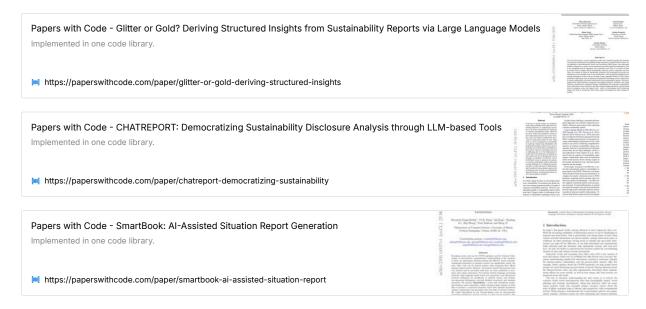
# **Report Generation**

### ▼ Paper with code on sustainability



## **▼** Glitter or Gold?

Report Understanding on sustainbility reports, similar to ChatReport.

Task Description: Get structured information from sustainbility reports in the form of PDF

https://arxiv.org/pdf/2310.05628.pdf

https://github.com/saturnmars/derivingstructuredinsightsfromsustainabilityreportsvialargelanguagemodels

#### Some Important Points we can learn

- · PDF pre-processing methods:
  - extract structured insights from sustainability reports;
  - o sections segmentation.

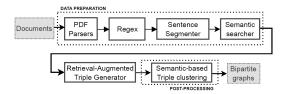


Figure 1: Our proposed approach and its components. Given a collection of textual documents as inputs and preprocessed using different NLP methods, semantically structured insights are extracted by the retrieval-augmented triple generator. Bipartite graphs are created after performing a semantic-based triple clustering.

Pipeline

Report Generation 1

# **▼** SmartBook

Report Generation on news.

Task Description: given large volumes of news data, produce a structured report with multiple hypotheses (claims) summarized and grounded (linked) to sources with factual evidence.

https://arxiv.org/pdf/2303.14337v2.pdf

https://github.com/blender-nlp/SmartBook

Report Generation 2