



Stage	Description	Inputs	Outputs
retrieval	Articles need to be retrieved from multiple sources	Public datasets, news feeds, news aggregators	Articles (headline, body, source, URL, publication date)
processing	Articles are processed: tokenisation, parsing, and attaching single-article framing features	Articles	Documents with: <ul style="list-style-type: none">- Parsing, segmenting- Embedded representation- Framing
linking	Similar articles are linked as they cover the same event, and similar sentences are recognised across documents	Processed articles	Similarity links between <ul style="list-style-type: none">- Articles- sentences