实习总结

李诚

在京东方实习期间，我主要工作是网络爬虫的维护，以及后面做的导诊机器人的自然语言处理工作。

1. 网络爬虫维护
2. 外部监控模块的搭建

目的是对爬虫的状态进行监控，以确定爬虫程序是否在工作。

具体做法是读取当前各个网站抓取的网页数量；与上一次的网页数量做对比，确定网页数量增量；将当前网页数量以及网页增量打包发送至个人邮箱。

程序在