FINNA

I- Context

Where does this project come from?

In the current context in France, a lot of false information is circulating. So we did some research and we found that 39% of French people used the digital press to get information and that almost 30% of French people think they cannot detect false information according to Statista*.

*https://fr.statista.com/statistiques/960803/verification-informations-lues-jeunes-france/

What is the problem identified?

The problem currently is that there is far too much information being said but not being sourced.

Our problem is therefore:
How can we simplify the comparison of information for Internet users in order to be sure of its value?

And what would be the solution?

As a user, I want a tool that provides me with the sources of information and possibly identifies its first appearance, so that I can verify the authenticity and origin of the data.
To do this, I download the Finna extension from our website then I use this extension, I select the article, the tweet or an extract from it which seems dubious to me, then I have it checked by Finna who tells me will give the sources of this extract and I will deduce the reliability of the information myself.

What is the goal of the project?

Our goal is precisely to provide a simple and quick to use web extension, allowing the user to be able to consult all the sources which contain the information to be verified as well as a small paragraph which summarizes the information thanks to an LMM AI as well as all the sources it will have available, so the user will be able to form an idea for themselves about the information they will read, either simply by reading the small paragraph generated or by going to see the sources themselves.

Which tracks did we choose for our project?

The track that seems most obvious to us for this project would be the solution track because we directly respond to the problem cited above.

II- Technical specifications

Technologies used:

- MySQL (DB)
- React (Framework Front)
- NodeJs (Framework Back)
- Bash (Scripting)
- Git (CI/CD)
- Python (Script Scraping)
- AI LMM (Language to be defined)

Planned features:

- Website
- A web extension
- An LMM AI
- Scripts
- A database
- A server
- CI/CD
- Unit testing

Milestone for the tek4:

- Learn the different languages used for our project.
- Define the language for our AI / search for a suitable model (Meta-Llama, google-gemma…)
- Rework the Figma model to obtain a better UX/UI
- Implement the front-end of the Finna extension in React. The objective is to have an extension with a text field allowing us to send the request for information to our server.
- Implement the Finna extension server in NodeJs. The purpose of the server is to communicate between the Finna extension, the AI and the database.
- Implement a landing page in React which will allow the extension to be installed and users to provide feedback on Finna.
- Creation of the scraping script to retrieve articles from different newspapers and store them in a database.
- Development of the search algorithm that will link user entries to articles in the database.
- Implementation of the integration and development chain.

User Story:

The user copies a text to the clipboard or via the context menu, he pastes it into the search bar of our extension/web application which will have the front-end which will work thanks to React, the text is checked if there is has a corresponding article in our db, if there is one or more, we display a summary of the article using our llm AI which will summarize it then we will cite the sources where the article(s) appeared.
In the event that we do not have any articles corresponding closely or remotely to the text provided by the user, we specify that we have not found any sources while specifying that this does not mean that the information is false because they can be non-public sources or even investigations where the source is the journalist.
To populate our db, we will use scrapping scripts which will search for as many articles as possible on as many media as possible.

III- Non-technical specifications

To develop our community of first users, we want to use social networks to advertise in order to attract potential users. We can also promote Finna within the Epitech community in order to get technical feedback. The meeting sessions will be organized on Finna's discord.
We will also have a feedback area on the Finna website, where we will be able to collect and organize user feedback in order to improve the user experience and streamline the interface.

We decided to prioritize the relationship with our audience. To do this, we will implement a chat channel on Discord. Users will be able to give us feedback and we will be able to share the progress of developments. We also plan to implement a way for users to provide feedback to us on the Finna website.

Additionally, we want to increase Finna's visibility through social media. By making recurring posts and using current topics to sort out fact from fiction to show Finna's effectiveness.
We want to use X which would be the main platform:
React to tweets and show sources with the explanation of the facts,
Instagram/facebook to show, in the form of an immersive photo, the progress of the project by the developers and highlight the work behind Finna. This could be done with photos or a short/real format video.