**数据采集与爬虫技术**

课程聚焦于利用Python编程语言进行网络数据抓取与处理。课程从网络爬虫的基础原理与架构讲起，涵盖HTTP协议、网页结构解析等知识要点，深入教授如何运用Python丰富的库（如BeautifulSoup、Scrapy等）编写高效的爬虫程序，能够精准地从各类网站抓取文本、图片、链接等数据，并详细讲解数据清洗、存储与管理的方法与策略，同时探讨网络爬虫在信息检索、数据分析、市场调研等多领域的广泛应用场景，注重培养学生独立开发网络爬虫项目的实践能力以及在合法合规前提下运用爬虫技术获取有价值信息的综合素养，为学生开启大数据时代数据获取与利用的大门奠定坚实基础。