

## Building question and answering module for improving campuswire search

### 1. Which tasks have been completed?

Tasks	Notes
Implement crawling function and cleaning up dataset	We have finalized our crawling function and dataset. We fetched over 1500 posts from campuswire for development and testing.  Code and data uploaded <a href="#">here</a>
Building question and answering module based on campuswire data	We have completed building a question and answering module.  It is based on Q-Q(User Question - Campuswire Questions) similarity matching which finds similar questions and gives answers from those posts.  Code and the note are uploaded <a href="#">here</a> . You can see the detailed notes about the flow of the executors and the sample results.

### 2. Which tasks are pending?

Tasks	Notes
Improving QA module	We are thinking about ways to improve the QA. Following is what we are considering <ul style="list-style-type: none"><li>Using the answers of a post<ul style="list-style-type: none"><li>Our QA is currently only using the first answer of a post.</li><li>There are several answers in a post. Moreover, answer threads are composed of the following questions and answers.</li><li>We are examining how to choose the answers from those threads.</li></ul></li><li>Additional Training<ul style="list-style-type: none"><li>We used the pre-trained Sentence Bert model for sentence embedding.</li><li>Considering additional training based on campus wire posts to improve the results.</li></ul></li></ul>

	<ul style="list-style-type: none"> <li>• The similarity between the document with long sentences and the one with short sentences.</li> <li>• (+) Additionally, we are thinking about which adapting BM25 ranking function for reranking the QA result will give the meaningful output.</li> </ul>
Run crawler on schedule to fetch newly added data.	Currently crawler fetch data from scratch every time. We are considering making a QA demo so data should be up to date.
Building web demo	Making a simple web demo which users can type in the question and get top-N answers. It may directly call the rest-api which our QA module provides.

3. Are you facing any challenges?

- No challenges until now. but we are considering several topics to improve the QA module(in the “Improving QA module” row). And those might be challenging topics for us to solve until the final submission.