

采购中心爬虫需求 处理流程规范

文件版本历史

文件版本	修订日期	修订人	审核人	批准人	修订说明
V1.0	2024/09/04	Lennon	Miang	Joel	初版

目录

一、 需求整理与提交	1
1. 对接负责人自行处理:	1
2. 邮件提交需求:	1
二、 邮件与附件格式	1
1. 邮件主题:	1
2. 邮件正文:	1
3. Excel 附件格式要求:	1
三、 需求处理	2
1. 需求处理流程:	2
2. 查重处理:	2
四、 其他事项	2

为提高爬虫需求处理的效率，确保信息传达的准确性和工作流程的规范化，现制定如下流程规范。请采购中心各事业部参照执行，并给予配合。

一、需求整理与提交

1. 对接负责人自行处理:

- 对于 RockAuto、Dorman、eBay、Cardone、Standard、SpectraPremium、FourSeasons 等网站的爬虫需求，因其操作简单且已提供完整代码，原则上由各事业部的对接负责人自行进行爬取。
- 如果在操作中遇到问题，可联系 Lennon 进行技术支持。

2. 邮件提交需求:

- 对于其他较为复杂的网站爬虫需求，请对接负责人整理需求，以 Excel 文件形式作为附件发送 Lennon，并抄送本部门部长和 Joel。

二、邮件与附件格式

1. 邮件主题: 请注明“【爬虫需求】+部门名称+简要内容”。例如：“【爬虫需求】项目投放部 - Steering Damper”。

2. 邮件正文:

- 需求描述: 简要描述爬虫需求的目标数据。
- 时间要求: 请明确需求的时间要求或对紧急程度做出说明。
- 其他说明: 包括文件输出样式等特殊要求，都请在邮件正文详述。

3. Excel 附件格式要求:

- Excel 文件应包含简洁但清晰的爬虫需求信息，第一列为需求序号，第二列为具体的目标网址。

	A	B
1	No	Url
2	1	https://www.dormanproducts.com/gsearch.aspx?type=keyword&origin=keyword&q=clip
3	2	https://www.dormanproducts.com/gsearch.aspx?type=keyword&origin=keyword&q=insert

- 请确保每个需求都有明确的网址，避免“爬一下 Dorman 家的卡扣”这种含糊不清的描述。

三、需求处理

1. 需求处理流程:

- Lennon 在收到需求后会进行初步评估，如有疑问会与对接负责人进一步沟通。
- 需求确认后，数据部将安排具体的任务，并尽量在规定的时间内完成。
- 爬虫任务完成后，数据部会通过邮件反馈获取到的数据。

2. 查重处理:

- 如无特殊要求，请各事业部对接负责人自行完成查重工作。
- 如果有特殊的查重需求，请在邮件正文中进行详细说明，并将邮件抄送 Nolan。

四、其他事项

- 请对接负责人在提交需求时，尽量确保信息准确，避免因描述不清导致的延误。
- 数据部会对每项爬虫需求进行记录，并定期向采购中心领导汇报进展情况。
- 如有紧急需求或特别事项，请提前告知 Lennon，以便及时协调处理。