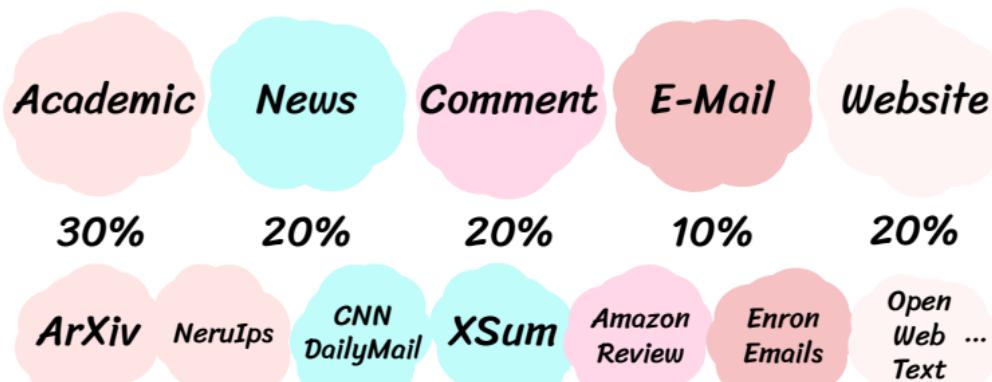


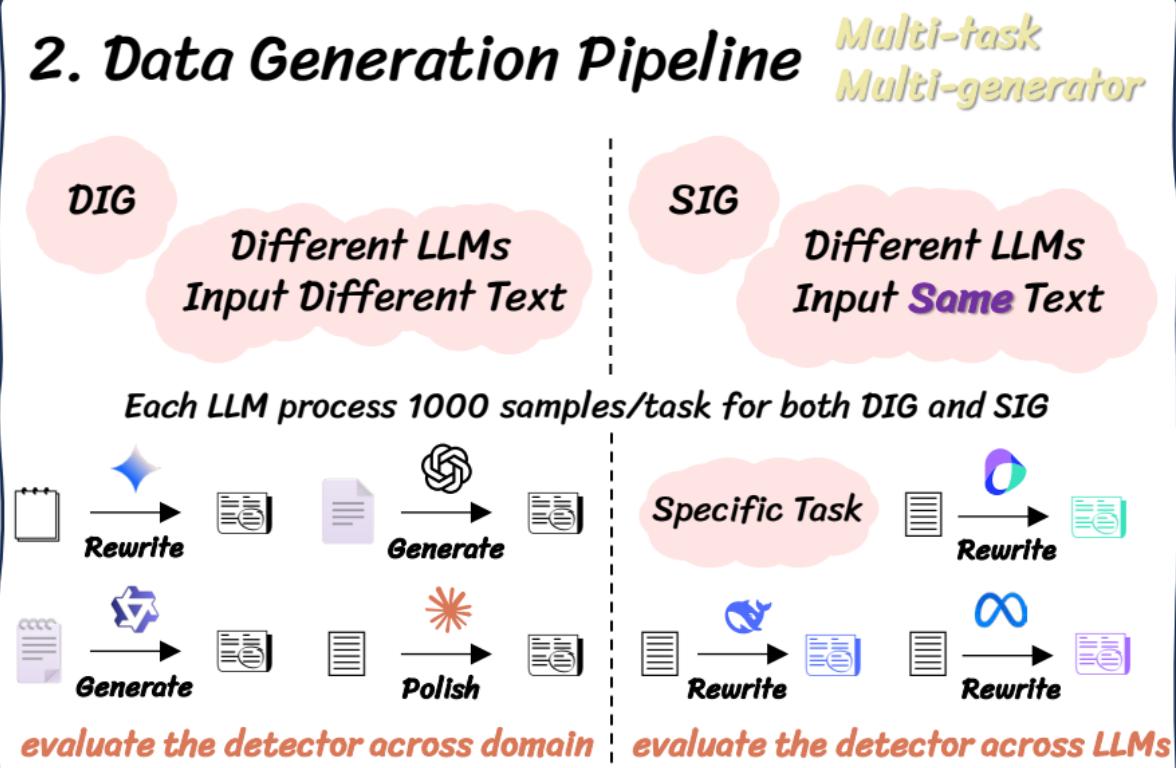
MIRAGE Benchmark Construction Workflow

1. Sampling Human-Written Text



Length in [100, 200)
Remove all the '\n' and '\r'

2. Data Generation Pipeline



3. Other

- Ensure Length in [90, 220)
(allow a 10% variation in length)
- Exclude harmful content
(finished while generating data using API)
- Remove meaningless text
(mostly consist of symbols or punctuation)
- Remove all the '\n' and '\r'
(Prevent relying on symbolistic features)

Length
Letter
Styles

16 common writing styles
formal, oral, academic, literary, critical, narrative, descriptive, lyric, objective, subjective, original, casual, expository, argumentative, journalistic, poetic