Dear XWiki Community,

I am excited to apply for the Google Summer of Code 2023 with XWiki. My name is Yajushreshtha Shukla and I am a Computer Science student at Symbiosis Institute of technology , with a keen interest in Open Source development.

XWiki Community Presence:

- XWiki Chat Name and ID: [Chat Name and ID]
- XWiki Forum profile URL: https://forum.xwiki.org/u/kronos667/summary
- GitHub profile URL: https://github.com/Soyvor
- XWiki Jira profile URL: [Jira Profile URL]

Project:

- Project Name: AI based tagging of pages
- Project Description: [The goal of this project is to allow pages to be automatically tagged using a machine learning framework of your choice, based on the content of the page and its metadata.
- There are two similar project proposals in the area of automatic tagging that could give you inspirations how this could be implemented with your chosen framework: content-based tag suggestions (using TF-IDF to suggest tags) and Organizing Knowledge Using Topic Models (using LDA for a global analysis).

- If I would be willing and able to do other projects instead, they would be **Automatic Screenshots in Documentation**

- I am interested in this specific project because As an avid user of software documentation, I know how frustrating it can be to have outdated or incorrect screenshots in the documentation. The Automatic Screenshots in Documentation project would solve this problem by automatically generating and updating screenshots whenever changes are made to the user interface. This would greatly improve the accuracy and usefulness of the documentation, and would also save a lot of time and effort for developers who would otherwise have to manually update the screenshots. As a developer, I am excited about the opportunity to work on a project that would have a tangible impact on users and would contribute to the improvement of open source software documentation.

- Tentative project architecture or approach: the tentative approach would be to use Natural Language Processing (NLP) techniques to analyze the content of the page and extract meaningful keywords and phrases. The extracted keywords and phrases can then be used to suggest relevant tags for the page.The first step would be to select an appropriate NLP library that supports entity recognition and keyword extraction. Then, the code would need to be written to extract the relevant information from the page content and store it in a data structure.Once the relevant information has been extracted, the next step would be to match it against a pre-existing set of tags. This set of tags would need to be curated beforehand, and would ideally cover a wide range of topics and subjects.Finally, the tags suggested by the AI-based approach could be displayed to the user, who could then select the ones that they think are most appropriate. These tags would then be added to the page, allowing it to be more easily discoverable and searchable.

- Overall, the approach would involve using NLP techniques to extract relevant information from the page content, and matching this information against a pre-existing set of tags to suggest relevant tags for the page. The user would have the final say in which tags are added, allowing for greater control and customization.

  Expected project plan and timeline with milestones:

  Week 1-2:

  Set up the development environment and familiarize yourself with the XWiki platform.

  Understand the existing tagging system in XWiki.

  Research and choose an appropriate machine learning model for the project.

  Week 3-4:

  Train and test the machine learning model on a small dataset.

  Integrate the model with XWiki and implement a basic tagging system.

Week 5-6:

Improve the machine learning model's accuracy by refining the training data and adjusting the hyperparameters.

Implement a user interface for the tagging system and test it on a larger dataset.

Week 7-8:

Test the system on a wide variety of pages to ensure that it works effectively in all scenarios.

Refine the user interface and add additional features such as the ability to create custom tags.

Week 9-10:

Optimize the system's performance to ensure that it scales well to large repositories.

Test and debug the system thoroughly.

Week 11-12:

Write documentation for the system and create a guide for users to easily understand and utilize the new AI-based tagging system.

Note that this is just a tentative plan and the exact timeline and milestones may be adjusted based on progress made during the project.

Technical Skills:

- Java/Javascript/Web knowledge or experience: Java: I have a good grasp of Java programming language and have worked on multiple projects using Java as the primary language.

- Javascript: I have experience with JavaScript, including working with front-end frameworks such as React and Angular.
- Web: I have experience developing web applications using frameworks such as Spring and building RESTful APIs using technologies such as Node.js.
- Velocity/Hibernate/Database experience: I apologize, but I don't have any previous context of your technical skills and experiences. Please provide me the necessary information so that I can assist you better.

Open Source:

- I am interested in Open Source development because Open Source development has many benefits and opportunities that align with my values and goals. I believe that technology should be accessible to everyone and that knowledge should be shared openly. Contributing to Open Source projects allows me to work with a diverse group of people and learn from their experiences. It also gives me the chance to create something that can have a positive impact on people's lives. GSoC and XWiki are great opportunities to gain experience in Open Source development and work on meaningful projects.

- Previous Open Source development experience: https://github.com/Soyvor

Background & Education:

- School Name: Lucknow public college
- Specialty/Major: computer science
- Years attended: 12

Summer Plans:

- City/Country for summer: Lucknow/Pune , India
- Expected project time availability: june-september
- Jobs, summer classes, and/or vacations: Not really , just college classes for 4 hours a day

GSoC Experience

- Applied for/planning to apply for other GSoC projects: [Projects]
- Reasons for applying to GSoC and XWiki: There are several reasons why I am excited to apply for the Google Summer of Code program and specifically for a project with XWiki:
-

- Professional Development: I am eager to gain real-world experience working on an open-source project, collaborating with other developers, and honing my technical skills.
-
- Contribution to Open Source Community: GSoC provides an excellent opportunity to contribute to an open-source community and make a meaningful impact on a project that is widely used by developers and organizations.
-
- Mentorship: Working with experienced mentors will allow me to gain valuable guidance, feedback, and support throughout the project.
-
- XWiki is an exciting platform that has the potential to transform the way teams collaborate and manage their content, and I am eager to be part of its development and evolution.
-
- Overall, I am excited about the opportunity to participate in GSoC and contribute to XWiki's continued growth and success.

And finally, I believe you should take me because of my passion and dedication towards Open Source development, my technical skills and experience, and my commitment to contributing towards the success of XWiki and its community.

Thank you for considering my application.

Best regards,

Yajushreshtha Shukla