



1.1 爬虫程序开发环境

深圳信息职业技术学院

Shenzhen Institute Of Information Technology

教师：黄锐军

目录

COMPANY

1.1.1 爬虫程序简介

1.1.2 python开发环境搭建



PART ONE

爬虫程序简介



one

爬虫程序是一组客户端程序，它的功能是访问web服务器，从服务器中获取网页代码，网页代码中包含了很多各种各样的数据信息，程序从中提取所关心的数据，把数据整理后存储在本地的数据库中，这些数据将应用在数据分析等领域中。

two

爬虫程序爬取的数据往往很多，而且相关的数据往往分布在很多不同的网页中，甚至分布在相关联的很多不同的网站中，爬虫程序必须能按链接自动往返于这些不同的网站中去爬取数据，一个爬虫程序爬取成百万上千万条的数据是常有的事，怎么样设计一个高效率的爬虫程序成了我们学习的重点

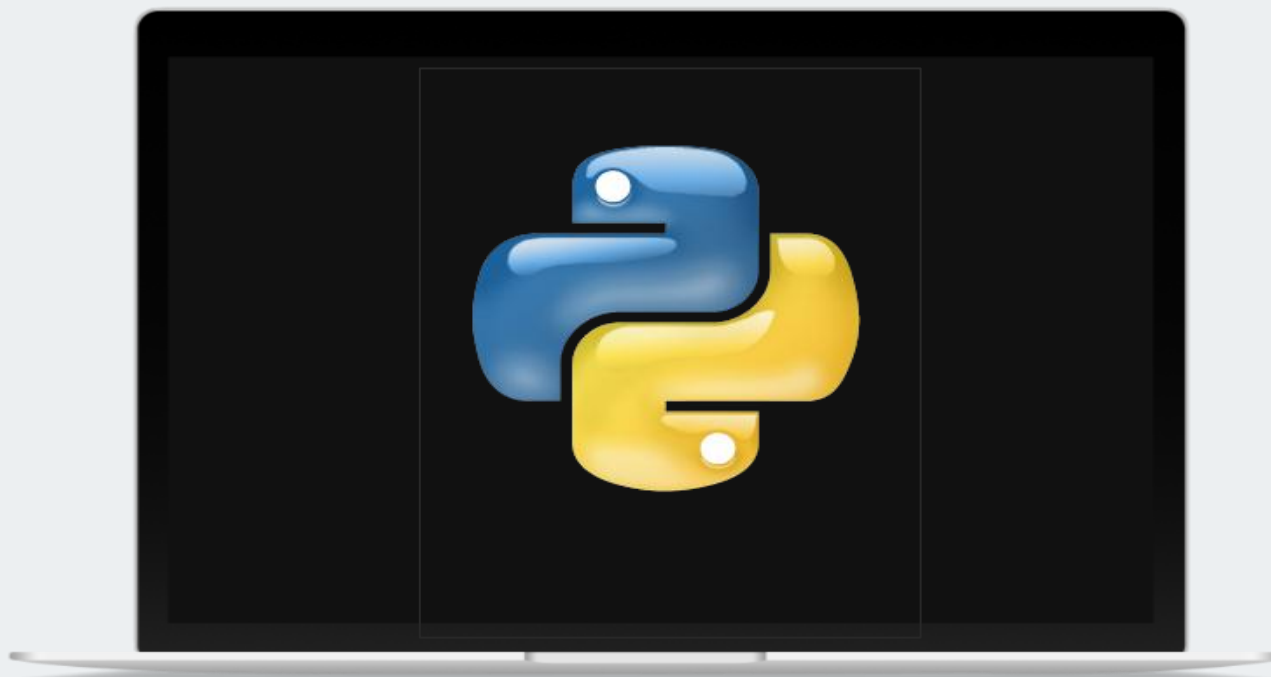


PART TWO

python开发环境搭建



1.1.2python 开发环境搭建



Python是一种面向对象的解释型计算机程序设计语言，由荷兰人 Guido van Rossum于1989年发明，第一个公开发行版发行于1991年。

python 语言具有以下特点：

语法简洁清晰，
强制用空白符
(white space)
作为语句缩进；

易读、易维护，
被大量用户所
欢迎的、用途
广泛；



开源、免费、
功能强大；



具有丰富和强
大的库，无论
你想实现什么
功能几乎都有
一个库能满足
你的要求；

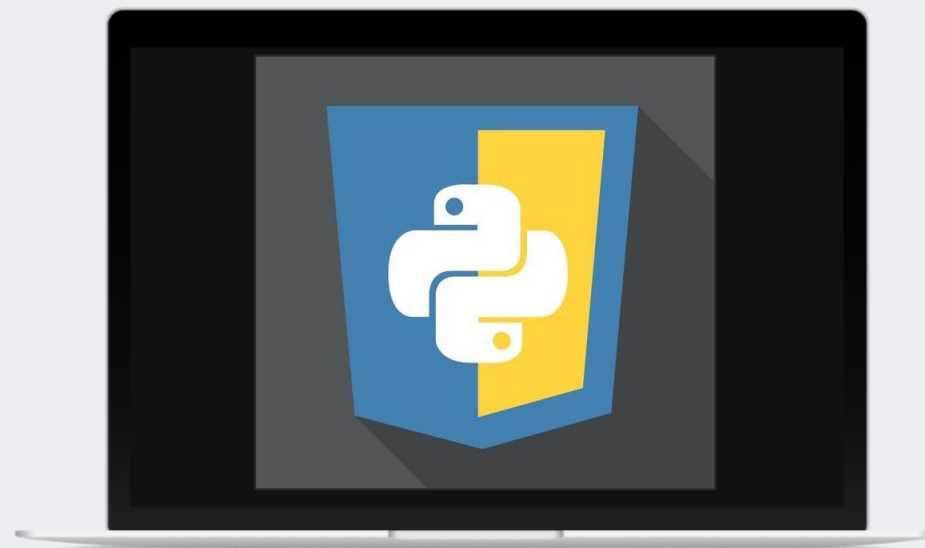


是解释性语言，
变量类型可变，
类似
JavaScript；

1、python自带开发环境

Python的开发环境也十分简单，用户可以到官网

<https://www.python.org/>中直接下载Python的程序包。目前Python有两个主流的版本，一个是Python2.7，另外一个Python3.6，这两个版本在语法上有些差异，本教程主要使用Python3.6。



2、PyCharm与python的开发环境

一个比较流行的开发环境是PyCharm，它的风格类似Eclipse，是一种专门为Python开发的IDE，带有一整套可以帮助用户在使用Python语言开发时提高其效率的工具，比如调试、语法高亮、Project管理、代码跳转、智能提示、自动完成、单元测试、版本控制。

读者可以到PyCharm的官网

<http://www.jetbrains.com/pycharm/>去下载免费的PyCharm Community版本，这个版本虽然不及收费的Professional专业版本功能强大，但对于一般应用已经足够了。



3、Anaconda与python的开发环境

一个比较流行的开发环境是Anaconda，这个程序比较庞大，但它是一个十分强大的Python开发环境，它自己带Python的解释器，也就是说安装Anaconda时就自动安装Python了，同时它还带有一个功能强大的IDE开发工具Spyder。Anaconda最大的好处是可以帮助用户找到与安装Python的各种各样的开发库，使得Python的开发十分方便与高效。另外Anaconda对Windows用户十分有用，因为Python的一些开发库在Windows环境下安装常常出现这样那样的问题，而Anaconda能顺利解决这些问题。读者可以到官网<https://www.continuum.io/downloads>下载Anaconda。





THANK YOU