Appendix

A An example of prompt used to extract the number of submitted/accepted papers

It contains two messages for the chat models where the first one is the system message and the second one is the human message. Furthermore, it contains two examples (two-shots) to specify the task as well as the output format. The {preface_text} variable in line 27 is replaced with the conference proceedings text from which the submitted/accepted paper counts to be extracted.

```
acceptance_ratio_system_template = ``You are a data entry clerk
     and you know how to read some text and extract the
     requested information. You always follow the expected
     output format shown in the examples precisely. You only
     extract information in the text and do not introduce any
     new facts."
 acceptance_ratio_human_template = ``Given a preface from a
     conference proceedings, extract the number of research
     papers submitted and accepted to each track. Output the
     results in the CSV format.
 Here are two examples:
 Preface: The main scientific program of ESWC 2023 contained 41
     papers selected out of 167 submissions (98 research, 23
     in-use, 46 resource): 19 papers in the research track, 9 in
     the in-use track, and 13 in the resource track. The overall
     acceptance rate was 24% (19% research, 39% in-use, 28%
     resource).
 Output:
 track, submitted, accepted
12 research, 98, 19
13 in -use, 23, 9
14 resource, 46, 13
 --- complete ----
 Preface: The research papers program received 245 full paper
     submissions, which were first evaluated by the Program
     Committees of the respective tracks. The review process
     included evaluation by Program Committee members,
```

discussions to resolve conflicts, and a metareview for each potentially acceptable borderline submission. After this a physical meeting among Track and Conference Chairs was organized to see that comparable evaluation criteria in different tracks had been used and to discuss remaining borderline papers. As a result, 52 research papers were selected to be presented at the conference. The proceedings also include ten PhD symposium papers presented at a separate track preceding the main conference, and 17 demo papers giving a brief description of the system demos that were accepted for presentation in a dedicated session during the conference.

```
Output:
track, submitted, accepted
research, 245, 52
PhD symposium, -, 10
demo, -, 17
--- complete ----

Now extract the number of research papers submitted and
accepted to each track from the following text. Output only
the CSV content and "--- complete ----" as the last line.

Preface: {preface_text}"
```

B SPARQL queries used to extract the papers of a given proceedings and their corresponding authors from the DBLP KG.

```
# get the list of papers in a given proceeding
#identifier <__proceedings_uri__>
PREFIX dblp: <https://dblp.org/rdf/schema#>
SELECT ?paper ?title ?doi ?pages ?year WHERE {
   ?paper a dblp:Publication, dblp:Inproceedings;
       dblp:title ?title;
       dblp:doi ?doi;
       dblp:pagination ?pages;
       dblp:yearOfPublication ?year;
       dblp:publishedAsPartOf <__proceedings_uri__> .
   FILTER (STRSTARTS(str(?doi), "https://doi.org/"))
}
# get the list of authors associated in papers of a given
# proceeding identifier <__proceedings_uri__>
PREFIX dblp: <https://dblp.org/rdf/schema#>
SELECT ?paper ?title ?name ?ordinal ?orcid ?wikidata ?scholar {
   ?paper a dblp:Publication, dblp:Inproceedings;
       dblp:title ?title;
       dblp:hasSignature ?sign;
       dblp:publishedAsPartOf <__proceedings_uri__> .
   ?sign dblp:signatureDblpName ?name;
       dblp:signatureCreator ?dblp_person;
       dblp:signatureOrdinal ?ordinal .
   OPTIONAL { ?dblp_person dblp:orcid ?orcid }
   OPTIONAL { ?dblp_person dblp:wikidata ?wikidata }
   OPTIONAL { ?dblp_person dblp:webpage ?scholar .
```

C SPARQL queries used to extract the papers of a given proceedings and their corresponding authors from the DBLP KG.

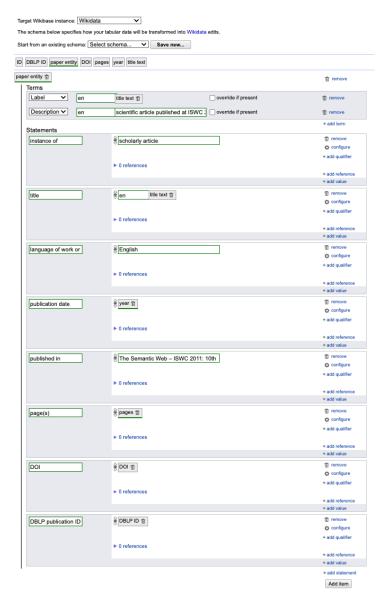


Fig. 7: Screenshot Open Refine schema that maps the papers extracted from the SPARQL query to Wikidata for generating Quick Statements.