

## 写在最后

更新时间：2019-07-12 09:44:50



“从不浪费时间的人，没有工夫抱怨时间不够。”

——杰弗逊”

首先很感谢您能读到这里，其次是感谢慕课给我们这样一个共享知识的大平台。完成这个专栏对我来说非常不容易，在撰写这个专栏的时候正好是我工作非常紧张的时候，而且专栏的更新时间又安排得非常紧，在家人的支持下也总算是挺过来了。

在约此专栏之前我出版了一本关于爬虫方面的书名为《Python绝技——虫术》，我本以为写专栏应该是一件比较轻松的事，但作为一名专业的软件从业者，我认为只是一味重复地做同样的事而没有任何的进取那是羞耻的。所以这个专栏是我花了比写书更多的心思准备的，包括所有示例都是全部重新设计重新编写代码和测试的。

这也成为了我一个难得的历程。能写到这里我对scrapy的认知也得到了不少的提升，相应地在爬虫方面的技术也有了长足的进步。在最后是有感而发，想给读者们共享一些关于学习python的心得。

我在微信群里面知道不少的童鞋是完全不懂python或者初学python想通过学习爬虫进入python这个领域。确实这是一个不错的选择。

之所以说这是一个不错的选择是因为爬虫技术的涉面其实非常广，虽然说这可以算是一门后端的技术，但是它却涉及了很多的相关领域，我们来盘点一下除了Scrapy之外我们还接触了什么：

1. HTML
2. CSS
3. RSS
4. HTTP 基本通信基础
5. Javascript
6. 正则表达式

7. Splash
8. Selenium —— E2E测试的基础
9. Docker —— 容器技术不单单只是方便安装服务还是DevOps的基础
10. Redis — Key/Value式的NoSQL
11. MongoDB —— Document式的NoSQL
12. 阿里云OSS
13. SQL (PostgreSQL)和SqlAlchemy —— 正统SQL
14. TDD —— 测试驱动开发

笼统地算了一下也有超过14个大知识领域了，对这可不是知识点而是一个领域，每个领域要深入进去都要花上不少的学习时间。但这是值的而且只要你坚持走这条路你将会接触得更多，这就是后端开发。

要精进爬虫技术我的方法就是疯狂地编码，爬取各种各样的网站，而且是那些最刁钻被认为是无可攻破的网站，例如微信和淘宝。当然这要以手法为前提，技术用对了才有价值。当完全精通了Scrapy，做过了多种多样的实战，你可以尝试一下其它的爬虫框架例如pySpider。Scrapy已经赋予你一种非常完善的爬虫架构思想，带着这种思想学习会让你看到更多，走得更远的。

爬虫技术仅仅只是为你打开了通向浩瀚的后端编程的知识大门，学会了scrapy也只是学到了python相关技术的冰山一角而已。Python本身不算强大，它只是简单而已。Python最强大的地方在于它的粘合性，有C的地方就会有它的身影，这也是她被称之为“胶水语言”的一个来由。作为一门教学级语言各个科学领域都为python社区贡献了不少及为出名的工具包，做爬虫我们可以用Scrapy框架、做Web可以用Flask、AI当然首选Tensorflow、做视频处理有OpenCV、做嵌入式开发有Micropython，诸如此类多不胜数了。

最后，不要试图在Windows环境下来进行python开发，那是一条邪道！Windows是没有良好的系统概念的，那只适合于.netframework生存，而只有linux才是python最好的生存土壤，也是能打开一扇更广阔天地的大门。有很多的python技术在windows上用是无限的巨坑，生命如此的短暂为何要为windows填坑呢？