Comparison through trained word vectors was not applicable in our domain specific case, due to the lack of data in the field. To overcome this limitation, nominal chunks are extracted from the user's text with the help of spaCy. spaCy's pre-trained Greek and English tokenizer and POS tagger components extract the chunks and their lowercased lemmas are compared to the lowercased lemmas of node names of the KG or their properties, by utilising a pre-trained spaCy lemmarizer. This is because the form and grammar of the user's message segment might differ from the form the items are stored in the KG. It is worth mentioning that Greek lemmatizers perform poorly in our domain specific case, so in case the interaction of the agent with the user is in Greek or the English lemmatizer fails matching the user's request to a museum's item, the edit distance of the words is calculated. If it is found relatively small, the correspondent item from the KG is displayed to the user. In case the user provides input via buttons, a query to the KG is directly performed without the need of data processing through NLP techniques.

If the user requests, via textual input, a specific exhibit referring to some of its characteristics, the user's message is classified into the "see specific intent" and the custom trained NER component performs entity recognition to the text in order to match mentioned characteristics to predefined entities. The extracted entities are then compared to the KG nodes, through the lemmatizing and edit distance techniques mentioned earlier. In case no entities are extracted from the NER component, the nominal chunk technique is performed as mentioned above. The exhibits that have a relationship with the nodes matching the extracted entities, or the ones that have as properties keywords matching the user's request, are then demonstrated to the user accompanied by related and personalised recommendations.

3.4 Recommender System

The goal of the RS is to aid the users during their virtual tour by providing recommendations about exhibits that match their preferences and are similar to the ones already examined. In our case we suppose that the user sees an exhibit and then several recommendations are provided.

The RS has two key features. The first one is the explainability of the recommendations. This means that the provided recommendations should share at least a common attribute with the current exhibit (i.e., belong to the same thematic area, have the same body part, found in the same geographical region, etc). This gives the opportunity to generate intuitive explanations about the similarity of the proposed exhibits to the one the user is currently seeing on the screen. The other feature, is to exclude any exhibits that the visitor has already seen from future recommendation lists.

The RS consists of two major components. The first one is a content-based (CB) mechanism that detects similar exhibits to the one the visitor is currently viewing. The second mechanism is a collaborative-filtering approach (CF) that matches the visitor's behaviour to the actions of past visitors. In the following subsections the operation of the two components of the RS is presented in greater detail.

3.4.1 Content-based Mechanism. Since no profiles for the visitors of the museum are created, we assume zero-knowledge about the visitor. The CB mechanism relies exclusively on the similarities between the exhibits of the museum. These similarities are extracted through the process of KG embedding. In the embedded space relevant nodes are closer together than with unrelated ones. After the KG is embedded we consider the euclidean distance of the vectors (representing nodes) as a measure of their similarity (i.e, the shorter the distance the more similar the nodes). For the current exhibit being examined by the visitor, the k nearest exhibit nodes are discovered.

3.4.2 Collaborative-filtering Mechanism. Despite the lack of complete visitor profiles, we can store in the CMS the exhibits viewed by each visitor during their virtual visit. Then, we can extract similarities among the visitors by examining these vectors with a suitable metric such as the cosine similarity. After the visitor has seen a predefined (by the platform) number of exhibits, then the l most similar users are selected. The summation of their vectors gives us the most popular exhibits for these users. These exhibits are potential recommendations for the user.

The flowchart of Fig. 4 describes the overall operation of the RS. Each component always detects available (i.e., not recommended before) exhibits. Finally, after the lists are composed, any non-explainable recommendations are filtered out and recommendations are grouped, accordingly depending on the explanation. Each exhibit is assigned a score that can be used for selecting a subset of the recommendations in order to provide only a certain amount of recommendations to the user. This score s, for every exhibit i, is given by the following formula:

$$s_i = \alpha \cdot cb_i + \beta \cdot cf_i, \tag{1}$$

where α and β are parameters that measure the importance assigned to the two mechanisms, CB and CF, respectively. For these parameters it holds that: $\alpha + \beta = 1$.

Regarding the cb_i and cf_i parameters, they indicate the score of exhibit i as computed by the CB and CF mechanisms respectively. If i is not included in a list (CB or CF) then its score is equal to zero. If, on the other hand, i is included in the list returned by CB (or CF) then its score is equal to its position in this list sorted in an ascending way. In other words, the more similar the exhibit (or the more popular for the similar users for the CF mechanism) the higher its score.

It is possible that only non-explainable recommendations could occur. This could be because all the relevant exhibits have already been seen by the user or there is insufficient information about the exhibit. In this case, the RS checks the total popularity of the exhibits. Then, it presents to the visitor a short list of the most popular exhibits.

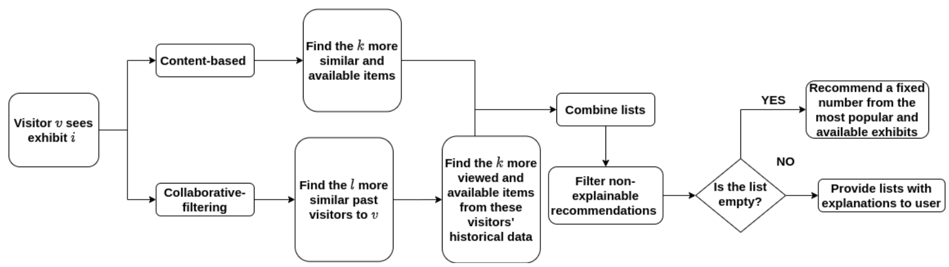

Fig. 4. Flowchart of the overall Recommender System

4 EVALUATION

In this Section, we evaluate the performance of the proposed system. First, we explain the setups which were used in the simulations and then, we discuss the performance of the NER mechanism. Finally, the operation of the RS is evaluated by employing specific metrics. All the developed algorithms are open-source and developed in Python3.

4.1 Experimental Setup

4.1.1 Recommender System. Concerning the employed dataset we can see some of its characteristics in Table 1 that presents the amount of instances of each entity type included in the KG and related to the Exhibit entity. As we can see, we can form 425 triplets of the form (h, r, t). These are split into train (360 triplets) , test (40 triplets) and validation (25 triplets) datasets and are given as inputs to a cross validation process to select the best hyperparameter setups for each one of the three selected embedding algorithms. Regarding the activity of past visitors, since we do not yet have access to real data, it was sampled at random, supposing we had 100 such users. For the collaborative filtering part of the recommendation algorithm we set the number of similar users, $l = 3$. In both the CB and the CF we search for the $k = 5$ more similar (and available) exhibits. For the purposes of the simulations we consider that a user picks at random an exhibit from the recommended ones.

4.1.2 NER training. For the NER components, we have generated custom training, validation and test data sets, both in English and in Greek as described in section 3.2. To this end, 32 sentence "casts" of the way a user could request for an exhibit by providing some of its characteristics (e.g., I would like to see a % s found in %s) have been created to generate the training, validation and part of the test data set (Data Set A) and 9 sentence "casts" separately for the test set (Data

Table 1. Attributes of KG

Exhibits	Thematic Area	Animal Type	Body Part	Habitat Type	Location	Age
52	8	45	16	30	46	7
Total Nodes:		204	**Total Relationships**:			425

Set B), both in English and in Greek. Similarly, lists of possible entity values have been created to fill the sentence "casts", leading to a sufficient amount of combinations and a big volume of generated data. It is worth noting that the data sets for Greek are larger in order for the model to train better on a language with more complicated grammar. The detailed numerical information depicting the number of sentences included in each dataset can be seen in Table 2.

Table 2. NER Datasets Info

	Training Set 75% of Data Set A	Validation Set 20% of Data Set A	Test Set Data Set B and 5% of Data Set A
English Data Set A	242926	64780	16195
English Data Set B	-	-	16996
Greek Data Set A	305879	81568	20392
Greek Data Set B	-	-	18856

The training of the NER components has been done with the help of SpaCy, using the available pipelines "en_core_web_lg" for English and "el_core_news_lg" for Greek. Early stopping parameters have been set, to avoid overfitting.

4.2 Custom trained NER component's Performance

The performance of the custom trained NER components in both languages is depicted in Table 3. From the F1-score (F), precision (P) and recall (R) presented both as a total score, and separately for each of the entities, we observe the high performance of the components. The method applied for synthetically generating data in order to train NER components for specific domain, has led to the efficient detection of entities inside possible user's messages, with English F1-score (96.46 %) being slightly better than the Greek F1-score (94.46 %). A similar behavior is observed for most of the other metrics, as well. This is expected, as NLP techniques have made remarkable progress mainly in English, while training in different languages still faces challenges. Nevertheless, this method generalizes well in both languages, indicating possible extension to more languages and more domains, if properly adapted.

Table 3. NER components' Performance

	English			Greek		
	Precision (P)	**Recall (R)**	**F1-Score (F)**	**Precision (P)**	**Recall (R)**	**F1-Score (F)**
Total	97.46	95.48	96.46	95.99	92.97	94.46
Animal Type	99.28	91.73	95.36	99.06	93.30	96.09
Habitat Type	97.02	96.73	96.87	98.61	95.73	97.15
Age	99.01	99.36	99.19	92.60	98.05	95.25
Body Part	96.62	94.98	95.80	92.55	83.89	88.01
Location	96.17	98.32	97.24	95.90	97.91	96.90

4.3 Recommender System Evaluation

4.3.1 Comparison of KG Embedding Algorithms. In this subsection, we perform a comparison of the three aforementioned KG embedding algorithms, namely TransE, DistMult and ComplEx. After the KG has been embedded we perform the following experiment. For each exhibit x in our collection we search for the 5 closest exhibits to it. Then, for each one of them, y, we examine if it has at least one common attribute with x (e.g., belong to the same thematic area, are from the same paleontological era, concern the same type of animal, etc). If this is the case, the recommendation of y after the visitor has seen x would be explainable and we could generate a smart textual description in order to nudge them to learn about y. If there exist no common attributes then this recommendation would be unspecified, meaning also that we cannot produce any reasoning about it without close examination of the entities x and y or without an expert's opinion. From Fig. 5 it can be seen that in all cases the amount of "Unspecified" recommendations is quite limited, thus strengthening our hypothesis that by embedding the formed KG we can detect relevant exhibits by employing simple similarity metrics (e.g., cosine similarity). Also, the ComplEx method has the least amount of such recommendations, proving the claim that it can handle complex relationship types more efficiently.

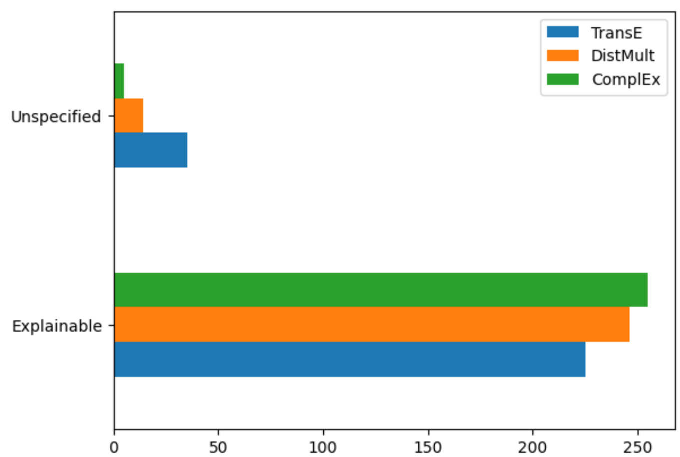

Fig. 5. Amount of Explainable and Unspecified recommendations using different KG embedding models

Having performed the above experiment we choose ComplEx in order to proceed to the next evaluation metric, but even using the simple TransE would be an acceptable choice provided we had a very large dataset.

4.3.2 Evaluating recommendations strength. Another important aspect of the RS is the quality of the provided recommendations. As an indicative trait of this quality we consider the common attributes shared by the current exhibit and a recommendation. In order to achieve this, we perform a simulation of the RS's operation and count the common attributes for each recommendation, before they pass through the filtering mechanism. We consider 20 rounds of exhibit selection and recommendations for the user, with a warm-up period of 3 epochs where only the CB component is active. In Fig. 6 we present a histogram that shows the number of recommendations that share common attributes ranging from zero (i.e., unspecified recommendations) to six (i.e., exhibits share all attributes) for different configurations of the α and β parameters. From this diagram, we can draw some useful insights. First of all, in each case, even in the collaborative filtering only ($\alpha = 0$), the explainable recommendations remain more than the unspecified and also, there exists a non negligible percentage of recommendations that share more than one attribute with the current exhibit the visitor is examining. Moreover, we see that by assigning importance to both components of the RS we can provide quality recommendations that oftentimes have more than one common attribute while in their generation both the exhibit-exhibit similarity but also the history of similar users is taken into account.

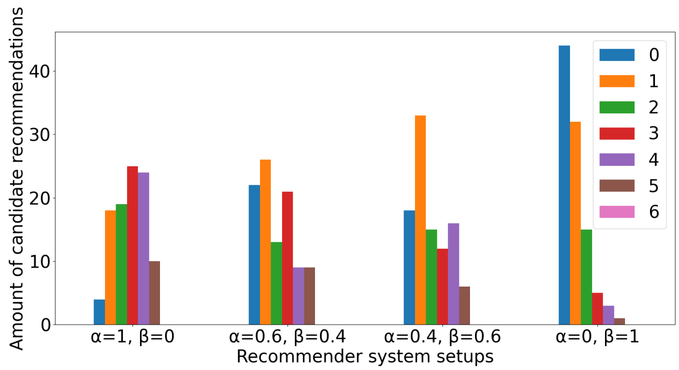

Fig. 6. Recommendations Strength Over Various Setups of α, β

5 CONCLUSION

This article introduces a framework that aims to improve the virtual museum tour experience for users. The proposed solution includes an open-source chatbot that has a custom-trained NER mechanism and an RS. A KG is used to store all relevant information in a structured way, which helps to operate the system efficiently. The developed hybrid RS provides the user with recommendations of relevant exhibits. We conducted simulations to demonstrate the performance of the NER mechanism and the behavior of the RS. In the future, we plan to evaluate the system's effectiveness in realistic settings using questionnaires that are appropriately formatted and filled in by the end users as well as develop a mechanism in order to address the needs of visitors that have limited time at their disposal but still want to learn about the museum's collection.

ACKNOWLEDGMENTS

This work is conducted within the «Research and Innovation Synergies in the Region of Attica» Action, and is realized within the framework of ESPA 2014-2020, co-financed by Greece and the European Union (European Regional Development Fund), under the project with title «Advanced Tour and Education Services in the virtual Paleontology and Geology Museum / v-Palm» (award number: MIS 5185025).

REFERENCES

[1] Deborah Agostino, Michela Arnaboldi, and Antonio Lampis. 2020. Italian state museums during the COVID-19 crisis: from onsite closure to online openness. *Museum Management and Curatorship* 35, 4 (2020), 362–372. https://doi.org/10.1080/09647775.2020.1790029

[2] Nazakat Ali. 2020. Chatbot: A Conversational Agent employed with Named Entity Recognition Model using Artificial Neural Network. *arXiv e-prints*, Article arXiv:2007.04248 (June 2020), 10 pages. https://doi.org/10.48550/arXiv.2007.04248 arXiv:2007.04248 [cs.CL]

[3] Pierre Baldi, Paolo Frasconi, and Padhraic Smyth. 2003. *Modeling the Internet and the Web.* Wiley Chichester.

[4] Tom Bocklisch, Joey Faulkner, Nick Pawlowski, and Alan Nichol. 2017. Rasa: Open source language understanding and dialogue management. *arXiv preprint arXiv:1712.05181* (2017), 9 pages.

[5] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. In *Advances in Neural Information Processing Systems*, C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger (Eds.), Vol. 26. Curran Associates, Inc., 9 pages. https://proceedings.neurips.cc/paper_files/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf

[6] Pedro JS Cardoso, João MF Rodrigues, João AR Pereira, and João DP Sardo. 2017. An object visit recommender supported in multiple visitors and museums. In *Universal Access in Human–Computer Interaction. Design and Development Approaches and Methods: 11th International Conference, UAHCI 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part I 11*. Springer, Springer International Publishing, Cham, 301–312. https://doi.org/10.1007/978-3-319-58706-6_24

[7] Janneth Chicaiza and Priscila Valdiviezo-Diaz. 2021. A comprehensive survey of knowledge graph-based recommender systems: Technologies, development, and contributions. *Information* 12, 6 (2021), 232. https://doi.org/DOI={10.3390/info12060232}

[8] Yuanfei Dai, Shiping Wang, Neal N Xiong, and Wenzhong Guo. 2020. A survey on knowledge graph embedding: Approaches, applications and benchmarks. *Electronics* 9, 5 (2020), 29 pages. https://doi.org/DOI={10.3390/electronics9050750}

[9] Google 2023. *Dialogflow*. Google. Retrieved March 29, 2023 from https://cloud.google.com/dialogflow

[10] Europeana Foundation 2023. *Discover Europe's digital cultural heritage*. Europeana Foundation. Retrieved March 29, 2023 from https://www.europeana.eu/en

[11] Alessio Ferrato, Carla Limongelli, Mauro Mezzini, and Giuseppe Sansonetti. 2022. The META4RS Proposal: Museum Emotion and Tracking Analysis For Recommender Systems. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. Association for Computing Machinery, 406–409. https://doi.org/10.1145/3511047.3537664

[12] Matthew Honnibal, Ines Montani, Sofie Van Landeghem, Adriane Boyd, et al. 2020. spaCy: Industrial-strength natural language processing in python. (2020).

[13] Anran Jiao. 2020. An intelligent chatbot system based on entity extraction using RASA NLU and neural network. In *Journal of physics: conference series*, Vol. 1487. IOP Publishing, 012014. https://doi.org/10.1088/1742-6596/1487/1/012014

[14] Stefan Kopp, Lars Gesellensetter, Nicole C Krämer, and Ipke Wachsmuth. 2005. A conversational agent as museum guide–design and evaluation of a real-world application. In *Intelligent Virtual Agents: 5th International Working Conference, IVA 2005, Kos, Greece, September 12-14, 2005. Proceedings 5*. Springer, 329–343. https://doi.org/10.1007/11550617_28

[15] Laurent Lescop. 2021. Expanding Museums: new tools and concepts for mediation and virtualization. In *Proceedings of the Museum Environments: Challenges and Opportunities conference 2021*. HAL, Cairo, Egypt.

[16] Olga Loboda, Julianne Nyhan, Simon Mahony, Daniela Romano, and Melissa Terras. 2019. Content-based recommender systems for heritage: Developing a personalised museum tour. In *Proc. 1st International 'Alan Turing'Conference on Decision Support and Recommender Systems*. DSRS-Turing, 11–16.

[17] Octavian-Mihai Machidon, Aleš Tavčar, Matjaž Gams, and Mihai Duguleană. 2020. CulturalERICA: A conversational agent improving the exploration of European cultural heritage. *Journal of Cultural Heritage* 41 (2020), 152–165. https://doi.org/10.1016/j.culher.2019.07.010

[18] Chihiro Maehara, Kotaro Yatsugi, Daewoong Kim, and Taketoshi Ushiama. 2012. An exhibit recommendation system based on semantic networks for museum. *Innovations in Intelligent Machines–2: Intelligent Paradigms and Applications* 2 (2012), 131–141. https://doi.org/10.1007/978-3-642-23190-2_10

[19] Yeo-Gyeong Noh and Jin-Hyuk Hong. 2021. Designing Reenacted Chatbots to Enhance Museum Experience. *Applied Sciences* 11, 16 (2021), 15 pages. https://doi.org/10.3390/app11167420

[20] Grand View Research. 2020. *Recommendation Engine Market Size, Share & Trends nalysis Report By Type (Collaborative Filtering, Hybrid Recommendation), By Deployment, By Application, By Organization, By End-use, By Region, And Segment Forecasts, 2021 - 2028*. Grand View Research. Retrieved March 27, 2023 from https://www.grandviewresearch.com/industry-analysis/recommendation-engine-market-report#

[21] S Reshmi and Kannan Balakrishnan. 2018. Enhancing inquisitiveness of chatbots through NER integration. In *2018 International Conference on Data Science and Engineering (ICDSE)*. IEEE, IEEE, Kochi, India, 1–5. https://doi.org/10.1109/ICDSE.2018.8527788

[22] Tuukka Ruotsalo, Krister Haav, Antony Stoyanov, Sylvain Roche, Elena Fani, Romina Deliai, Eetu Mäkelä, Tomi Kauppinen, and Eero Hyvönen. 2013. SMARTMUSEUM: A mobile recommender system for the Web of Data. *Journal of Web Semantics* 20 (2013), 50–67. https://doi.org/10.1016/j.websem.2013.03.001

[23] Giancarlo Sperlí. 2020. A deep learning based chatbot for cultural heritage. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*. Association for Computing Machinery, New York, NY, USA, 935–937. https://doi.org/10.1145/3341105.3374129

[24] Aleš Tavčar, Antonya Csaba, and Eugen Valentin Butila. 2016. Recommender system for virtual assistant supported museum tours. *Informatica* 40, 3 (2016), 279–284.

[25] Miguel Torres-Ruiz, Felix Mata, Roberto Zagal, Giovanni Guzmán, Rolando Quintero, and Marco Moreno-Ibarra. 2020. A recommender system to generate museum itineraries applying augmented reality and social-sensor mining techniques. *Virtual Reality* 24 (2020), 175–189. https://doi.org/10.1007/s10055-018-0366-z

[26] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Eric Gaussier, and Guillaume Bouchard. 2016. Complex Embeddings for Simple Link Prediction. In *Proceedings of The 33rd International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 48)*, Maria Florina Balcan and Kilian Q. Weinberger (Eds.). PMLR, New York, New York, USA, 2071–2080. https://proceedings.mlr.press/v48/trouillon16.html

[27] Bishan Yang, Scott Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2015. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In *Proceedings of the International Conference on Learning Representations (ICLR) 2015* (proceedings of the international conference

on learning representations (iclr) 2015 ed.). ICLR, 12 pages. https://www.microsoft.com/en-us/research/publication/embedding-entities-and-relations-for-learning-and-inference-in-knowledge-bases/

Received 20 April 2023