

批量下载某研究方向重要文献

目录

安装Python及依赖包	1
python下载安装	1
依赖包安装	1
BeautifulSoup	1
requests	1
xlrd	1
通过AI搜索关键词获取文献列表	1
爬取文献保存至本地	2
参考	3

安装Python及依赖包

python下载安装

- windows用户请到Python官网<https://www.python.org/downloads>下载相应的版本, 本教程在 **version 3.10.0** 测试可行, 建议安装3.7以上版本。
- Linux系统自带python, 一般不需单独安装, 除非版本太低, 则需升级。

依赖包安装

BeautifulSoup

以下命令可以在Windows中的CMD/Powershell或Linux终端中运行。

```
pip install BeautifulSoup4
```

requests

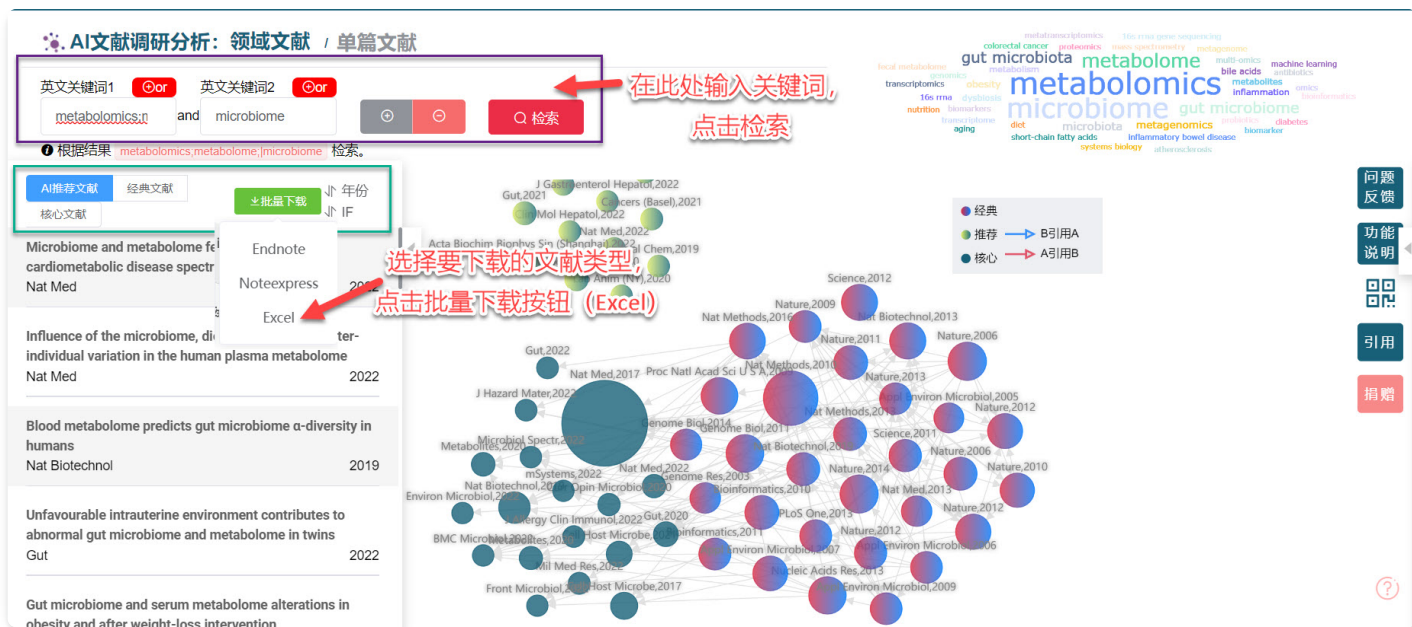
```
pip install requests
```

xlrd

```
pip install xlrd
```

通过AI搜索关键词获取文献列表

假设我们需要查找 **代谢组学** 和 **微生物组** 联合研究的文章, 进入AI based文献检索网站<https://www.citexs.com/Paperpicky>, 输入关键词“metabolomics;metabolome”和“microbiome”, 并点击批量下载, 保存为Excel格式(如下图所示)。



AI文献检索

下载后会得到“文献.csv”，打开文件检查各列是否与下图匹配。从左至右依次为Title、Journal、IF、DOI、PMID、Pub_Date、Url，若不匹配，请先修改，如果第二例为作者信息，那么可将该列删除。

Title	Journal	IF	DOI	PMID	Pub_Date	Url
Microbiome	Nat Med	87.241	10.1038/s41591-022-01688-4	35177860	2022/2/15	https://www.citexs.com/detail?pmid=35177860
Influence of	Nat Med	87.241	10.1038/s41591-022-02014-8	36216932	2022/11/15	https://www.citexs.com/detail?pmid=36216932
Blood	Nat Biotechnol	54.908	10.1038/s41587-019-0233-9	31477923	2019/10/1	https://www.citexs.com/detail?pmid=31477923
Unfavourable	Gut	31.793	10.1136/gutjnl-2021-326482	35387876	2022/12/15	https://www.citexs.com/detail?pmid=35387876
Gut	Nat Med	53.44	10.1038/nm.4358	28628112	2017/7/1	https://www.citexs.com/detail?pmid=28628112
Metabolomics	BMC Microbiol	4.465	10.1186/s12866-022-02486-1	35264111	2022/3/9	https://www.citexs.com/detail?pmid=35264111
Alterations of	Arthritis	15.483	10.1002/art.42331	36054683	2022/8/23	https://www.citexs.com/detail?pmid=36054683
Of microbiome	Lab Anim (NY)	12.625	10.1038/s41684-020-00670-2	33046900	2020/11/1	https://www.citexs.com/detail?pmid=33046900
Role of dietary	Cell Host	31.316	10.1016/j.chom.2020.12.012	33440171	2021/3/10	https://www.citexs.com/detail?pmid=33440171
Microbiome	J Gastroenterol	4.369	10.1111/jgh.16043	36287112	2022/10/26	https://www.citexs.com/detail?pmid=36287112
Metabolome-	Metabolites	5.581	10.3390/metabo10050181	32369899	2020/5/1	https://www.citexs.com/detail?pmid=32369899
Influence of	Gut	23.059	10.1136/gutjnl-2019-319188	31953253	2020/8/1	https://www.citexs.com/detail?pmid=31953253
Microbiome	Clin Mol	8.337	10.3350/cmh.2022.0171	35850497	2022/7/15	https://www.citexs.com/detail?pmid=35850497
Altered faecal	Gut	23.059	10.1136/gutjnl-2020-323565	34035120	2021/5/25	https://www.citexs.com/detail?pmid=34035120
Microbiome-	Acta Biochim	3.511	10.3724/abbs.2022149	36269135	2022/1/25	https://www.citexs.com/detail?pmid=36269135
Association of	J Allergy Clin	14.29	10.1016/j.jaci.2022.02.005	35196534	2022/8/15	https://www.citexs.com/detail?pmid=35196534
Fusarium	Microbiol	9.043	10.1128/spectrum.01226-22	35766498	2022/8/31	https://www.citexs.com/detail?pmid=35766498
Illuminating	Environ	5.476	10.1111/1462-2920.16242	36222155	2022/11/15	https://www.citexs.com/detail?pmid=36222155
Biodegradable	J Hazard	14.224	10.1016/j.jhazmat.2021.127282	34607029	2022/2/15	https://www.citexs.com/detail?pmid=34607029

文献.csv内容

爬取文献保存至本地







本教程脚本基于 **大阔同学** 脚本修改而来，添加了参数，避免用户修改源代码；增加了随机user-agent，避免下载次数过多被屏蔽。原理是基于文献DOI，利用爬虫通过SCI-HUB下载文献。因此，必需要在 **文献.csv** 中提供DOI。有些文章不在SCI-HUB中，或者网络环境较差，则会下载失败，失败信息写入 **error.log** 中。

将 **DownloadPaper.py** 和 **文献.csv** 放在同一目录下，并在该目录下 **shift+右键** 打开Powershell窗口(俗称cmd)，输入下面的命令，回车即可下载。如果电脑安装了WSL Ubuntu，也可以进入Linux终端。

```
python DownloadPaper.py -f 文献.csv -o Papers -c 20
```

参数解析：

- f:指定包含文献信息的文件
- o:将文献下载至该参数指定的路径中
- c:影响因子阈值, 低于该阈值的文献将不会下载

名称	修改日期	
 2017-07-01-Nat Med-53.440-Gut microbiome and serum metabolome al...	2023/4/24 21:43	M
 2019-10-01-Nat Biotechnol-54.908-Blood metabolome predicts gut mic...	2023/4/24 21:42	M
 2020-08-01-Gut-23.059-Influence of gastrectomy for gastric cancer treat...	2023/4/24 21:43	M
 2021-03-10-Cell Host Microbe-31.316-Role of dietary fiber in the recove...	2023/4/24 21:43	M
 2022-02-15-Nat Med-87.241-Microbiome and metabolome features of t...	2023/4/24 21:42	M
 2022-11-15-Nat Med-87.241-Influence of the microbiome, diet and gen...	2023/4/24 21:42	M

下载结果

参考

- [快速获取某方向研究重要文献并批量下载](#)
- [Github](#)