1 备份旧的survey.xlsx

2 将在新收集.xlsx中的新论文人工加入survey.xlsx 中

3 执行3-check.py文件检查excel中的论文信息和dblp检索到的论文信息是否一致

4 4-unrepeat.py同一类别下的同名论文去重，注意可以允许一个paper在多个类别里

5 excel排序，先ctrl+A全选，然后点击排序和筛选，然后自定义排序，然后按图2执行即可

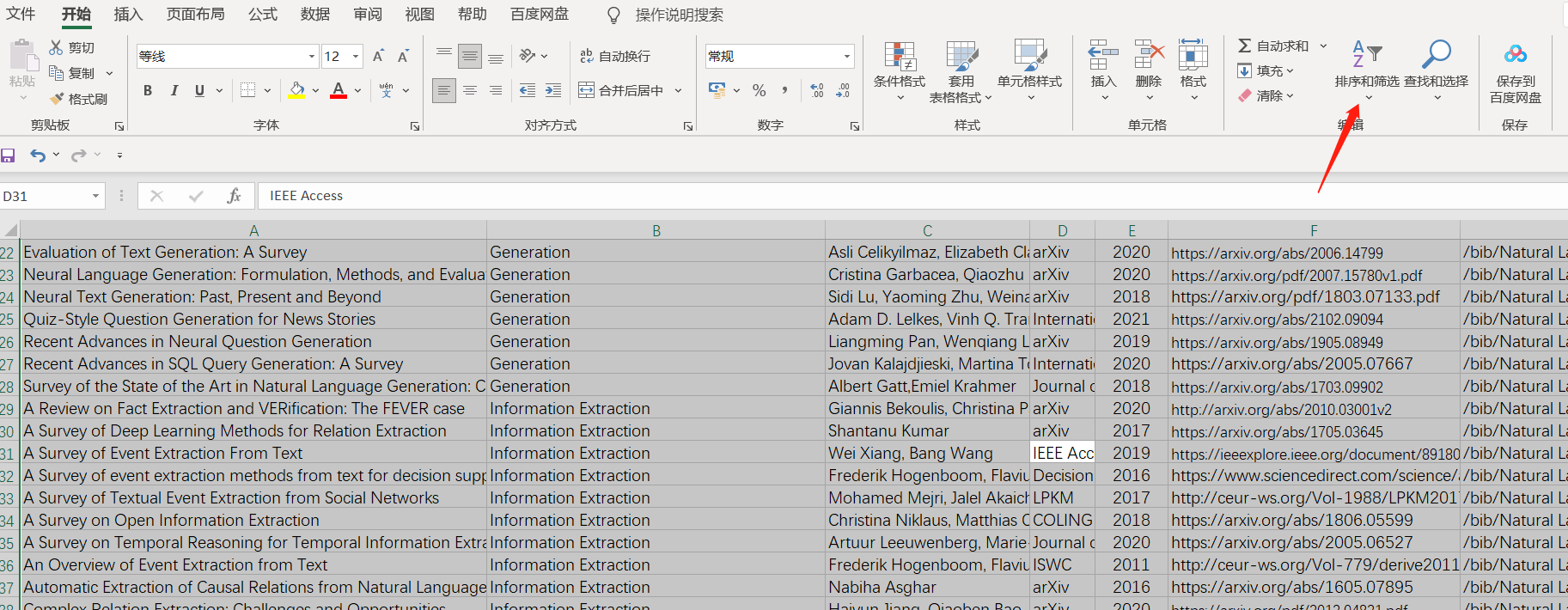


图1 内容全选

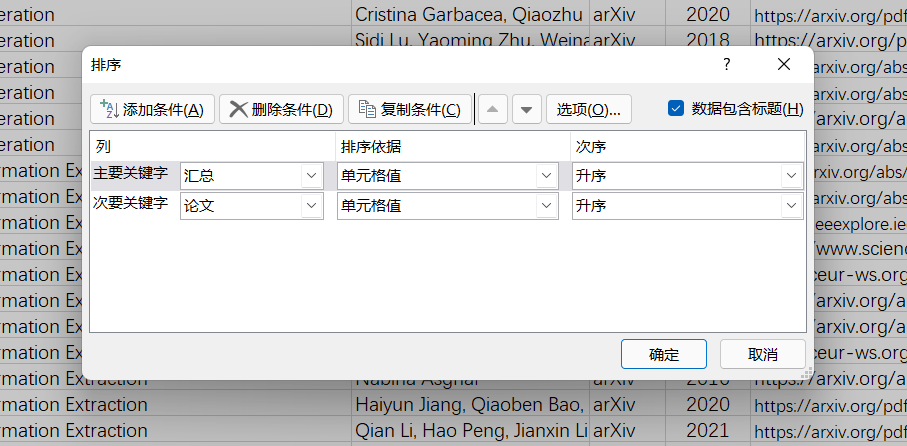


图2 排序规则

6 excel表格bib筛查，执行bib-excel1.py

根据“论文首个单词、时间、作者姓氏”检测出excel中错误填入的bib信息，然后人工进行修改

注：bib同名情况，是以第一步.pdf中后面加数字来解决的。执行bib-excel1.py

脚本出现这样的报错时不用管

7 excel表格同一汇总下，同名bib筛查，执行bib-excel2.py

如果出现同名的情况，按照第一步.pdf中的规则：在excel表格的最后一列中给其后面加数字，如：

Su2020A.md Su2020A1.md Su2020A2.md ……

8 文件夹bib筛查 bib-files.py（该脚本一般要运行两次，得到的输出信息更完善）

检测出文件夹中缺少那个paper的bib，然后该脚本会自动创建一个空的md文件，我们需要手动查找bib信息并添加到该空md文件中。

同时，检测出文件夹中多余的md文件，输出到extra-bib文件夹中，该步骤结束之后，我们可以手动删除这个文件夹

9 检查md文件，执行md.py

检测缺少``` 和 文件名与第二行后面的索引不匹配的错误，并自动修改

10 执行excel.py 得到out.txt

11上传github

git init

git remote add [git@github.com:NiuTrans/ABigSurvey.git](mailto:git@github.com:NiuTrans/ABigSurvey.git)

git fetch origin master:temp

git merge temp

更新文件（使用VSCODE更新git会方便很多）

git add/rm xxx

git commit -m “”

git push origin master