### **A logo of a university Description automatically generated with medium confidenceA blue and yellow badge with a picture of a building and a gold ribbon Description automatically generatedUniversity of Bucharest**

**FACULTY OF MATHEMATIS AND INFORMATICS**

**DEPARTMENT OF INFORMATICS**

**Bachelor's Thesis**

**HistoryMapper**

**Graduate**

**Neaga-Budoiu Maria**

**Scientific coordinator**

**Lect.dr. Dobrovăț Anca**

**Bucharest, June 2024**

**Abstract**

History has always been a major subject for creating a solid basis of general culture, and exploring it could be facilitated using technology. This is the main objective of the HistoryMapper application, to make the process of learning History more accessible and appealing using technology, and, in particular, to help the exploration of historical events by forming visual connections between their position in time and space.

Therefore, the platform provides an interactive map, obtained with the usage of Google Maps API. It changes through a timeline axis and highlights historical events and routes manipulated using different local APIs. Even if the map is the core of the application, other important features are provided as well, including the possibility for users to navigate through historical periods and take quizzes related to the events they discover on map.

The backend of the application is written in Django framework for Python, while the interface is written in Angular. The data used by HistoryMapper is stored in an Oracle SQL database, which is partially populated with information for DBPedia database, the main storage system for Wikipedia.

#### **Rezumat**

Istoria a fost întotdeauna un subiect principal în crearea unei baze solide de cultură generală și explorarea ei poate fi facilitată de folosirea tehnologiei. Acesta este principalul obiectiv al aplicației HistoryMapper, să facă învățarea istoriei mai accesibilă și fascinantă cu ajutorul technologiei și, în particular, să ajute la explorarea evenimentelor istorice prin crearea de conexiuni vizuale între localizarea lor în timp și spațiu.

Astfel, platforma oferă o hartă interactivă, obținută prin folosirea API-ului de Google Maps, care se modifică printr-o axă temporală și evidențiează evenimentele și rutele istorice manipulate cu ajutorul diferitelor apeluri de API-uri. Deși harta este nucleul aplicației, alte funcționalități sunt, de asemenea, oferite, inclusiv posibilitatea ca utilizatorii să navigheze prin perioadele istorice sau să dea quiz-uri legate de evenimentele pe care le descoperă.

Partea de backend a aplicației e scrisă în framework-ul de Python Django, în timp ce interfața e creată în Angular. Datele folosite de HistoryMapper sunt stocate într-o bază de date Oracle SQL care e parțial populată și cu informații preluate din baza de date DBPedia, principalul sistem de stocare folosit de Wikipedia.

1. [Introduction](#_bookmark0) 4
   1. [Motivation](#_bookmark1) 4
   2. [Concurency](#_bookmark2) 5
   3. Exploited Domains 5
   4. [Project](#_bookmark2) Structure 5
      1. Data Manipulation 5
      2. Map Feature 5
      3. Users’ System and Actions 4
2. Data Manipulation 7
   1. Database [Structure](#_bookmark1) 4
   2. Back-end APIs 4
   3. Google APIs 4
   4. Historical Periods 4
   5. Populating the Database 4
      1. Manually 5
      2. From DBPedia Database 5
      3. Authomatic Location Mapping 5
      4. Authomatic YouTube Videos Insertion 5
3. Map Feature 6
   1. [Google](#_bookmark1) Maps API 4
   2. Timeline Axis 4
   3. Events on Map 4
      1. Markers 5
      2. Search and Filter Options 5
      3. Historical Routes 4
      4. Event's Information and YouTube Videos 4
      5. Clustering Algorithm 4
4. Users’ System and Actions 8
   1. [Authentication](#_bookmark1) System 4
   2. Admin Actions 4
   3. Quizzes 4
5. Technologies
   1. [Back-end](#_bookmark1) 4
   2. Front-end 4
   3. Database 4
6. [Conclusions](#_bookmark5) 9
   1. [Summary](#_bookmark1) 4
   2. Further Development Possibilities 4
7. Bibliography 10

# **Introduction**

## **Motivation**

History has always been a subject of interest for me and I consider that today’s younger generations are losing interest in the subject because of the lack of modernization of the methods for its learning. I believe that the use of technology could make the domain more attractive for those who explore it, so I have chosen an interractive and facile approach for presenting historical facts in my project.

HistoryMapper is addressed to anyone who wants to explore the domain, from historians or teachers who need quick access to data to users who want to discover facts of the past. Therefore, the platform has an educational utility as it provides information materials for the studying process, including data about the events and all the historical periods, and quizzes that test the knowledge of the users.

I decided that the platform should be centered on the map feature because it creates visual connections between the historical timeline and the physical locations of the events, which makes HistoryMapper an interractive application and the navigation more suggestive. Apart from the fact that visual learning is a powerful method for memorizing information, I also believe that this way of presenting History is not much explored by other sources and has a lot of potential to be further exploited.

## **Concurency**

Fusce fermentum quis tellus vitae congue. Interdum et malesuada fames ac ante ipsum primis in faucibus. Proin convallis congue sapien semper volutpat. Curabitur rhoncus et nunc quis tempor. Quisque sollicitudin dolor id varius luctus. Donec sapien dolor, eleifend quis justo id, euismod ultrices nulla. Suspendisse leo quam, luctus in velit ac, venenatis molestie odio. Aliquam vitae feugiat elit. Sed auctor gravida leo, ut dictum enim tempus non. Suspendisse in dapibus urna, vitae pretium ligula. Morbi ut diam laoreet, convallis leo non, semper arcu. Sed vel lorem tortor. In auctor a diam sit amet convallis. Etiam ipsum ligula, luctus quis interdum ac, consequat vitae orci. Maecenas hendrerit turpis.

## **Exploited Domains**

For creating this project, I considered approaching domains covered in my academic experience. I chose to focus on:

* **Web Applications Development,** used for creating the overall project structure, the API endpoints, the authentication system and user actions, the front-end elements, and the connection with the back-end, and the database
* **Software Engineering,** highlighted in establishing the goals of the application, as well as the development process, organized on tasks with different priorities scheduled over a period of time
* **Databases,** for the creation of the models and the relations between them, and for the queries used in the back-end APIs

## **Project Structure**

The project is roughly divided into: the aspects reagarding database manipulation and completion, the elements related to the map feature, and the system of users with the corresponding actions. In the sections below, I will describe the development aspects in regard to each of these parts.

## **Data Manipulation**

One of the most important particularities that I established for the application from the beginning was the possibility to store a large amount of events in order to cover all periods of Human History as much as possible. This is the reason why the Oracle SQL Datbase used as a storage system has as a center-entity the one with events data, which is connected to the entities storing locations and historical periods, both essential for placing events in a tempo-spatial context. In this section I will cover one of the major parts of the development process, data creation and handling, which will include the logic of the application’s APIs, the connection to DBPedia database, the storage of historical periods, and the authomatic population of the locations entity with the algorithms for establishing geocoordinates.

## **Map Feature**

The creation of this functionality involves all the operations for personalizing and manipulating the map, from correlating it to the timeline axis to features such as searching and filtering events on map or displaying historical routes. For this section, I will also explain the process of Google Maps API integration for displaying the map, and the usage of different Google Maps classes for markers and routes. One other aspect that will be covered is the clustering algorithm, a method for grouping events based on zoom level of the map. Nevertheless, the platform has integrated YouTube videos that complete the information about the events on the map, so I will describe the method for calling the YouTube API and obtaining them.

## **Users’ System and Actions**

The platform is intended to be user-oriented, so I created a system for authentication, which also provides access for users to the quizzes functionality. The performance data for each quiz will be stored per-user, logic which was created with the help of a quiz history entity in the database. Apart from the access to questions, the authentication functionality also gives the administrator of the database access to special CRUD actions for events, including the authomatic population of the events entity.

**Data Manipulation**

## **Database Structure**

All the information for the application is kept in an Oracle SQL Database, linked through a connection to the back-end part of the project. The connection is established in the *settings.py* file of the back-end application by specifying the required database parameters. Django framework provides a facile mechanism for handling data through the models, sources of data simulating the tables in a database. When a model is created, a new entity perfectly matching it is added authomatically to the database along with the corresponding constraints. Moreover, Django provides methods for executing SQL Queries in the Python code, facilities that made the code of the application more compact and the development process faster.

HistoryMapper’s database is focused on the Event entity, that contains all the data needed for the historical facts to be placed on the map. This table is directly connected to almost all the other ones, for having access to additional information about the events such as the location, the historical period or the category they are part of.

Below there is an ERD (Entity-Relationship Diagram) schema of the database.

A screenshot of a computer

Description automatically generated

The Event entity contains foreign keys (referneces) to the Ids of the location and period they are part of, the type (battle, treaty, truce and so on), and the category. While the type is introduced to make identifying an event easier or to group facts in a visual manner in the platform, and is used for the filter option, the category is required for future authomatic insertions and for the creation of quizzes, functionalities described in the following chapters.

Apart from the logic for historical facts, the database also includes the users’ system. Because Django already provides such a system through a specific library (*django.contrib.auth*) imported in the back-end, the User model inherits the already built-in class with the same name, and is necessary only for the creation of a mapping table in the database. This entity is linked to a table storing a quiz result history related to each registered user.

* 1. **Back-end APIs**

All the pages in the platform handle information from the database under different forms. This logic; either being fetching data, modifying the entries or creation of new ones or delete operations; is done in the back-end part of the application in methods called APIs. They define a specific logic for handling the data required by the front-end, and each time the resource is requesting an operation, one API is called and retrieves a needed response or a punctual error message. For example, displaying the events on the map requires the existing database events in order to place them as markers, so the page makes a request to obtain this resource. Not only do these back-end methods obtain data, but they also transform it to a desired format and add additional elements to it, if they are needed in Angular.

In the implementation of HistoryMapper, I created all the APIs using REST (Representational State Transfer) framework, an architectural style defining constraints to be used for creating web services. Fot this, I installed the Django REST framework package, which also brings as a benefit a user-friendly interface for testing each API. The logic imposed by REST framewrok implies that all the methods are of one of the types GET, POST, PUT or DELETE, depending on the type of request the client (the Angular application) sends to the server (the Django part). All API calls are HTTP requests sent to the server via a unique web url, and the response is then sent back to the client in JSON format. I chose this format not only because it is the most popular for Web Services today, and also because it is easy to access the fields of the response back in the Angular application. For more information about the usage of REST framework, consult the chapter regarding the technologies used in HistoryMapper.

One important API for the platform is the one retrieving events between two given years because it fetches the data for the map page, later used by many other functionalities such as clustering, filtering, searching or establishing routes for events. It is called every time a change in the timeline axis occurs. Four paramteres are sent to the server via the url, the start and end years and eras (AD or BC). Diving deeper, into the implemetation of the API, it is created as a class containing a method of type GET, because it should only retrieve data from the server.

The method converts years BC to negative values, action necessary for executing the SQL query that will retrieve entries from the Event entity, performed direcly in the Python code thanks to the object raw method supplied by Django. After that, the method calls a global function for getting coordinates for each event and populating the MapLocation entity, converts all the data to JSON format and returns it as the response to the client.

Similar to this implementation are all the other GET functions in the API classes. However, the methods of types PUT and POST will receive data to be updated, respectively added to the database via a parameter called request explicitly passed to them when the APIs are called in Angular. This parameter simulates an entire entry information for the database and thus can not be specified in the url as the parameters similar to the ones described in the chapter above. The last type of method, DELETE, is used in HistoryMapper only for removing an event in the database and only requires its ID, making the method similar to the GET ones.

Endpoints of different types can be placed under the same API (and thus under the same url as well), so how does the application know what method to choose? The answer is that the selection depends on the format and type of the request sent from the client.

Now, turning to the client side, I will briefly describe the structure of the Angular project, with accent on the API calls. A more detailed description is included in the chapter destined to the technologies used in HistoryMapper.

The front-end side consists of components, each handling a different page or section of the platform. They have access to the APIs in the back-end and to some parameteres shared between components via services, which are classes with private fields, public getters, setters and other methods. The access to the services from components is realized via the principle of Dependency Injection, so the service with the needed API call is injected into the constructor of the component. After that, the service method that fetches API responses is called in an asynchronous manner with the async-await mechanism, and the result of the API call is saved in component for use.

* 1. **Google APIs**

HistoryMapper utilizes various Google API endpoints in its implementation, either in the local APIs found in Django, which are presented in the section above, or directly in the front-end, for the map page. They serve different needs of the platform: displaying a map, obtaining geocoordiantes, fecthing YouTube videos and so on.

Users have access to Google APIs only after loging into the Google Cloud Platform for Developers, and requesting a private Google API Key. This one must be linked to the platform projects where APIs will be enabled. I firstly created a new project in the Developers Console, and added the private key used by it in HistoryMapper’s database because it is required by the method calls inside the code of my project, yet its value must be kept private for security reasons. However, the key can be safely fetched from the application’s storage system.

Google Maps Javascript API is the one that permits the display of the project’s map, as well as the creation of the individual map markers or of the clusters, and the presence of historical routes as roads connecting map points. Information about the structure of the API’s requests and responses, as well as the methods for its integration into applications, is found at its official documentation page[[1]](#footnote-1).

Another such method is the Google Geocoding API, responsable for converting locations into physical coordinates (latitude and longitude) and vice versa, and it is useful for determining the map geolocation of the historical events in the project. One requirement here is that the locations should be specified in a given format: the name of the locality followed by the prefix of the country, the reason why this particular structure is found for the names of the locations in HistoryMapper’s database as well.

Apart from these Google APIs, the YouTube one is also included in the project, for fetching videos related to each of the events.

* 1. **Historical Periods**

For highlighting the educational particularity of HistoryMapper, I wanted to give users access to information about ten periods of time covering almost all the Human History from the prehistoric times to the present. All the events fall into one of these periods based on the date they took place in, so I consider that additional information resources about the historical eras could offer more insight about the explored events as well, and give users the impression of a journey through time.

The HistoricalPeriods front-end page dispalys these eras in chronological order after getting them via an API call. In case of a click on one of the ten sections, a new page is opened, showing information regarding the selected period. But how does the page with historical period information know what period it should display data about, and, more specifically, what period was selected in the HistoricalPeriods page? The answer is that the two components, the Historicalperiod and HistoricalPeriodInfo, share via a service a common parameter, the ID of the currently selected event. So, when the user clicks on the section for a specific era in the first page, a private ID field in the periods service is set to the unique key of that selected historical period. When the new page, the one with detailed description, loads, it gets the saved ID value from the same service, fetches a period by ID with the call of an API, and displays the data in HTML. This how a field is shared between different components, the value is kept in a private field in a service injected in the components, and the data is handled with public getters and setters.

One last feature related to this part of HistoryMapper is that each historical era can be visually seen on the map. The user can navigate directly from HistoricalPeriodInfo page to the map and see the events of the selected period displayed. In order to create this, I made the values of the timeline axis change to the period’s interval of time whenever the user clicks to be redirected to the map from the historical period page.

* 1. **Populating the Database**

For storing a large quantity of events in the database, manually populating it entirely was not a good option, and I needed to find a method to add more entries at same time authomatically. This represented a real challenge especially because the Event table contains a lot of foreign keys to other tables. This fact leads to the necessity of authomatically adding correspondent entries in some of the linked tables as well, for example inserting new locations for the new events to be stored in the database. But the fact that additional information about the events must be stored is not the only reason other entities must be populated with an authomatic algorithm: some information, such as geocoordinates for the map or YouTube videos for events, have to be obtained with the help of some Google APIs, so the process of calling them can be regarded as an authomatic data creation. In this section, I will go through the process for populating the database, from the data which was added manually to the one inserted via back-end algorithms.

## **Manually**

One entity which was populated entirely manually is the HistoricalPeriod because I decided to split the human History into ten main eras and thus the amount of information is small. Another table with the same characteristics is the EventType, including only sixteen main types the historical facts can fall into. In the first stages of development, I needed to have some test data for other tables as well, so I manually added some entries in Event, MapLocation, and Category as well. Some of the tables do not appear in this enumeration because of several reasons: the User entity depends on the authentication system, so it is basically populated at log in, the QuizHistory depends on the users’ actions on the quizzes page and is populated authomatically, and the Video table is strictly related to data fetched by the YouTube API.

Adding new entries by hand can be done either with SQL queries written in a database console application or through the interface provided by Django. I opted for the second choice because it is a faster process as the Django template project includes a page where the admin of the application can add entries on the interface. The only required aspect is that the models be included in the *admin.py* file of the project’s structure in order to appear on the interface. After that, the Django application is running on the port 8000 and the required page is accessed at */admin* by entering a previously chosen username and password.

A screenshot of a computer

Description automatically generated

## **From DBPedia Database**

DBPedia is a project for handling structured information from the Wikipedia platform and it mades it available on the Internet. Users can query properties of chosen Wikipedia sources and obtain them in a given format, which is described on the DBPedia official webpage[[2]](#footnote-2).

In HistoryMapper, obtaining historical events from DBPedia to be inserted in the local database is done by specifying the Wikipedia category they belong to, which is different from the actual category, the one they will be mapped to in the SQL database. I decided not to insert categories with the same name as the ones in Wikipedia because they can be too restrictive in some cases and the application should have event categories that characterize a wider range of database entries.

Now, one challenge in the process of adding events from DBPedia is transforming the information to the format required by the local database. To this aspect is added the fact that events of different types are retrieved in different formats from DBPedia, and these issues leaded to the necessity of creating sepparate functions for types of events that differ. I chose to write three sepparate functions that fetch the data for battles, treaties, and successions from Wikipedia, but the application can be further developed to obtain entries for the other types as well, by adding new similar functions.

In my project, I opted for the Public SPARQL Endpoint for DBPedia which can be accessed via the *<http://dbpedia.org/sparql>* url. The query language used is thus SPARQL, effective in retrieving data across various Web sources in a specific format called RDF. This format is organized as a directed, labeled graph structure, so one important step in the process is transforming the RDF data into a JSON response. A complete description of the SPARQL language can be found at the page of its official documentation[[3]](#footnote-3).

As a first operation, I imported the SPARQLWrapper library in my project to be able to execute SPARQL queries directly in Python. In each of the functions calling the DBPedia API, the endpoint is wrapped, and then there is created the query for retrieving data by Wikipedia category, query that is, after this, executed and its results converted to JSON. But, for example, the information retrieved for battles included in a certain category is not the same as the one obtained for treaties, even if they belong to the same historical group in Wikipedia. What is more, some results of the query do not define complete future entries in the local database and some fields might be missing. So, the information must be completed and turned to the desired format before retrieving it, which is an operation handled differently by each of the three functions.

For all the event types, one example that illustrates that some data might be missing is the one of the event locations found in the response of the SPARQL query. Sometimes the position of an event is specified as a string name when the actual needed data are the geolocation coordinates. In such case, another query is created and run for converting the Wikipedia location to actual latitude and longitude values.

The requests for populating the database with data from DBPedia are sent from the front-end of the application, so, in the back-end, there must be established what type of events are added, what is the Wikipedia category they are fetched from, and what is the new category, the one in the local database, they will fall into. These values are passed as parameters of the request in the URL of the API endpoint that adds data, and based on the specified type (either battle, treaty, or succession), the corresponding function that fetches DBPedia data is called. Its returned response is processed in the API method: the date is turned into a format accepted by Oracle SQL, and the foreign keys to Category and EventType are obtained. In case the current value of the category does not point to any entrance in the database, a new category is inserted. More than this, the location names are transformed to the structure that is needed by the Google Maps API: the name of the locality followed by the prefix of the country.

Only after all these operations, the API method inserts new rows in Event table with the obtained Wikipedia returns them as a response message back to the client.

## **Authomatic Location Mapping**

Each event in the database needs to be placed on the map at the historical location where it took place, represented by a latitude and a longitude value. The problem is that multiple events could have taken place at the same location, meaning that, in order to avoid repetion in the database and, thus, violation of correct database creation principles, the geocoordinates are not stored in the Event entity. Instead, a many-to-one relationship is established between two separate tables, Event and MapLocation.

When adding events into the database in an automatic manner, and even when they are inserted manually, mapping them to locations by hand can be difficult either because of the large quantity of events or to the fact that many-to-one relationships must firstly be checked. To this aspect is added the fact that the process of individually and manually obtaining the geolocation for all the places of the events is a slow process which can be much improved with Google API calls. Therefore, HistoryMapper’s back-end side contains methods that fetch geocoordinates for locations, and then add them to the storage system in case it does not yet include them.

Each time a new event is to be used in the response of an API, the application verifies if its location is already stored in the MapLocation table and, otherwise, adds a new entry consisting of the name, latitude and longitude of the respective place. I decided to perform this check every time events are to be returned as a response and not at the moment of event creation because this provides additional assurance that no event location is missing before performing operations with them, such as trying to display the events on the map.

The logic of populating the MapLocation entity is divided into three main parts: establishing the new locations to be added, obtaining their geocoordinates, and adding them in the database.

The first part is implemented in a global function that will be called by all the APIs that retrieve certain events, and it receives as parameter a list of events that should have correspondent locations in the database. I created a set of the names of all these locations, another one of the names already stored in the database, and then made a difference operation between them, with the purpose of obtaining the missing data. The coordinates need to be calculated for all these locations, so, in a repetitive manner, the logic for providing geocoordinates is performed.

For the second part, the one of converting locations into physical coordinates, I turned to the Google Geocoding API. As all the Google APIs utilized in the development of the application, it firstly had to be enabled in the Google Cloud Developer Console after creating a personal project there and a secret Google API Key granting access. The API call is done via an URL including the private Google API Key value.

## **Authomatic YouTube Videos Insertion**

1. https://developers.google.com/maps/documentation/javascript/overview [↑](#footnote-ref-1)
2. https://dbpedia.org/page/DBpedia [↑](#footnote-ref-2)
3. https://www.w3.org/TR/sparql11-query/ [↑](#footnote-ref-3)