

Document
Retrieval

Event Extraction
& Timeline

Multi-model
Scoring

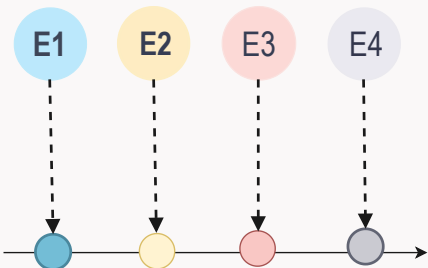
Filtering &
Verification

Question
Construction

- Relevant Docs
- Distractor Docs



Sentence-level Events



Timeline Construction

Extract → Merge →
Disambiguate

Candidate Pairs
(Cause → Target)



GPT-4.1

S1



Gemini-2.0

S2



Claude-3.7

S3

Compute Variance



Variance



Keep

Variance



Discard



Annotation

Majority → Keep

Disagree → Remove

Target Event

4 Options

- A. ...
- B. ...
- C. ...
- D. ...

Multi-select answers

Output: "B,C"