COLX 523 Annotation Materials

Bingyang Hou, Jae Ihn, Jiexie Kuang, Jinhong Liu, Min Zeng

Provide an initial draft of your annotation guidelines. This should include a basic description of the task and explanation of each possible label that an annotator might choose, with examples.

MIT OpenCourseware offers educational resources online, consisting of written syllabi containing course descriptions, lecture schedules, and other relevant information. Our goal is to annotate the reading materials and categorize them as either required or optional. However, curriculum designers may use different terms to label the materials. Here are some examples of the different labels that may be used, and these labels are the words we expect to annotate:

REQUIRED
- abbreviation is made, and is referenced in table for chapters
- assigned readings
- readings for discussion
- recommended (*should check context, sometimes they mean required, sometimes they mean optional.)
- required
- students should have personal copy
- students should have physical copy
- texts/textbooks (some are mentioned as optional, recommended, or additional textbooks--then it should go below)
- Almost required reading
- *, **, ***

OPTIONAL
- additional
- exploratory reading
- for more information
- further reading
- helpful
- interesting
- main, mainly
- optional
- primary
- recommended

- references
- related
- relevant
- supplementary
- useful

Annotation target includes
1. books
2. academic papers
3. journal articles
4. news articles

No need to include
1. handouts
2. lecture notes
3. generic websites--do include if it's a specific post, with title and author

A small amount of data (at least 10 examples) that has been processed into a format appropriate for annotation (e.g. a csv for upload to Mechanical Turk). You should include a link to the code that produced it from your corpus.

example link:
https://github.ubc.ca/MDS-CL-2022-23/COLX523_BiyaIvanJaeJessieMin/tree/min/milestone_2/annotator%20sample