

Agenda

1. Retrieval & Pooling
2. Relevance Annotations
3. Access & Outlook

Building a Test Collection

Retrieval & Pooling

- ❑ 10 retrieval systems will be used to retrieve documents
 - varied set of retrieval techniques (lexical, neural, query expansion, ...)
 - run via the TIRA platform
 - these baseline systems will be visible on the leaderboard too
- ❑ top-10 documents from each system will be pooled
 - pooled documents are deduplicated across systems
 - we expect 75 ± 25 documents per topic
 - should be between 1 – 2 hours of annotation work

Building a Test Collection

Relevance Annotations

- ❑ Binary Relevance: judging a document as non-relevant (0) or relevant (1)
- ❑ Relevance dependent on:
 - **Query** – provides information that helps the user make a decision
 - **Narrative** – provides information in the scope of the topics' narrative

Building a Test Collection

Relevance Annotations

- ❑ Binary Relevance: judging a document as non-relevant (0) or relevant (1)
- ❑ Relevance dependent on:
 - **Query** – provides information that helps the user make a decision
 - **Narrative** – provides information in the scope of the topics' narrative
- ❑ Relevance is not dependent on:
 - **Credibility** – a user can find non-credible documents useful, or misjudge their credibility, thus not influencing relevance
 - **Correctness** – judgments are made from the POV of a search user who does not know the correct answer

Building a Test Collection

Relevance Annotations

- ❑ Binary Relevance: judging a document as non-relevant (0) or relevant (1)
- ❑ Relevance dependent on:
 - **Query** – provides information that helps the user make a decision
 - **Narrative** – provides information in the scope of the topics' narrative
- ❑ Relevance is not dependent on:
 - **Credibility** – a user can find non-credible documents useful, or misjudge their credibility, thus not influencing relevance
 - **Correctness** – judgments are made from the POV of a search user who does not know the correct answer
- ❑ Documents are automatically non-relevant if they:
 - are written in a language other than English
 - contain adult material or spam
 - are unreadable or broken

Building a Test Collection

Access

- Relevance annotations will be handled through Doccano [\[link\]](#)
- Each group will receive by mail:
 - their finalized group name and TIRA account
 - a Doccano account per topic (linked with the topics' project)

Account data: **End of week** (Thursday/Friday)

Annotation Completion Due: **06.05.2024**

Outlook

Materials in Preparation for M-III

`https:
//github.com/tira-io/teaching-ir-with-shared-tasks`