EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2023)
A. Bucciero, B. Fanini, H. Graf, S. Pescarin, and S. Rizvic (Editors)

HerMeS - HERitage sMart social mEdia aSsistant: from requirement elicitation to data modelling for feeding Artificial Intelligence recommendation system

Alberto Bucciero[1], Alessandra Chirivi[2], M.A. Jaziri[3], I. Muci[4], A. Orlandini[5] and A. Umbrico[6]

Abstract

In the last years, several mobile APPs have been developed within the cultural tourism domain to give new impetus to this sector which is booming both in Italy and worldwide. In the wake of the increasing importance of technologies based on artificial intelligence, even mobile applications for the use of cultural tourism heritage are increasingly taking advantage of these techniques. Machine learning strategies are increasingly used to recommend points of interest and itineraries that are compatible with the user's preferences, requirements and constraints. The quality and integrity of the data acquired become the starting point for training and implementing AI models. By levering well-structured data, these algorithms can offer valuable insights, personalised recommendations, and enhanced user interaction in the cultural tourism domain.
The HerMeS APP that we present in this paper was designed starting from these premises. The application aims to provide a wide range of artificial intelligence-based features to enhance the enjoyment and exploration of cultural heritage, both tangible and intangible.

CCS Concepts
• ***Applied computing*** → ***Arts and humanities;*** *Digital libraries and archives;* • ***Computing methodologies*** → ***Artificial intelligence;*** • ***Human-centered computing*** → *Interactive systems and tools;*

Introduction

Several recent studies have highlighted a new trend in tourism known as the "always-connected traveller" [MLP18] With the increasing use of mobile devices like smartphones and tablets, more and more people are constantly connected while travelling and enjoying experiences. This has prompted tourism service providers to explore innovative ways of engaging with these connected travellers. The paper focuses on the profound transformations and challenges that digital technologies bring about in the enhancement of cultural heritage and the dissemination and promotion of sustainable cultural tourism. The innovative mobile APP HerMeS (HERitage sMart social mEdia aSsistant), that we present, fits into this scenario. The design and development of the APP are the goals of the namesake project proposed and developed by Consiglio Nazionale delle Ricerche (CNR), i.e., Istituto di Scienze del Patrimonio Culturale (ISPC) in collaboration with Istituto per l'Analisi dei Sistemi ed Informatica "Antonio Ruberti" (IASI), Istituto di Scienze e Tecnologie della Cognizione (ISTC), and DigiLab Centro interdipartimentale di Ricerca at Università La Sapienza. The project aims at offering tools and innovative services to favour the fruition of (tangible and not tangible) Cultural Heritage in the Lazio region through advanced AI and ICT methodologies and technologies.

The APP allows users to share their expertise insights, and experiences, thereby enriching the collective knowledge base surrounding Artificial Intelligence algorithms. These inputs from users serve as essential building blocks, enabling the system to learn and improve, ultimately enhancing its capabilities and effectiveness. Through machine learning algorithms and advanced analytics, AI systems can process and interpret large volumes of data with exceptional speed and accuracy [CCDB*20]. By harnessing the power of AI, we can automate tasks, streamline operations, and enhance overall efficiency. The first section of the paper focuses on data modelling, specifically discussing the conceptual model and the significance of a data dictionary. The second section delves into data processing, highlighting the process and technical methods involved. The third section examines the interoperability of data, particularly its application in facilitating the exchange of information and services. Lastly, the paper explores the implementation of an intelligent recommendation system for cultural itineraries, leveraging the insights gained from data modelling, data processing, and interoperability to enhance user experiences and provide tailored

DOI: 10.2312/gch.20231151

recommendations. In the last section, there are the conclusions and further perspectives.

1. HerMeS platform goals

HerMeS aims to promote the enhancement and enjoyment of Lazio's Cultural Heritage, in its many territorial manifestations and multiple meanings. Through a collaborative and participatory system, stakeholders become an active part of the system by sharing experiences, opinions, services, and tools [VB00].

A preliminary survey about the APP and other social platforms for touristic-cultural content sharing, as well as the current use of AI systems in cultural tourism promotion platforms and their impact in the reference context, was conducted. Also in light of this, there are two main considerations on which the design of our idea is based:

- **From institutional information to social experiences**: the amount of institutional, promotional and marketing information available and specially prepared to provide valid support in choosing the places of cultural-tourist interest to visit is so large, it is often considered useless, misleading, and sometimes, even deceptive. On the other hand, we all know that in the moment of embarking on a trip or planning a visit, we rely more on the impressions (reviews, evaluations, stories, etc.) of other visitors and not on the "institutional" or "official" information that describes the place of interest. In other words, we not only trust the "experiences" of those who preceded us, but in many cases, we try to replicate them.
- **From data to information**: in a completely connected world, where information arrives from all over, the common expectation, both during the planning of a trip and during its course, is to have a considerable amount of data precisely at the moment needed. Furthermore, this information must have a high "quality" and must be able to respond to particular needs and specific needs of users, linked, for example, to the presence of children, the disabled, etc.

The idea of creating an innovative social platform to support 'cultural discovery', in which the same users produce, describe and upload the content, comes from the first consideration. On the other hand, for the proposed system to be able to provide valuable information, it must be intelligent. It must understand the expectations, and interests of users and, at the same time, take into account contextual situations (opening hours, availability of places, characteristics of the itinerary, presence of constraints or special needs). The APP must therefore be able to propose personalized paths in time and space.

2. Requirement elicitation

There are different techniques for the analysis of requirements (requirement elicitation in English). In the specific case of the HerMeS project it was decided to use the AWARE methodology (Analysis of Web Application REquirements) [BP04]. AWARE envisages a process of identifying the actors/stakeholders and their objectives (goals) and subsequent refinement of each goal into one or more requirements. The analysis of the scenarios shows the existence of different actors with different goals and objectives. In detail, the actors identified are:

- tourist who can be local (could have own transport) or non-local;
- group of tourists;
- moderator;
- expert/ Institutional expert;
- financing body
- administrator
- domain related professionals

Through a goal-oriented analysis, the most relevant use cases were first defined and analyzed. Finally, the scenarios considered most relevant were implemented in order to detect any user requirements that had been omitted. The following table (fig n° 1) shows the basic information of each stakeholder, in order to clearly and simply define their role. From this table, it will be possible to identify the key functions that must be developed in the application.

STAKEHOLDERS	GOALS
Tourists (local and no-local, tour groups), Tourist guide	- plan a visit (search for POIs/tourist routes/cycles paths); - book a visit; - review the PoI or route already visited; - create content; - content sharing;
Moderator	- check and verify the information relating to PoI and the contents connected to it;
Expert/ Institutional Expert	- produce content for proven scientific relevance; - recognition of the authority of the professional figure; - recognition of the authorship of the published contents;
Financing body	-monitor the progress of the works; -obtain useful information to activate new activities in support of the Culture Heritage;
Administrator	-have access to the back-end platform to add, edit, delete (contents, PoIs, users)
Domain related professionals	-follow the development of the sw components and carry out testing operations; -use open APIs and datasets for development of the application solutions

Figure 1: *Stakeholders involved and corresponding goals*

This analysis gave rise to a relevance matrix that cross-referenced user requirements with the corresponding stakeholder (fig. n° 2).

	STAKEHOLDERS						
REQUIREMENTS	1.Tourists/ Touristic Guide	2.Local Expert	3.Institutional Expert	4.Moderator/ Reviewer	5.Financial body	6.Administrator	7.Domain related professionals
1.Access to APP functionalities	!!!	!!!	!!!	x	!!!	!!	!!!
2.Personalize the personal profile	!!!	!!!	!!!	x	x	x	x
3.Choose interests	!!!	!	!	x	x	x	x
4. Research itineraries	!!!	!!!	!!!	x	x	!!	!!
5. Visualize itineraries	!!!	!!!	!!!	x	x	!!	!
6.PoIS and Trips creation	!!!	!!!	!!!	x	x	x	x
7.Sharing into App's community	!!!	!!!	!!!	!!!	x	x	x
8.Manage the Trips	!!	!	!	x	x	!	!!
9.Visualize PoiS info	!!	!	!	x	!!!	!	!
10.Check PoI's info	!!	!	!	x	!!!	x	!
11. Alert notification referring to events or Trips	!!!	!	!	x	x	!	x
12.Start the trip	!!!	!!	!!	x	x	x	x
13.Modify the trip	!!	!	!!	x	x	!!	!
14. Disseminate value trips	!	!!	!!	x	x	x	x
15. Increase authority	!	!!!	!!!	x	!!!	!	!
16.Scientific recognition	x	!!!	!!!	x	x	x	!!!
17.Checking the real trip	!	!!!	!!!	!!!	!	x	!
18. Send notification to Admin	!	!!!	!!!	!!!	x	x	!
19. Visualize all published contents from APP users	x	!!!	!!!	!!!	!!	!!!	!!!
20. Algoritm that alert harmful content	x	x	x	!!!	x	!!!	!!!
21.Valuation	!!!	!	!	x	!!!	!	!

Figure 2: *Relevance matrix*

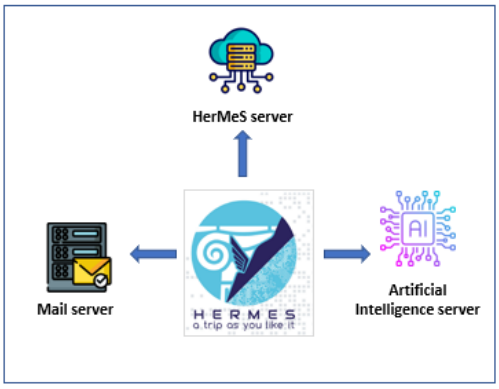

Figure 3: *Actors involved in HerMeS*

The requirements elicitation highlighted the presence of additional actors who interact with HerMeS. They are able to satisfy the main functional requirements such as Registration and creating a new account (fig. n°3).

They are:

- Mail Server: it is a system that should provide us with the functionality to send mail.
- Artificial Intelligence Server: it is a system that should inform us the tourist the right choice to take.
- HerMeS Server: it is a system that should interact with the requests sent by different applications.

The last diagram that we show in this chapter is Entity-Relationship [Che76] diagram related to HerMes entities (fig. n° 4). The systematic analysis defines and describes what data is created and which is needed for processes in HerMeS domain. Thanks to the remarkable outcomes achieved, we had the opportunity to thoroughly examine the data provided for HerMeS, as will be illustrated in the subsequent paragraphs. In fact, we successfully extracted the following entities: Users, Categories, Category_Interests, User_Interest, Trips, Favorite_Trip, PoIs, PoIs_Trips, and Interests. Each of these entities is interconnected with others, establishing relationships that are cardinal in nature.

3. Data Modelling

Starting from the previously described requirement elicitation phase, a data model has been created, identifying entities within the system, and determining where they fit in relation to each other. Data modeling encompasses different approaches, each serving specific purposes. Conceptual data modeling focuses on understanding the high-level relationships between entities, emphasizing the overall structure and the key concepts involved. This type of modelling provides a conceptual framework that serves as the foundation for further development.

For the design of the overall architecture, we adopted Conceptual data modeling. This technique is particularly suitable as it provides a holistic understanding of the relationships between entities, serving as valuable input for the AI algorithms. By levering conceptual data modeling, we can create a structured representation that aids in organizing and utilizing data effectively for our AI algorithms.

In order to show the innovation of the HerMeS project, we organized the data understanding which is the data that dives into Artificial Intelligence. In the following paragraphs, we underline which data are involved.

3.1. User Information

Our AI algorithms heavily rely on gathering information directly or indirectly from users. HerMeS offers users the opportunity to share their data while utilizing the application. This data acquisition process begins with the creation of an account, either through social media integration or traditional registration, ensuring a seamless experience for users seeking to plan their trips.

Each component of the data we collect holds significant importance for yielding optimal results through our AI algorithms. Figure 5 exemplifies the various information fields that tourists can populate. We present users with a standard registration form that includes essential details such as username, email, and password. Furthermore, we encourage users to complete an additional optional form designed specifically to cater to our artificial intelligence requirements.

Several crucial features play a pivotal role in nourishing our AI

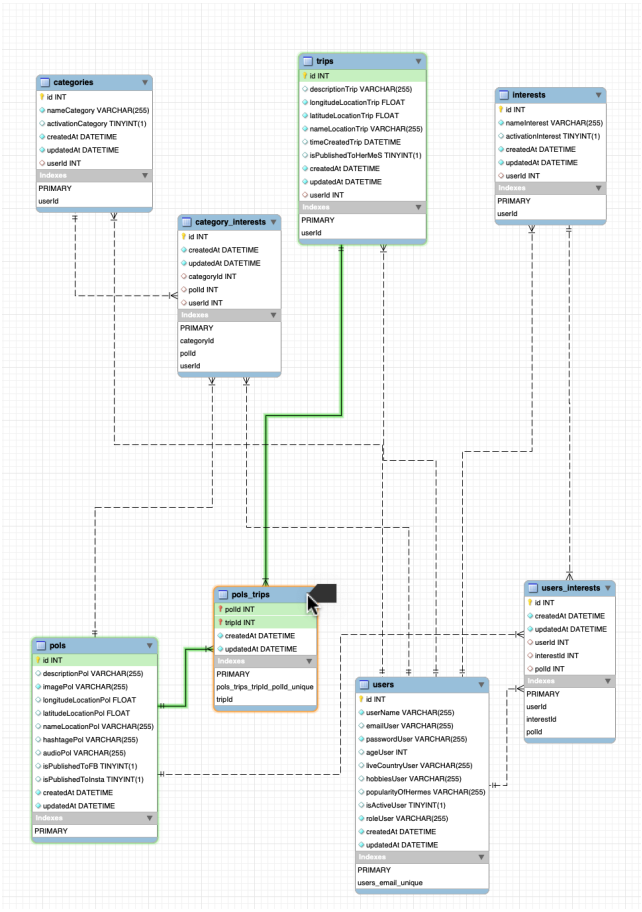

Figure 4: *Entity Relation Diagram*

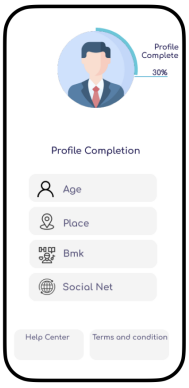

Figure 5: *User registration*

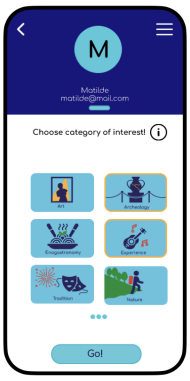

Figure 6: *User set interest*

algorithms. It is imperative to ascertain the age of each tourist to determine their respective category, enabling us to curate tailored experiences. Knowing the country of origin allows us to gain insights into the tourist's cultural and religious background, enabling us to propose suitable trips and recommend appropriate dining establishments. Additionally, our AI algorithm seeks information regarding the tourist's hobbies and preferred activities during their leisure time, facilitating meticulous planning of engaging and enjoyable activities for each trip. Lastly, the traditional registration process incorporates a social network component, which helps us understand the most prevalent sources through which tourists discover and utilize the HerMeS application, thereby informing our strategic decision-making processes.

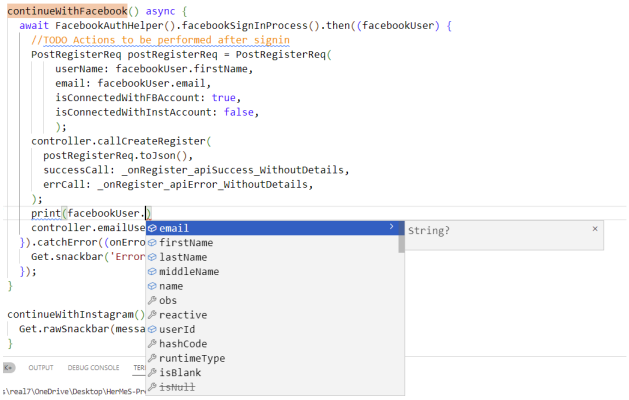

Figure 7: *Public Information on Social Media Account*

In addition to traditional registration, the HerMeS application offers an alternative approach through social media account integration, which holds equal significance in obtaining comprehensive tourist information. Users are presented with the option to continue their journey using either their Facebook or Instagram accounts, as illustrated in figure 7, showcasing the Facebook account selection process.

When opting for the social media login method, our application ensures the utmost security and privacy. The social media platform's robust security system prompts the user to grant authorisation for accessing their private information. Once the tourist grants this authorisation, the application seamlessly transfers all relevant data directly to our artificial intelligence algorithms, enabling them to perform the same data analysis and processing as with the previously described methods.

3.2. Set Interest

The data acquisition in HerMeS application continues with the section concerning the choice of interests, also in this session, AI Algorithms are on stage. HerMeS application gives chances to the user to receive a complete list of *Point of Interest* which match the user's taste. The setting of what the user likes to do after its registration is possible thanks to the section figure n° 6. Based on the selected interest, which is saved on the Database and in order to be sent to the algorithm the user will receive a complete list of Point Of Interest belonging to the Lazio region, according to the selected category of interest. After registration - login, the user will always have the option to modify his/her interest.
After the user participation in setting the interests, the data are saved on the database in order to be sent to the AI algorithm.

3.3. App session

The efficacy of our artificial intelligence algorithms relies on adhering to specific features and requirements. By providing the necessary input and fulfilling these criteria, we can achieve exceptional outcomes. This is exemplified by figure 8 and figure 9, which depict these essential features. As illustrated, the application prompts tourists to configure their trip preferences, enabling them to receive a curated list of results generated by our AI algorithms.

While setting up their trip, tourists should exercise caution and attention to detail. Although certain features may not be explicitly mandatory, we can assure users that incorporating all relevant information will yield optimal and favourable outcomes from our AI algorithm. By ensuring comprehensive data input, tourists can expect the best possible recommendations and results tailored to their preferences and needs.

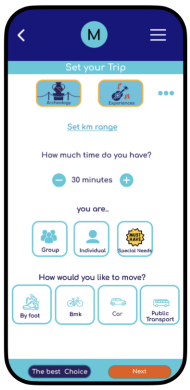

Figure 8: *Set a trip*

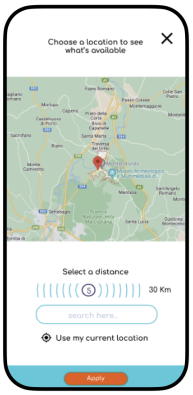

Figure 9: *Select a KM range*

As previously highlighted, the input of crucial information by tourists is essential for nourishing our artificial intelligence algorithms. Within the application, the Explore session serves as a vital repository of data. This section encompasses the direct interaction between users, specifically tourists, and the trip settings. Consequently, all the aforementioned features are intelligently presented by our artificial intelligence algorithms. Once the trip settings data is established, it becomes imperative to store this information in our database, as any subsequent interaction carried out by the tourist within our application can prove invaluable for our algorithms.

However, before storing this data in the database, it must undergo a conceptual modeling phase. This phase encompasses the overall structure, content, and delineation of business rules involved. Amongst the key components of the conceptual model, the Class Diagram [JH02] assumes paramount importance. The Class Diagram represents the structure of entities managed by the system, outlining their relationships.

Figure n° 7 provides a glimpse into a segment of our Class Diagram, with a specific focus on the Point of Interest and Trips entities. These entity classes [Che76] form the foundational core of our AI algorithms, playing a pivotal role in generating personalized recommendations and optimizing the user experience.

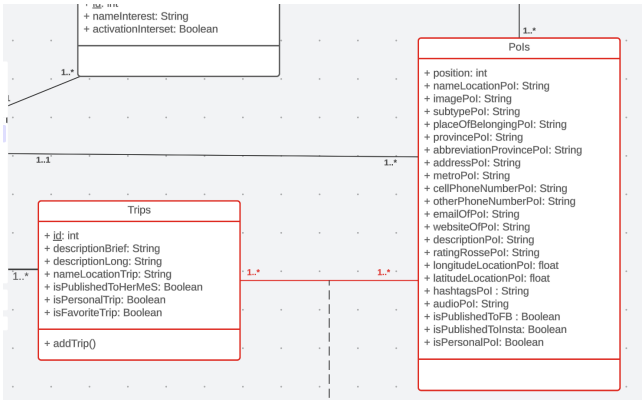

Figure 10: *Association in detail between two entities*

3.4. Data-Dictionary

To ensure utmost clarity within the development ecosystem, it is considered a fundamental practice to provide a meticulously constructed *data-dictionary* [RE15] that comprehensively outlines the properties associated with each entity. In our specific case, we have in detail organized this information into tables, with each table encompassing the entity's name, element value, description, data type, character length, acceptable values, as well as indicators denoting requirements and permissibly of null values. This meticulous arrangement not only facilitates a thorough understanding of the data but also streamlines its utilization within the development ecosystem. By adhering to this approach, we enhance the overall efficiency and effectiveness of data-driven processes, enabling seamless integration and data utilization in various contexts (fig. n° 11). Furthermore, to exemplify the practical implementation of a data dictionary, we have included a visual representation (fig. n°12), specifically illustrating the data dictionary of the "Users" table. This example serves as a concrete demonstration of how the data dictionary provides detailed insights into the properties and characteristics of the data elements within the table. By examining the data dictionary, one can gain a deeper understanding of the structure, data types, allowable values, and other relevant information pertaining to the "Users" entity. This exemplification highlights the invaluable role that a data dictionary plays in facilitating data comprehension and supporting the seamless integration of artificial intelligence algorithms.

This example serves as a concrete demonstration of how the data dictionary provides detailed insights into the properties and characteristics of the data elements within the table. By examining the data dictionary, one can gain a deeper understanding of the structure, data types, allowable values, and other relevant information pertaining to the "Users" entity. This exemplification highlights the invaluable role that a data dictionary plays in facilitating data comprehension and supporting the seamless integration of artificial intelligence algorithms.

Users							
Spreadsheet tab	Element or Value Dispaly name	Description	Data Type	character Lenght	Acceptable Values	Required?	Accepts null value?
users	id	PK	INT		N/A	yes	n
users	userName	Words used by user to identify system	varchar	255	N/A	yes	n
users	emailUser	email of the user	varchar	255	N/A	yes	n
users	passwordUser	Personal password of the user	varchar	255	N/A	yes	n
users	ageUser	Personal age of the user	INT	NULL	N/A	yes	n
users	liveCountryUser	The place where the User lives	varchar	255	N/A	yes	n
users	hobbiesUser	Free time of the user	varchar	255	N/A	yes	n
users	popularityOfHermes	Where the user knew about HerMeS	varchar	255	N/A	yes	n
users	isActiveUser	The status of the user (active/not active)	tinyINT	1	N/A	yes	n
users	roleUser	the activity that distinguishes the user	varchar	255	N/A	yes	n
users	createdAt	Creation date of user	DATETIME	NOT NULL	YYYY-MM-DD hh:mm:ss	yes	n
users	updatedAt	Update date of user	DATETIME	NOT NULL	YYYY-MM-DD hh:mm:ss	yes	n

Figure 11: *User properties table*

Unique Keys			
Type	Table	Key	Columns
Primary	users	pk_users	id
Unique Keys	users	uk_users	emailUser

Figure 12: *Primary key detail*

4. Data processing: process and technical method

Data processing starts with data in its raw form and converts it into a more readable format (graphs, documents, etc.), giving it the form and context necessary to be interpreted by computers and utilised by employees throughout an organisation [Fre96].

The HerMeS database is a critical component of our project and plays a vital role to achieve our objectives. To fully leverage its potential, we must conduct a thorough review and validation of all data to be included.

Our primary data sources are provided by the Touring Club [https://www.touringclub.it/], one of the HerMeS'partner. The culture association provides a data-set equipped with (user, trip, PoIs etc.) The other data-set is generated by Artificial Intelligence algorithms. As the sole providers of data, it is due that we submit their outputs to rigorous verification procedures. Procedures which include checking data before giving the result to the users avoid the error which AI can produce [SAE14]

To ensure the accuracy and reliability of the data, we rely on two key techniques: data cleaning [Wu13] and data processing [Fre96]. The data cleaning method involves the identification and removal of any errors or inconsistencies in the data, while data processing transforms the raw data into a usable format for analysis. These techniques are critical to the success of our project, as they help to ensure that the insights we derive from the data are accurate and reliable.

4.1. Data Cleaning

In our work, we used Python programming language and Pandas library to perform data cleaning [https://pandas.pydata.org/]. The first step in our data cleaning process involved removing empty rows from the data-set. Empty rows can cause errors and inconsistencies in data analysis, and therefore, removing them improves the accuracy and quality of our data. We achieved this by using the "dropna()" function of the Pandas library, which drops all rows that contain missing values.

Next, we addressed the issue of duplicate rows in our data-set. Duplicate rows can occur due to various reasons, such as data entry errors, data collection issues, or system errors. To remove duplicates, we used the "drop_duplicates()" function of the Pandas library, which drops all rows that have the same values across all columns.

To visualize the impact of data cleaning on our data-set, (Fig. n° 13 and 14) demonstrated the changes made to the data. These figures helped you to understand the distribution of data, identify potential outliers, and confirm that the cleaning process was successful.

```
# Importing the pandas library, we can use pd as a shorthand reference
import pandas as pd

# Read a Excel file and create a pandas DateFrame for the two sheets
dfItinerario = pd.read_excel('Data/CampioneItinerarioRomaTCI.xlsx', 'Itinerario')
dfTappe = pd.read_excel('Data/CampioneItinerarioRomaTCI.xlsx', 'tappe')
```
✓ 0.1s

```
# Drop rows where the rows are empty in the sheet of Itinerario
dfItinerario = dfItinerario.dropna(how= "all")
new_dfItinerario = dfItinerario.fillna({
    'denominazione': "",
    'descrizione_breve': "",
    'descrizione_lunga': "",
    'numero_tappe': "",
})
# Drop rows where all data is the same
new_dfItinerario = new_dfItinerario.drop_duplicates()
new_dfItinerario.head()
```
✓ 0.0s

Figure 13: *Remove duplication rows*

4.2. Data Processing

The transformation of data from a CSV file to a database, we need to use one of the techniques of data processing. This process involves converting data from one format to another, which is necessary when integrating data from multiple sources. In this case, transforming the CSV data to a database format allows for easier data analysis and management. One of the key data processing techniques we used was converting our data from a DataFrame to a list.

A DataFrame is a two-dimensional data structure in Pandas library [https://pandas.pydata.org/] that consists of rows and columns, similar to a spreadsheet. It allows for easy manipulation and analysis of data and is a popular tool for data processing.

To convert our DataFrame to a list, we used the "tolist()" function in Pandas, which transformed each column of the DataFrame into a list. This allowed us to perform further data processing tasks on our data in a more flexible way.

Once we had transformed our data, we inserted it into our

```
# Drop rows where the rows are empty in the sheet of tappe
dfTappe = dfTappe.dropna(how= "all")
new_dfTappe = dfTappe.fillna({
    'entità': "",
    'posizione_tappa': "",
    'nome': "",
    'categoria': "",
    'tipologia': "",
    'sottotipologia': "",
    'luogo_di_appartenenza': "",
    'provincia': "",
    'sigla_provincia': "",
    'indirizzo': "",
    'metro': "",
    'telefono': "",
    'cellulare': "",
    'altro_recapito_telefonico': "",
    'email': "",
    'sito_web': "",
    'descrizione': "",
    'rating_rosse': "",
    'latitudine': "",
    'longitudine': "",
})
# Drop rows where all data is the same
new_dfTappe = new_dfTappe.drop_duplicates()
new_dfTappe.head()
```

Figure 14: *Remove empty rows*

database. To accomplish this, we used the SQLAlchemy library, which provides a powerful and flexible way to connect to databases and perform SQL operations.

Using the SQLAlchemy library [https://www.sqlalchemy.org/], we connected to our database and created an engine object. We then used the "to_sql()" function in Pandas to insert our data into the database. This function allows us to specify the table name, the engine object, and other parameters, making it easy to insert large amounts of data into a database quickly and efficiently.

To visualize the impact of our data processing techniques on our data-set. The figures 15 and 16 helped us to confirm that the data was transformed and integrated correctly and that it was consistent and accurate.

```
from sqlalchemy import *
from sqlalchemy.orm import sessionmaker, query, relationship
from sqlalchemy.ext.declarative import declarative_base
from sqlalchemy.sql import func

engine = create_engine("mysql+pymysql://hermes:123@localhost/hermes_db?charset=utf8mb4")

Session = sessionmaker(bind=engine)
session = Session()
Base = declarative_base()
```

✓ 0.3s

Figure 15: *DataBase Connection*

```
# Convert all the column of our DataFrame to a Data Structure (List)
# Those list are the features of the point of interests
dfTappeOnListNome = new_dfTappe['nome'].tolist()
dfTappeOnListSottotipologia = new_dfTappe['sottotipologia'].tolist()
dfTappeOnListMuogoDiAppartenenza = new_dfTappe['luogo_di_appartenenza'].tolist()
dfTappeOnListProvincia = new_dfTappe['provincia'].tolist()
dfTappeOnListSiglaProvincia = new_dfTappe['sigla_provincia'].tolist()
dfTappeOnListIndirizzo = new_dfTappe['indirizzo'].tolist()
dfTappeOnListMetro = new_dfTappe['metro'].tolist()
dfTappeOnListTelefono = new_dfTappe['telefono'].tolist()
dfTappeOnListCellulare = new_dfTappe['cellulare'].tolist()
dfTappeOnListAltroRecapitoTelefonico = new_dfTappe['altro_recapito_telefonico'].tolist()
dfTappeOnListEmail = new_dfTappe['email'].tolist()
dfTappeOnListSitoWeb = new_dfTappe['sito_web'].tolist()
dfTappeOnListDescrizione = new_dfTappe['descrizione'].tolist()
dfTappeOnListRatingRosse = new_dfTappe['rating_rosse'].tolist()
dfTappeOnListLatitudine = new_dfTappe['latitudine'].tolist()
dfTappeOnListLongitudine = new_dfTappe['longitudine'].tolist()

# This two list are foe the categories and interests tables
dfTappeOnListCategoria = new_dfTappe['categoria'].tolist()
dfTappeOnListTipologia = new_dfTappe['tipologia'].tolist()

# Those list are the features of the Trips
dfItinerarioOnListDenominazione = new_dfItinerario['denominazione'].tolist()
dfItinerarioOnListDescrizioneBreve = new_dfItinerario['descrizione_breve'].tolist()
dfItinerarioOnListDescrizioneLunga = new_dfItinerario['descrizione_lunga'].tolist()
```

✓ 0.0s

Figure 16: *Convert a DataFrame to List*

In conclusion, data cleaning and data processing are crucial techniques that play a vital role in ensuring the accuracy and reliability of the data we use in our project. By carefully verifying and validating the data we add to the HerMeS database, we can ensure that the insights we derive from it are valid and reliable, ultimately contributing to the success of our project.

5. Interoperability of data applied at the change of info and services

Interoperability refers to the ability of different systems or components to work together effectively and efficiently. In the context of data, interoperability allows for the seamless exchange and use of data across different systems, applications, and organizations. In the case of the HerMeS database, interoperability can be applied in several ways to enhance the functionality and value of the system. For example, by adopting open standards and protocols, HerMeS can ensure that its data is compatible with other systems and services, allowing for easier integration and exchange of information.

Interoperability can also be applied to the way in which HerMeS presents and shares its data with users. By using standardized interfaces and APIs, HerMeS can make it easier for users to access and use its data, regardless of the specific device or platform they are using. In addition, interoperability can be used to facilitate the sharing and integration of data from other sources. For example, HerMeS can integrate data from other Cultural Heritage databases, tour operators, or social media platforms, allowing for a more comprehensive and personalized user experience. Overall, interoperability is a key consideration for the HerMeS database, as it enables the system to work effectively with other systems and services, while also providing a more seamless and user-friendly experience for tourists and other stakeholders.

To pursue interoperability, HerMeS follows a rigorous approach by defining a novel ontology suitable to characterize the *cultural heritage* domain [GOS09]. A formal model of defining general concepts and properties supports the integration of existing data sources and catalogs as well as proper and uniform interpretation of knowledge.

5.1. The ArCO Ontology

HerMeS relies on the ArCO ontology specifically designed for cultural heritage [CGM*19, CGM*21]. The key advantage of the ArCO ontology is its modularity which supports flexible integration and usage within HerMeS. ArCO defines a number of general concepts and properties that are suitable to interpret pieces of knowledge and integrate existing thesaurus e.g., PICO 4.1 https://www.vocabularyserver.com/pico/it/index.php. However, the model of ArCO mainly focuses on *tangible cultural entities* that are only part of the cultural entities addressed within the project. One of the key contributions of HerMeS indeed is the capability of integrating tangible and intangible cultural entities. HerMeS extends some of the ArCo concepts by refining territorial concepts in order to support a structured (and layered) description of a territory identifying parts (areas) and sub-parts that are relevant from a heritage perspective. In addition, HerMeS introduces new types of properties supporting the description of the topological structure of a territory.

HerMeS especially extends ArCO by refining the representation of *intangible cultural entities* and correlating them with tangible cultural entities in order to support the *narrative* needs of the System. A contextual representation of intangible cultural entities and the capability of correlating them to tangible entities are indeed crucial to unlocking hidden relationships between places, history, religions, food, and local traditions.

6. Towards an Intelligent Recommendation System for Cultural Itineraries

The HerMeS project aims to provide a diverse range of AI-based functionalities to enhance the enjoyment and exploration of Cultural Heritage, both tangible and intangible. HerMeS facilitates the connection of various heterogeneous needs and interests among different stakeholders and more in particular of tourists. Through a bottom-up participation model and leveraging advanced IT technologies, including AI algorithms, personalized itineraries, and valuable information is proposed to tourists and visitors. Using AI and machine learning algorithms, HerMeS can analyze various variables such as user preferences, historical data, and current trends, to generate personalized itineraries for tourists. The app recommends specific cultural sites, events or activities based on the individual's interests and preferences.

To this aim, the information stored in the databases developed for Hermes will collect knowledge from different and highly heterogeneous sources. And, a recommendation system considers the information related to users and their preferences in order to select a set of *cultural items* and propose a personalised itinerary. To support such AI-based features, a Knowledge Base (KB), based on ArCo, was designed to characterize a wide set of information concerning the cultural heritage of a territory considering geographic and structural features as well as tangible and intangible cultural qualities of tangible entities that are part of a specific territory and are relevant from a heritage point of view.

6.1. Linking Tangible and Intangible Entities

To support a contextualized experience of users it is important to define *meta-knowledge* defining *semantic linkgs* between tangible and intangible cultural entities described within the KB. Indexing cultural properties according to different topics and points of view enables the construction of contextualized narratives [MBM21]. In this regard, the definition of a well-structured taxonomy of *topics* and *themes* supports the (contextualized) filtering, and retrieval of (sub-sets of) knowledge entities that are relevant and coherent with respect to a certain set of topics. In other words, topics "tag" the description/content of a certain cultural entity as relevant with respect to the selected topics. A basic topic, already defined in ArCo, can be leveraged to define perspectives (e.g., religion, social, art) that "classify" the type of content associated with the modeled cultural entities. Figure 17 shows an excerpt of the taxonomy defined within HerMeS by extending the abstract concept Topic defined within ArCO.

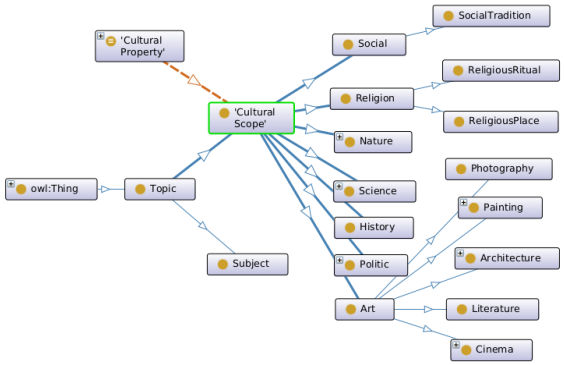

Figure 17: *Excerpt of the HerMeS taxonomy of topics*

6.2. Contextual Recommendations and Temporal Planning

Topics represent a semantic bridge between user preferences and the cultural content of the knowledge base. Topics in particular characterize coherent perspectives and points of view correlating tangible and intangible cultural entities (i.e., meta-knowledge). Having a well-structured list of topics and knowing the sub-set of topics that are relevant to a user is crucial to generate contextualized recommendations and personalized itineraries [CCDB*20]. Within HerMes we in particular design an advanced recommendation system taking advantage of the effective combination of semantic representations and automated planning [UCOC20]. Topics and semantic reasoning support knowledge filtering and the selection of the list of entities that fit users' profiles the most. Such knowledge is then used within a temporal planning system based on the timeline-based formalism [CMOU16] to synthesize personalized itineraries for the users by following "thematic" narratives.

It is important to point out that in order to synthesize feasible itineraries knowledge about topics is not sufficient. The selection of the entities to be concatenated within a certain cultural path (i.e., narrative) depends on many factors. The time available for the whole visit should be enough to complete the synthesized plan.

Some entities could have been already visited by a user in previous experience. Some entities could be congested entailing significant delays and tedious waiting queues for the user. From a planning point of view, the synthesis of contextualized and feasible itineraries can be seen as an optimization problem aiming at maximizing the user experience (e.g., maximizing the number of visited cultural entities) while satisfying the duration and temporal constraints of the physical visit. To this aim, HerMeS should integrate technical knowledge about the time needed to complete a visit (the duration could be further refined by taking into account different "levels of detail" of a visit depending on the level of expertise of a user), as well as the expected number of users visiting the same entities simultaneously. Such knowledge would be crucial to allow the planner to synthesize visiting paths that achieve a good level of experience in the given total time.

7. Conclusion

As part of the Hermes project, the innovative mobile APP was designed to offer tourists, visitors and others advanced and customised solutions for discovering the material and immaterial cultural heritage of the Lazio region. Through a bottom-up model of participation and leveraging advanced IT technologies, including artificial intelligence algorithms, personalised itineraries and valuable information is offered. To achieve these goals, a solid conceptual model was defined and a rigorous data model was developed. At the same time, an extension of the ArCO ontology was defined to represent and correlate intangible cultural entities with tangible ones to meet the narrative needs of the System.

8. Acknowledgements

The research leading to these results has received funding from the POR FESR LAZIO 2014–2020 – A0375E0110.

The authors are especially grateful to Dr. Costanza Miliani, director of the Istitute of Heritage Science (CNR-ISPC), for her constant and tangible support.

References

[BP04] Bolchini D., Paolini P.: Goal-driven requirements analysis for hypermedia-intensive web applications. *Requirements Engineering 9* (2004), 85–103. 2

[CCDB*20] Cesta A., Cortellessa G., De Benedictis R., De Medio C., Limongelli C., Sciarrone F., Tassarotti . G., Palombini A.: Personalizing technology-enhanced learning for cultural visits. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (New York, NY, USA, 2020), UMAP '20 Adjunct, Association for Computing Machinery, p. 333–339. URL: https://doi.org/10.1145/3386392.3399278, doi:10.1145/3386392.3399278. 1, 8

[CGM*19] Carriero V. A., Gangemi A., Mancinelli M. L., Marinucci L., Nuzzolese A. G., Presutti V. ., Veninata C.: Arco: The italian cultural heritage knowledge graph. In *"The Semantic Web – ISWC 2019"* (Cham, 2019), Ghidini C., Hartig O., Maleshkova M., Svátek V. ., Cruz I., Hogan A., Song J., Lefrançois M., Gandon F., (Eds.), Springer International Publishing, pp. 36–52. 8

[CGM*21] Carriero V. A., Gangemi A., Mancinelli M. L., Nuzzolese A. G., Presutti V., Veninata C.: Pattern-based design applied to cultural heritage knowledge graphs. *Semantic Web 12* (2021), 313–357. 2. 8

[Che76] Chen P. P.-S.: The entity-relationship model—toward a unified view of data. *ACM Trans. Database Syst. 1*, 1 (mar 1976), 9–36. URL: https://doi.org/10.1145/320434.320440, doi:10.1145/320434.320440. 3, 5

[CMOU16] Cialdea Mayer M., Orlandini A., Umbrico A.: Planning and execution with flexible timelines: a formal account. *Acta Informatica 53*, 6-8 (2016), 649–680. 8

[Fre96] French C.: *Data Processing and Information Technology*. Thomson Learning, 1996. URL: https://books.google.it/books?id=zVCdg7Tg6-AC. 6

[GOS09] Guarino N., Oberle D., Staab S.: *What Is an Ontology?* Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 1–17. 7

[JH02] Johnson J., Henderson A.: Conceptual models: Begin by designing what to design. *Interactions 9*, 1 (jan 2002), 25–32. URL: https://doi.org/10.1145/503355.503366, doi:10.1145/503355.503366. 5

[MBM21] Meghini C., Bartalesi V., Metilli D.: Representing narratives in digital libraries: The narrative ontology. *Semantic Web 12* (2021), 241–264. 2. 8

[MLP18] Marino V., Lo Presti L.: Engagement, satisfaction and customer behavior-based crm performance: An empirical study of mobile instant messaging. *Journal of Service Theory and Practice 28*, 5 (2018), 682–707. 1

[RE15] Ramez Elmasri S. B. N.: *Fundamental of Database Systems*, seven edition ed. Ramez Elmasri, 2015. 5

[SAE14] Stuart Armstrong K. S., Eigeartaigh S. S. O.: The errors, insights and lessons of famous ai & predictions – and what they mean for the future. 157–164. 6

[UCOC20] Umbrico A., Cortellessa G., Orlandini A., Cesta A.: Modeling affordances and functioning for personalized robotic assistance. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference* (2020), AAAI Press. 8

[VB00] Varvasovszky Z., Brugha R.: A stakeholder analysis. *Health Policy and Planning 15*, 3 (09 2000), 338–345. URL: https://doi.org/10.1093/heapol/15.3.338, arXiv:https://academic.oup.com/heapol/article-pdf/15/3/338/9806369/150338.pdf, doi:10.1093/heapol/15.3.338. 2

[Wu13] Wu S.: A review on coarse warranty data and analysis. *Reliability Engineering System Safety 114* (2013), 1–11. URL: https://www.sciencedirect.com/science/article/pii/S0951832013000100, doi:https://doi.org/10.1016/j.ress.2012.12.021. 6