

20240128

1. Hackathon**
2. Moogles + agent
3. HE

Hackathon

1. 问NL对于rule其他concern
2. 成功和公司分不开
3. 问大家有没有其他comments

Hi everyone, I am Yuqi. It is a great honor to share with you our 2023.

As you can see in the video just now, in September last year, our team - name is Italian Cannon, participated in the 24-hour hackathon. Everyone enjoyed that and very happy to contribute to society and help others. Thanks Citi for providing such a great chance.

And we are very lucky that our solution was finally selected by the charity. *Recently, my luck is really good!*

Before I present our solution, I must thank NL for being our sponsor.

And I know that Citi pays great attention to brand all the time and is very cautious about any open source.

So we have carefully read the FOSS training again and I want to commit here:

1. Firstly, all team members are direct hire tech employees, as non-tech, contractors or consultants may not contribute on behalf of Citi directly.
2. Secondly, each developer will have a peer reviewer to ensure that any commit will NOT contain Citi business content, non-public and internal information.
3. And according to FOSS, as a Citi moderated project, we will only use Citi approved communication channels.

Thank you again NL for being our sponsor.

问题 (NL) : Do you have any other concern for the rules or anything else we need pay more attention, I will definitely keep it in mind and convey to my team members.

-- 1号停顿 --

Now, let me introduce our solution.

Our category is PA Youth Vote. As you know, a new round of US president election begins, and looks like Trump is about to win again.

During such kind of election, non-profit charity found that young people do not actively participate in voting.

They talked with some young people, most of them think voting is too complicated and also it's hard to obtain candidates' information.

So young people are unwilling to vote, and then so candidates will not pay attention to them, and then young people will not vote even more, a vicious cycle, this is the challenge now.

And for charity side, there're multiple platforms about voting, this makes it even more complicated to use.

So they want to build an unified platform and they prefer a mobile version, best like tiktok they said.

-- 打开手机和平板 --

Therefore, we created a responsive UI application that supports mobile, tablet and web page as well.

It supports register, online and offline vote, points for prize, and AI search for candidate video information.

-->> 翻页 - 第二页

We use angularJS in frontend, java and spring framework in backend, and python, faiss and embedding model for AI search.

And for the code, we are contacting Citi Open Source Office to put to Citi Public Github. We will continue work with charity this year for further cooperation.

At the end, regarding the reason for being selected, from judges' feedback, one is because the overall good UI and another is a new way to search candidates' video information via AI.

For this, I have to say thank you to Dane, James and Wendy because, to encourage innovation, both Shanghai and Dalian often organize innovation competitions, tech sharing and coding dojos.

We have gained a lot of tech and LLM experience from these events, and that helps us implement this tool so fast in this hackathon. Thank you.

问题 (all) : And till now, may I ask what do you think of our solution, any comments will be appreciated, that will help us present Citi's professional and brand better I believe.

5'25"

Moogles and Agents

1. 视频搜索, 摘要, 字幕
2. **paper**高亮
3. agent

The video search tool mentioned just now is the same one I shared last time, Moogles. If there is still time, I'd like to share some new progress while drinking tea. Ok?

-->> 翻页 - 第三页

In Moogoo, we currently have a total of 154 videos, with a total duration of more than 180 hours, and 220 users from China, Japan and the US. Average 267 page view in the past two weeks.

Greatly saves employees' time and improve productivity.

For example, I searched a question just now: 'What is market mid?' There's related videos below, we can go through the contents and choose one, and in this 55 mins video, the answer starts at 23 min timesampt.

And, Moogoo also helps generated some basic information, such as the mentioned names, techs and summary, topics. And there're the short description of the names and techs below.

In addition, last time, Hong raised a very good improment for us -- to add subtitles, because many good videos are in Chinese, for examle, these are the most viewed videos, the first one is from a senior desk user to introduce business, this is about clean code by Wendy, and this is about ChatGPT by Rob, in order to let all employees can understand them, we added English subtitles for all videos in Moogoo and Japanese for part of them. Thanks Hong for the idea!

-- 停顿 --

And there is one more feature in Moogoo which is for documents.

Because many colleagues are interested in new technologies, like Dane, they often read the latest papers, but usually they are very busy and reading an academic paper is really time-consuming. Let me share Moogoo's solution on this.

For example: search: 'How many hours' audio data used in whisper?'

Whisper is an audio to text paper published by openai, and is also one of the models used in Moogoo.

We can see that Moogoo found some docs here, open one of them, we can see that in this 28 pages paper, the answer is on the second page, and we have highlighted the answer here. six hundred and eighty thousand labelled audio data.

Through this, we hope you can find the information you care about much faster.

-- 轻松一刻 --

And there's also something interesting in this paper. There's a graph shows that, as the corpus increase, the error of model transcribe goes down. This line is the average of the error, and there're 3 languages obviously above this line, Hebrew, Korean and Chinese. They are much more difficult for model to transcribe. But I heard that NL have learnt both Hebrew and Korean, so I guess Chinese might be anther good choice for you!

Agent

Another feature provided by Moogoo is the agent which is a bridge established by LLM between natural language and computer language.

For example, I can ask 'What is the stock loan rate for stock which fii = xxx'. Moogoo will call a tool instead of generate by LLM. The service behind the tool is an expert system for predicting rates that we created by decision tree by learning users' manual rate change history.

Currently it sends email report daily, contains the most liquid stocks. Through the agent, users can query any stock at any time with chat.

In the future, chat based on LLM may be a direction for us to provide rapid iteration and delivery for clients.

-->> 翻页 - 第四页

All of this, the LLM and agent function, we provide API for others, currently there're average 100 requests a day in the past two weeks.

HE

Actually for the LLM API, there're some concerns about information security and computing efficiency from Wendy because this may disclose users' privacy.

She talked with our Dalian Tech Guild some times, and after research, we found that HE should be a viable solution for neural networks calculation. It allows calculate in ciphertext and user can decrypt the output with private key.

There are more details, but the time is really limited today, so hope we can talk next time.

Thank you!