Minutes of Meeting

2017/7/27-视频会议（瞩目软件）

Recorder:李宜瑾

Participant：黄凯波（老师）、陈栋、闵志强、陈玉玲、李宜瑾

I Meeting Content

1. 黄老师将电脑屏幕共享，使用pycharm创建测试的脚本，使用脚本来创建数据库。创建user顶点变量和friend边变量，执行后在OrientDB中可以看到数据库结构。将脚本上传至GitHub中，供练习使用。
2. 找可以进行分类和爬取的网站，找到VIVO网站（https://vivo.usda.gov/individuallist?vclassId=http%3A%2F%2Fxmlns.com%2Ffoaf%2F0.1%2FOrganization），可以进行下一步的分析。
3. 已经找到了56个期刊，分为26个类，已提交到GitHub agricultural form中了。

II Plan

确定了实施步骤：

1. 建立基础，研究库，分配到个人；
2. 解析；
3. 可视化（orientDB）；
4. 研究一下可视化效果mike。

在数据和结构方面：

确定爬取的内容和要当下来的内容。

1. 之前找的农业期刊进行精选，选择4-5个进行研究，侧重点在表格中标为绿色和橙色的部分。另外优先找到企业的具体信息。
2. 爬取VIVO中的信息，爬下来之后在库中生成一个脚本。

在技术方面：

1. 认识和了解爬虫。每人配置好Python环境。

执行人：全体小组成员（闵志强负责）

时间：7月30日之前

1. 进行脚本导入、orientDB的连接，每人机器上成功运行起orientDB。

执行人：全体小组成员

1. 分享爬虫代码给小组成员。

执行人：陈栋

**预期成果**：库的数据填充。

**下一次会议的展示**：图库结构。