

MOOC Wikification System

Jan 19, 2018

AN Yahui

anyahui.120@gmail.com

Previous Work

- Paper Writing
- Supplementary Experiments -- Spiders

Current Work

- Paper Writing

Reconstructed the writing content

| # Courses | #Posts | # Posts with key words | # Posts with concrete mentions |
|-----------|----------|------------------------|--------------------------------|
| 179 | 116, 962 | 20, 712 | 1, 110 |

- Experiments

1. annotate structure design

2. Mention extraction accuracy

| Ratio for posts have key words | Ratio for posts with concrete mentions |
|--------------------------------|--|
| 17.7% | 5.36% |

| Course name | ID of post | Post content | Mention content | SCI (Y/N) | Checked (Y/N) | Ture (Y/N) | Actual URL | Resolved URL | Resolved (Y/N) |
|-------------|------------|--------------|-----------------|-----------|---------------|------------|------------|--------------|----------------|
|-------------|------------|--------------|-----------------|-----------|---------------|------------|------------|--------------|----------------|

Key Words: ([Pp]roblem|[Qq]uiz*|[Qq]uestion|[Mm]odule|[Ll]ecture|[Ss]lide|[Vv]ideo|[Hh]omework)

Mentions: Key Words + (*[#]*)(*[0-9]+ *?[\.-]? *?[0-9]*)(?![0-9]*%)

Next Step

- To annotate more posts to finish the extraction accuracy and resolve accuracy;
- To finish the presentation of a typical example (meaning finishing the workflow of the system).