

System requirements

Project name

Classification unclassified
Status in progress
Program name
Project number
Project manager Christian Galley, Kenny Manixab
Version 0.1
Date 19. Dezember 2023
Project sponsor Project sponsor
Author(s)
Distribution

List of changes

Version	Date	Change	Author
0.1			

Table 1: Change control

Description

The system requirements describe requirements for the They are structured according to Hermes 5 together with related system models and prototypes.

Contents

Contents.....	2
List of Figures	3
List of Tables.....	3
1. Rough requirements.....	3
2. System overview.....	4
2.1. High-Level Overview	4
2.1.1. Information System Overview	4
2.1.2. Main Use Cases and Features	4
2.2. IT Infrastructure	7
2.2.1. Components of the IT Infrastructure	7
2.2.2. Technical Requirements	8
2.3. Planning Studies	9
2.3.1. Technology Evaluation	9
2.3.2. Feasibility of Use Cases	12
3. Detailed Requirements	13
3.1. Functional Requirements.....	13
3.1.1. Category 1	13
3.1.2. Category 2	14
3.2. Non-Functional Requirements	16
3.2.1. Category 1	16
3.2.2. Category 2	16
3.3. User Interface Prototype	18
3.4. Relevance Criteria	21
4. Abbreviations and Glossary	22

List of Figures

Figure 1: High level overview, e.g., information flows.....	Erreur ! Signet non défini.
Figure 2: Use Cases	Erreur ! Signet non défini.
Figure 3 : IT Infrastructure, e.g., information flows or a class diagram.	Erreur ! Signet non défini.
Figure 4: Detailed models, e.g., use cases, sequence flows, classes (if not presented before)	Erreur ! Signet non défini.
Figure 5 : User Interface Prototype	Erreur ! Signet non défini.

List of Tables

Table 1: Change control.....	1
Table 2: Rough requirements	3
Table 11: Abbreviations and glossary	22

1. Rough requirements

Short overview of the main features of the system.

ID	Requirement	Acceptance criterion	Importance	Urgency
A1	Chatting	Validates all criteria of the related detailed requirements, technology evaluation successful, personalized LLM	2	2
A2	Group Chatting	Validates all criteria of the related detailed requirements, technology evaluation successful, access of authorized data	4	4
A3	Summarize Document	Validates all criteria of the related detailed requirements, technology evaluation successful	5	4
A4	Data Retrieving	Validates all criteria of the related detailed requirements, access of authorized data	5	4
A5	Planning	Validates all criteria of the related detailed requirements technology evaluation successful, access of authorized data	3	3

Table 2: Rough requirements

2. System Overview

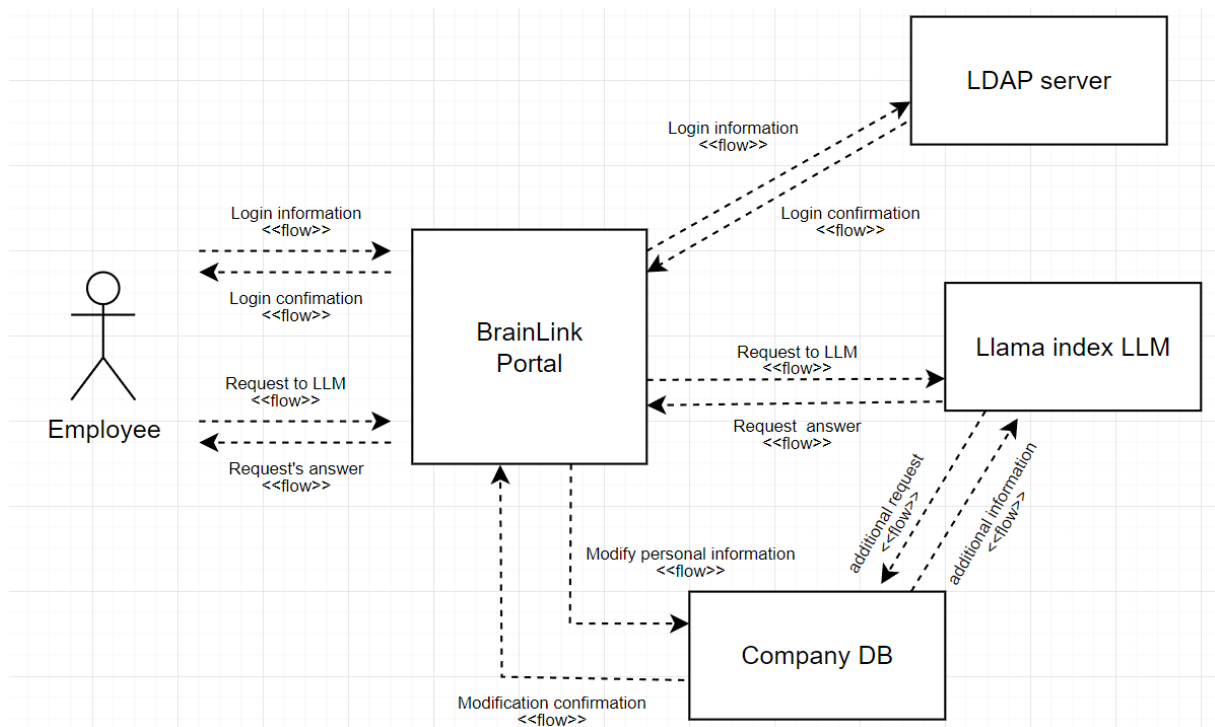
A high-level overview of the system.

2.1. High-Level Overview

The high-level overview shows that *BrainLink Portal* is a tool embedded in a company's system, which has a web interface. The tool allows employees to be assisted in some tasks, such as summarizing documents, creating a group chat based on some criteria, day planning. People high up in the hierarchy have also the possibility to retrieve employee's data if needed.

2.1.1. Information System Overview

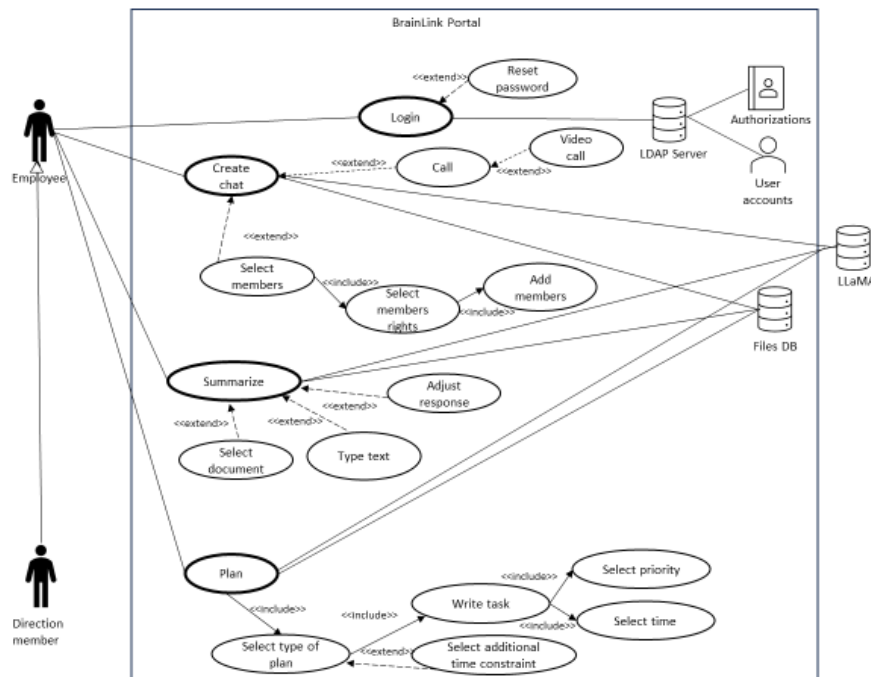
The system structure interacts with several actors. *BrainLink Portal* can ask information to the LDAP server, which store all employee's data for login. It also interacts directly with the LLM, which can access the company DB if required. In the case of trivial request from the user (changing user information for example), The tool changes the data directly in the company DB. All information is displayed to the user.



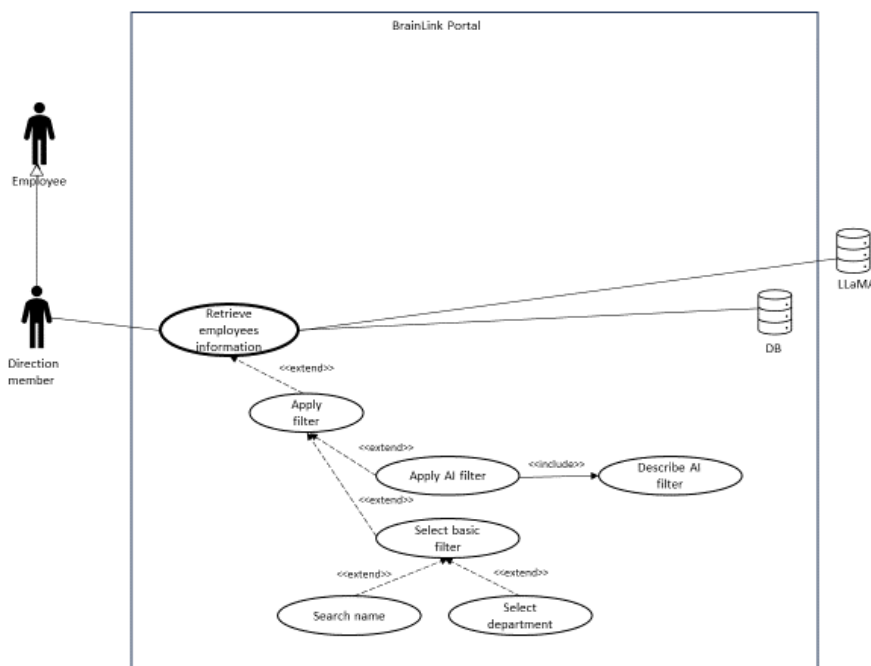
2.1.2. Main Use Cases and Features

The main use cases and features. If they correspond to the use cases, the personas and scenarios can also be included here or in a separate section.

Employees have four different available features, which includes the chat and group chatting, the text summarizing and the planning. (chatting is not represented in this use case, as it has no more functionalities than usual chat bot available online)



Direction members have an additional feature which is data retrieving. The following use case show all possibilities that this feature offers.



Personas

Alex (age : 55, function : direction member) faces challenges related to time management. The constant need for information and communication with team members often consumes a significant portion of Alex's day, leaving less time for strategic planning and project development. He regularly uses email, messaging apps, and video conferencing tools

Frank's (age : 32, function : employee in the sale team) objective in the company is to be well trained and be an expert in his domain. He always has objectives for the end of the month and tries his best to achieve all of them and if not, he would try to find what didn't work. He thinks he would need more training and feels like he works too much to achieve all of his objectives, but his supervisor is also drowned in work and tells him he has no time to train him. At the end of the month, he always spends a lot of time writing the reports.

Scenario

Alex is the supervisor of quite a few projects for the company, he has to add a new weekly meeting for the project Z, he tries to look at his physical calendar but cannot find an empty black where to fit this meeting. He logs in BrainLink Portal and uses the assistant to make a planning for the next month that would fit all of his meetings. He has now a working planning for the next month.

Alex want to create a new team for the project Z, this project aims to enhance and strengthen the cybersecurity by implementing security measures and audits. He needs employees from the IT department as well as employees from the communication department to create posters. He goes to both departments and realize there are a lot of employees in both department. He logs in BrainLink Portal and apply filters on employees that would fit the project the best and forms the new team.

Alex wants to make communication with different teams straightforward and keep it all in one place. Currently, people use different channels based on their preferences. Alex's solution is to use the chat feature on the BrainLink Portal. This way, he can create chats with the relevant people and have all communication in one central channel.

Frank was asked to perform multiple customer services such as product recommendation, order processing or documentation reporting. He is often overwhelmed by the amount of work. BrainLink Portal could help him manage his time in a more effective way by telling him organization advice and even doing his timetable of the week.

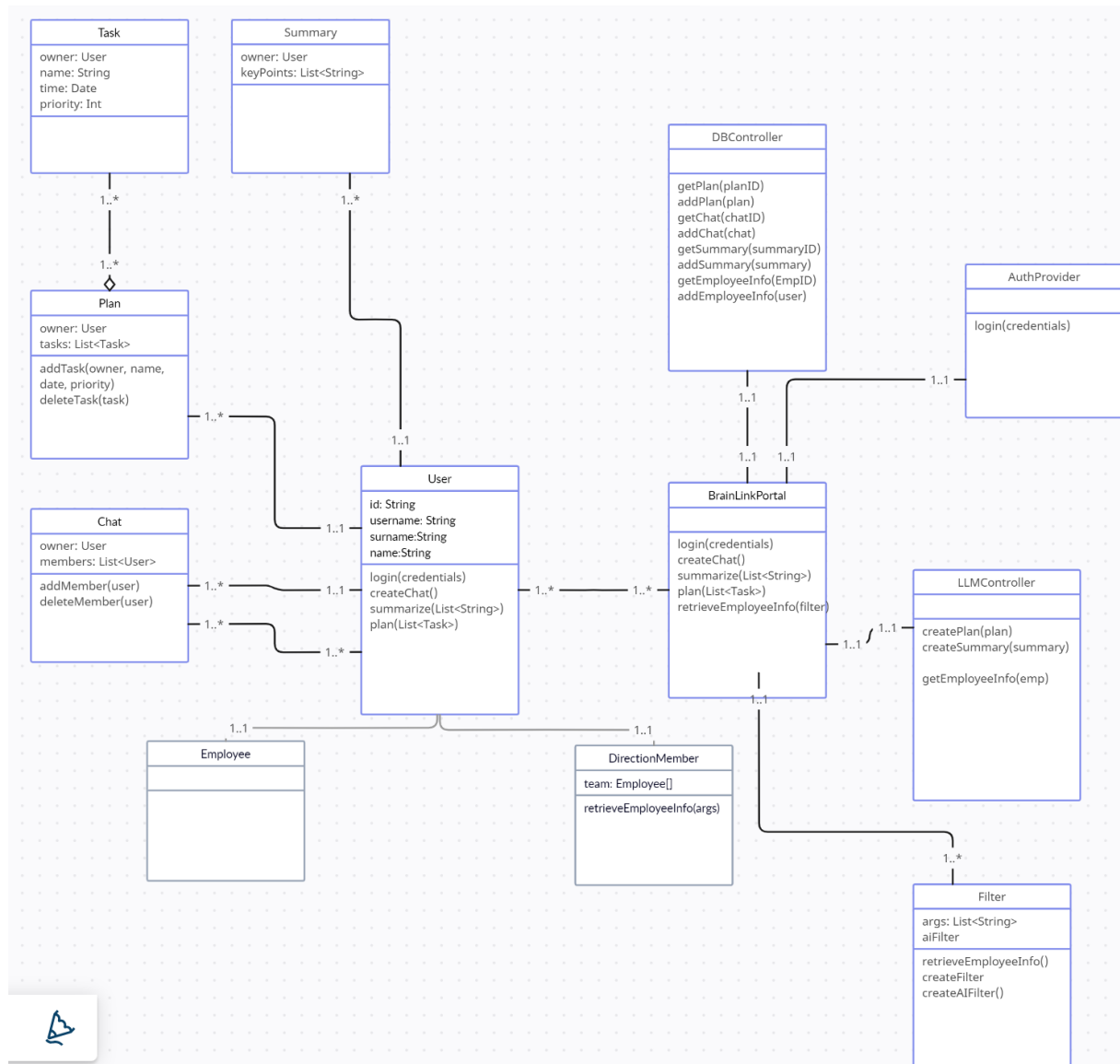
Frank is always asked to write reports at the end of the month. This is a tiring task as he has no writing skills, and this take him too much time to do it. Brainlink Portal can write a nice report within a few seconds based on the criteria Frank gives to the AI.

2.2. IT Infrastructure

The required IT infrastructure involve an LDAP server to store all user's login information, a server that hosts the LLM, and another server that store all the remaining information into a database (PostgreSQL for example). All those components are linked to a unique website: BrainLink Portal.

2.2.1. Components of the IT Infrastructure

The IT infrastructure in terms of information flow or class diagrams.



2.2.2. Technical Requirements

The LLM will be hosted into a server with the following requirements:

- CPU: Core i5
- GPU: > 8GB VRAM
- RAM: > 16GB
- Storage: > 256 GB

It requires a lot of power to perform well. GPU(s) and CPU(s) are the parts that deliver the power needed. It also must handle several instances at the same time, that's what RAM is useful for. The LLM does not store anything on its server, but we need to store the LLM, which takes a lot of space depending on the model used. As we decided to use a small model, 256 GB should be enough and futureproof.

The LDAP server, which stores all users' login information, only requires lots of storage. Since the company has approximately 50 employees, 32GB should be enough, but because storage is not the most expensive part, we allow us to go with 256 GB for futureproof.

One more server stores all remaining information (chat history, all summarized texts, etc.) into a database (PostgreSQL for example). Like the LDAP server, the most important part is the storage. This time, we need a lot of storage. We will begin with 2 TB of storage. We can still scale it if needed and add more storage.

2.3. Planning Studies

2.3.1. Technology Evaluation

Evaluation of the main technologies imposing challenges, in particular LLM abilities to give coherent answer.

Prompt 1 (chat):

Write me a Python function that calculate the standard deviation. the parameter should be a list.

Answer:

Sure! Here's a Python function that calculates the standard deviation of a given list:

```
def standard_deviation(numbers):  
    # Calculate the mean of the list  
    mean = sum(numbers) / len(numbers)  
    # Calculate the squared differences between each number and the mean  
    deviations = [(x - mean) ** 2 for x in numbers]  
    # Calculate the square root of the variance  
    variance = sqrt(sum(deviations))  
    # Return the standard deviation return variance
```

Quality:

LLM can indeed provide information when requested, if it's not too specific (what is the capital of Indonesia, how many bones there are in the human body...). However, complex question that require extreme precision like code generation for example, does not work: it creates random function that does not exist, or forget to import useful packages... For example, this code is not useable, as sqrt is not a function that exist. We should import the math package before, and use it as follow: math.sqrt(...)

Prompt 2 (summarize documents):

based on the following text: [...] summarize the text. Additionally, answer the next question. There is only one possible option for each question.

Question 1 ...

Question 2 ...

... → complete query is available in annexes file

Answer:

Based on the text you provided, here is my summary: [...]

Now, please answer the questions:

1. The text's introduction presents: (d) Two opinions about environmental problems
2. ...

Quality:

It performs very well. The only drawback is that it doesn't support all file types. For the moment, you must copy/paste the text into the LLM. But it understands texts well, it can summarize it and it can even answer questions about it. The answer might be a bit confusing because of the sentence structure, but in case we do not understand the answer, it is still possible to ask to rephrase it.

Prompt 3 (planning)

here is my week schedule:

Monday:

- * 9:00 AM - 10:00 AM: Attend lecture on data structures and algorithms
- * 10:00 AM - 12:00 PM: Work on assignments related to data structures and algorithms

Tuesday:

- * 9:00 AM - 10:00 AM: Attend lecture on computer systems architecture
- * 10:00 AM - 12:00 PM: Work on assignments related to computer systems architecture

Wednesday:

- * 9:00 AM - 10:00 AM: Attend lecture on database management systems
- * 10:00 AM - 12:00 PM: Work on assignments related to database management systems
- * 2:00 PM - 3:00 PM: Attend lab session on data warehousing and business intelligence
- * 3:00 PM - 4:00 PM: Work on assignments related to data warehousing and business intelligence

[...] → full query available in annexes file

Answer:

Based on your schedule, Wednesday appears to be the day with the least amount of work. You have a lab session from 2:00 PM to 3:00 PM, followed by a work session from 3:00 PM to 4:00 PM, which gives you a few hours in the afternoon to fit in an extra lesson or tutorial session.

Quality:

Does not a great job understand planning. It gives random choice about free time.

Prompt 4 (retrieve employees' information):

Based on the following database from a company:

ID	Name	Sex	Age	Number of Current Projects	master's degree	Programming Skills	Languages
----	------	-----	-----	----------------------------	-----------------	--------------------	-----------

1	John Smith	M	35	3	Yes	Python, Java, JavaScript	English, Spanish
---	------------	---	----	---	-----	--------------------------	------------------

2	Jane Doe	F	28	5	Yes	C++, Python, SQL	French, Arabic
---	----------	---	----	---	-----	------------------	----------------

3	Bob Johnson	M	42	2	No	Java, Python, HTML/CSS	Chinese (Mandarin), Japanese
---	-------------	---	----	---	----	------------------------	------------------------------

4	Alice Brown	F	32	0	Yes	Ruby, Swift, MySQL	Korean, Vietnamese
---	-------------	---	----	---	-----	--------------------	--------------------

5	Mike Davis	M	47	2	Yes	C#, Python, Docker	German, Italian
---	------------	---	----	---	-----	--------------------	-----------------

6	Sarah Taylor	F	29	2	Yes	PHP, JavaScript, WordPress	Portuguese (Brazilian), Russian
---	--------------	---	----	---	-----	----------------------------	---------------------------------

7	Tom Wilson	M	39	4	No	Go, Rust, Kubernetes	Hindi, Punjabi
---	------------	---	----	---	----	----------------------	----------------

8	Emily Chen	F	30	5	Yes	TypeScript, Node.js, MongoDB	Thai, Cambodian
---	------------	---	----	---	-----	------------------------------	-----------------

9	David Lee	M	45	5	Yes	Python, Java, Android	Korean, Vietnamese
---	-----------	---	----	---	-----	-----------------------	--------------------

10	Rachel Patel	F	27	6	Yes	React, Angular, CSS	Gujarati, Marathi
----	--------------	---	----	---	-----	---------------------	-------------------

... → full query available in annexes file

built a team of 5 employee for a website project. It should contain only employee with a master's degree, and they must be more than 30 years old.

Answer:

Of course! Based on the provided data, here are the employees who meet the criteria for being part of a team for a website project:

* John Smith (Master's degree, 35 years old)

* Jane Doe (Master's degree, 28 years old)

* Bob Johnson (Master's degree, 42 years old)

* Alice Brown (No degree, 32 years old)

* David Lee (Master's degree, 45 years old)

These five employees have a master's degree and are over 30 years old, meeting the criteria for the team.

Quality:

It gives random answer, and even create information that are not true about employees. In this example, LLM says that Alice Brown has no degree, even though she has one. In addition, Jane Doe is less than 30 years old, which is not what we wanted. Building the query first and add some keyword to replace in the query is probably a better way to test LLM capabilities for data retrieving in a database.

2.3.2. Feasibility of Use Cases

Main Challenges of Main Use Cases.

Use case 1 (chatting): This feature is probably the easiest to implement, as it does not require any adjustment. This feature is a basic chat with the LLM, which is already operational on several server hosted LLM. The only challenge for this feature is tuning the LLM to behave like an employee: it should have the same ideology as the company.

Use case 2 (group chatting): The main challenge for this feature is the data retrieving in the database (id of employees) to create a chatgroup that regroup all users involved. Once the instance is created, the LLM does not interact with users anymore.

Use case 3 (summarize documents): This feature already works well for PDFs. Main challenges would be to implement a way to read and understand several types of files, such as PowerPoint.

Use case 4 (data retrieving): The main challenge of this feature is the query building that should be exact, otherwise it wouldn't work at all. The LLM should already have some query templates with keywords to replace with user inputs. This would limit possible mistake from the LLM. Once the query is built, the LLM runs it in the database and retrieve all information that corresponds to query's criteria. At the end, LLM should be able to display the retrieved data to the user.

Use case 5 (planning): This feature is probably the most difficult to make it viable. The LLM should understand patterns in timetable, which does not work very well without any tuning. That is feasible, but it will require a lot of adjustments.

3. Detailed Requirements

Detailed descriptions of the functional and non-functional requirements in the tabular format mandated by Hermes 5 together with related models. In addition, further detailed descriptions of user interface requirements

3.1. Functional Requirements

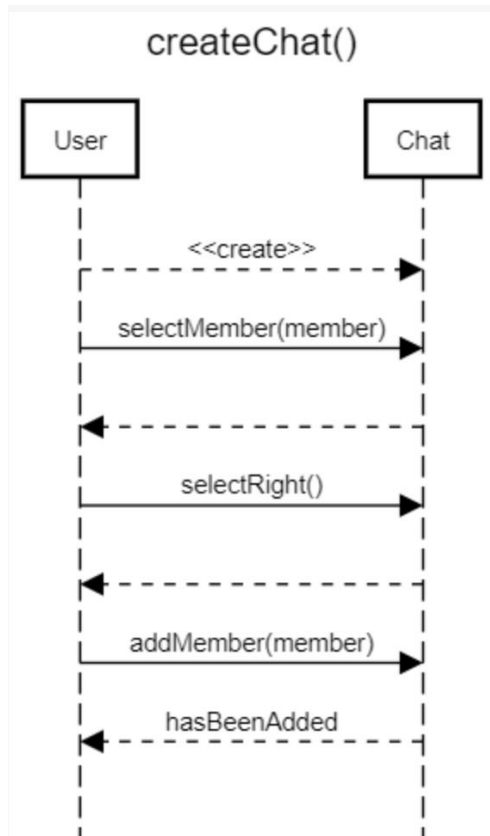
3.1.1. Category 1

These use cases / features concern all chatting features.

ID	C1-F1	Source	Model, Technology Evaluation, IT Infrastructure, ...	Author		Date		Status	approved
Name	Chatting								
Description	Users have the possibilities to talk with the LLM, asks for advice, asks for its point of view about different subjects...								
Acceptance criteria	The LLM should be precise enough to be viable. Giving wrong advice would result in bad work. If the LLM could not give any answer, it should not invent something, but instead just tell the user that it is not possible to give any answer.								
Importance ¹	4	Urgency ²	4	Risk ³	2	Outlay ⁴	3		

ID	C1-F2	Source	Model, Technology Evaluation, IT Infrastructure, ...	Author		Date		Status	approved
Name	Group chatting								
Description	LLM access the company DB to retrieve all IDs of involved user to create an instance of chat. Once created, LLM does not interact with users anymore. All chats are stored in the server.								
Acceptance criteria	LLM should only access required data, which are names of employee and their corresponding ID. It should not, at any moment, access other data.								

Importance ¹	4	Urgency ²	4	Risk ³	2	Outlay ⁴	3
-------------------------	---	----------------------	---	-------------------	---	---------------------	---

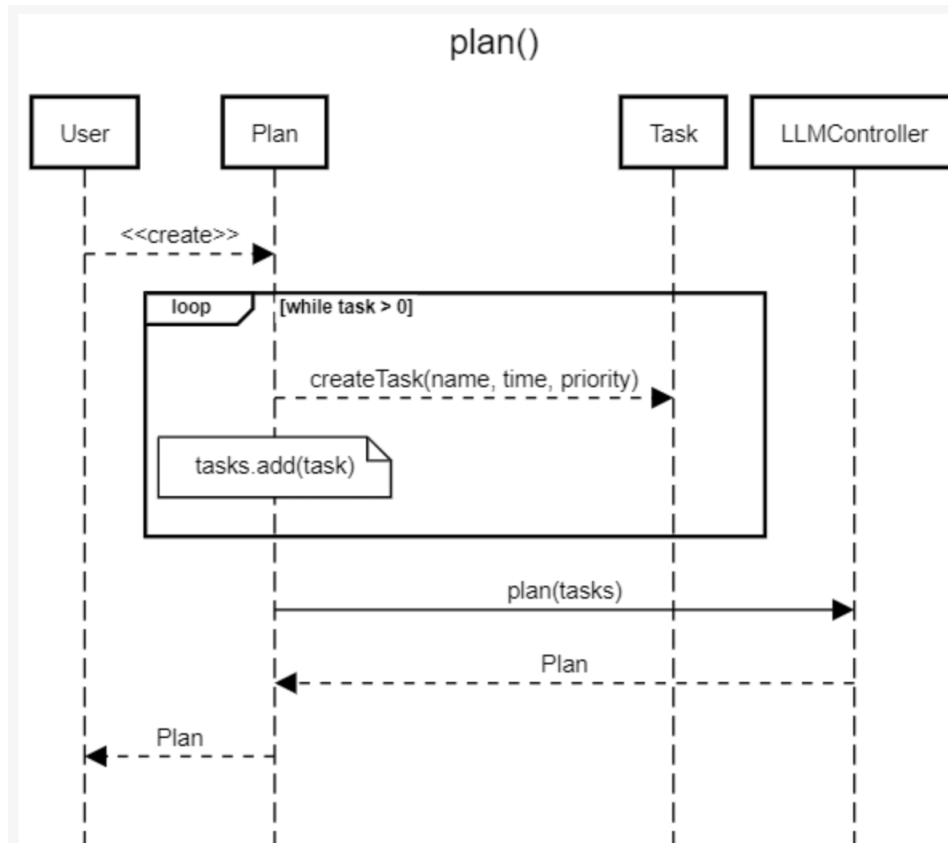


3.1.2. Category 2

These use cases / features concern the planning feature

ID	C2-F1	Source	Model, Technology Evaluation, IT Infrastructure, ...	Author		Date		Status	approved
Name	Planning								
Description	LLM should be able to manage a timetable and optimize the schedule of employee.								
Acceptance criteria	LLM should understand the concept of scheduling and apply it. It should not advice something wrong. It should only help users to manage their time better.								

Importance ¹	4	Urgency ²	4	Risk ³	2	Outlay ⁴	3
-------------------------	---	----------------------	---	-------------------	---	---------------------	---



These use cases / features concern the data retrieving.

ID	C2-F1	Source	Model, Technology Evaluation, IT Infrastructure, ...	Author		Date		Status	approved
Name	Data retrieving								
Description	LLM should access data from employees depending on the request.								

Acceptance criteria	LLM only has read option (no write, no update). Based on a pre-defined query, it access the data in the company DB and retrieve all information returned to the direction member.						
Importance ¹	4	Urgency ²	4	Risk ³	2	Outlay ⁴	3

3.2. Non-Functional Requirements

3.2.1. Category 1

These use cases / features concern LLM capabilities and limitations.

ID	C1-NF1	Source	Model, Technology Evaluation, IT Infrastructure, ...	Author		Date		Status	approved
Name	Workload capacity								
Category	Server specifications								
Description	Considering available features and the number of employee (max. 50), the server hosting the LLM should be strong enough to interact with at most 5 employees at a time (assuming not all employees will use BrainLink Portal at the same time). Server should also be futureproof, in the case more features are added.								
Acceptance criteria	LLM should be responsive enough (depending on the task, it can take more time. A simple text summary should take around 3 seconds to be considered responsive.)								
Importance ¹	5	Urgency ²	4	Risk ³	1	Outlay ⁴	2		

3.2.2. Category 2

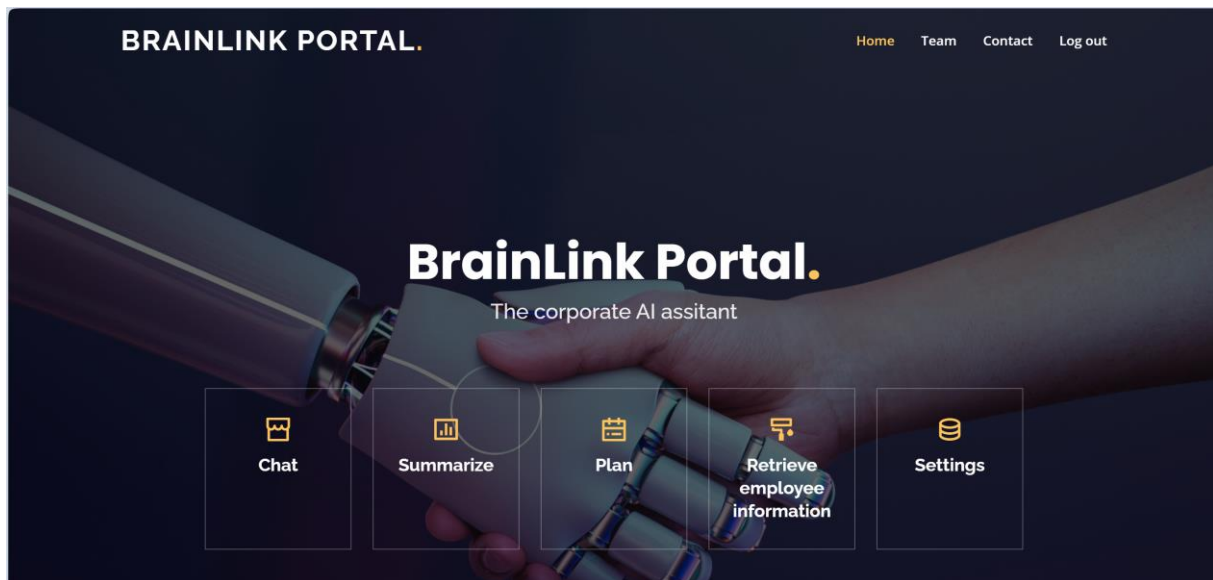
These use cases / features concern the tool interface.

ID	C2-NF1	Source	Model, Technology Evaluation, IT Infrastructure, ...	Author		Date		Status	approved
Name	Ease of use								
Category	Website interface								
Description	The tool must be implemented in a way that is intuitive to users. The tool is here to help employees, not to make them wasting time. The tool will have a nice graphical interface with a menu to select which feature to use.								
Acceptance criteria	Users should not overthink on how to use the Tool. It should be intuitive for everyone.								
Importance ¹	4	Urgency ²	3	Risk ³	2	Outlay ⁴	2		

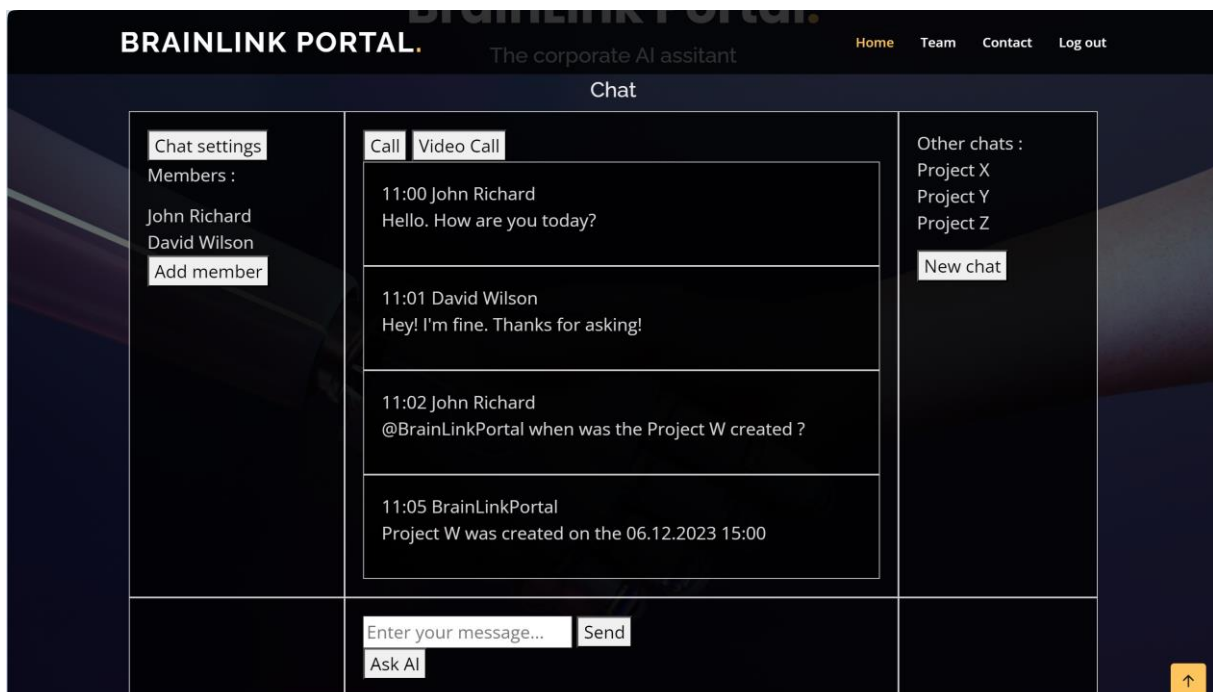
3.3. User Interface Prototype

Short prototype description and link to interactive prototype.

The image below shows the main page of BrainLink Portal. It contains a menu on the top right corner, and all available features on the bottom.



The image below shows a prototype of the chat feature. Several settings are available, such as chat setting (that can modify chat color, show/hide metadata to each message, etc.) and add members. We can navigate to other existent group chat or creating new ones.



The image below is the prototype of the summarizing documents. We have to possibility to directly copy/paste the text, or to import an existing document. The LLM will then begin to summarize the document. An history is also available on the right for quick access.

The screenshot shows the 'BRAINLINK PORTAL.' interface with a 'Summarize' tab. The main content area displays a summary for 'Minutes-06-12-2023.pdf'. The summary is structured with sections: 'Project Updates:', 'Client Feedback:', 'Budget Review:', and 'Upcoming Events:'. Each section contains a brief description of the content. On the right side, there is a 'History' section listing three documents: 'Minutes-03-12-2023.pdf', 'Minutes-04-12-2023.pdf', and 'Minutes-05-12-2023.pdf'. Below the history is a 'New chat' button. At the bottom, there is a form with three input fields: 'Enter your text', 'Select document', and 'Summarize'. A yellow 'Up' arrow button is located in the bottom right corner.

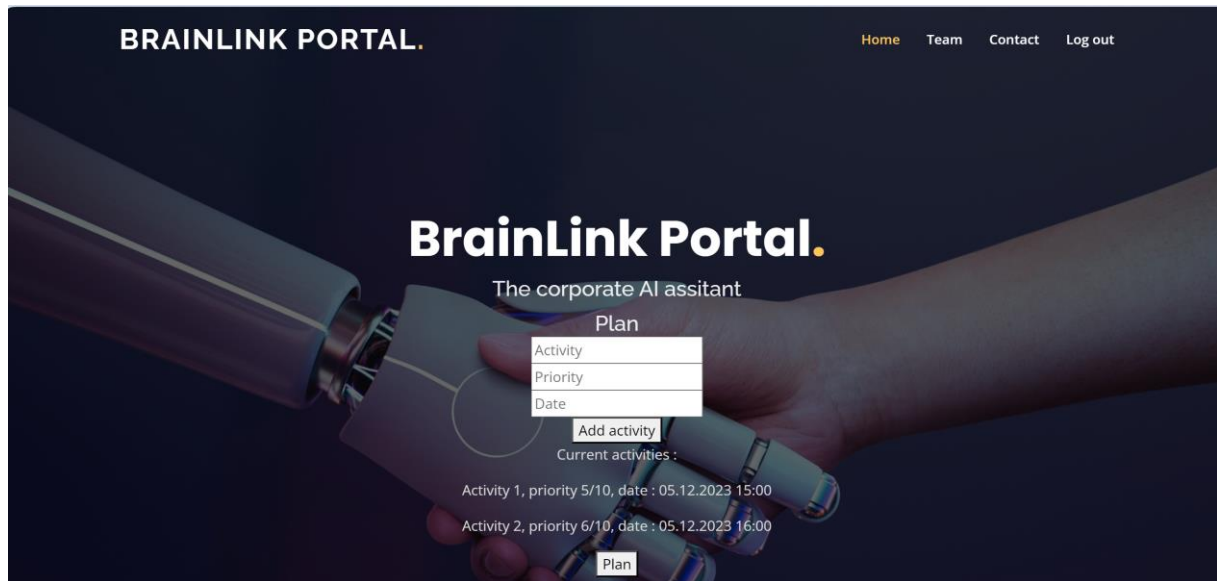
BRAINLINK PORTAL.		Summarize		Home Team Contact Log out	
<p>Summary of : Minutes-06-12-2023.pdf</p> <p>Project Updates: Team members provided updates on current projects, highlighting progress, challenges, and upcoming milestones.</p> <p>Client Feedback: Feedback from recent client interactions was discussed, and actions were identified to address any concerns or suggestions.</p> <p>Budget Review: Financial reports were presented, including a review of the current budget status and any adjustments needed.</p> <p>Upcoming Events: Discussion on upcoming conferences, workshops, or industry events that team members will be attending.</p> <p>The meeting concluded with any final remarks, expressions of gratitude, and a reminder of the next meeting date.</p>		<p>History :</p> <p>Minutes-03-12-2023.pdf</p> <p>Minutes-04-12-2023.pdf</p> <p>Minutes-05-12-2023.pdf</p> <p>New chat</p>			
<input type="text" value="Enter your text"/>		<input type="text" value="Select document"/>	<input type="text" value="Summarize"/>		

The image below shows the prototype of the data retrieving feature. It has some parameters to filter the retrieval.

The screenshot shows the 'BRAINLINK PORTAL.' interface with a 'Retrieve employee information' section. The section is titled 'The corporate AI assistant' and 'Retrieve employee information'. Below the title, there is a 'Use AI assistant' button. Underneath, there are two dropdown menus: 'Choose a department: IT' and 'Choose a seniority: Junior'. At the bottom, there is a 'Retrieve employees' button. The background features a stylized image of a hand holding a robotic arm.

BRAINLINK PORTAL.		Home Team Contact Log out	
<p>The corporate AI assistant</p> <p>Retrieve employee information</p> <p>Use AI assistant</p> <p>Choose a department: IT</p> <p>Choose a seniority: Junior</p> <p>Retrieve employees</p>			

Finally, the image below shows the planning feature. We have the possibility to create a new event with its name, the level of priority and the date.



(no link to prototype for the moment.)

3.4. Relevance Criteria

Each requirement is described with ...

- Importance: 5 = mandatory implementation; 4 = very important; 3 = important; 2 = normal; 1 = not important
- Urgency: 5 = must be implemented immediately, 4 = very urgent, 3 = urgent, 2 = normal, 1 = not urgent
- Risk/critical nature: 5 = unacceptable risk, 4 = very high risk, 3 = medium risk, 2 = low risk, 1 = no risk whatsoever
- Outlay: 5 = unacceptable outlay, 4 = very high outlay, 3 = high, 2 = reasonable, 1 = negligible or no outlay

4. Abbreviations and Glossary

Abbreviation	Explanation
HERMES	Methodology for projects and programs HERMES 5 is an eCH standard
KPI	Key performance indicator (Quantitative measure for NFR's)
NFR's	Non Functional Requirements

Table 3: Abbreviations and glossary