

# Project-Driven Reform and Practice of the "Data Acquisition and Web Crawling" Course Teaching Mode

Yun Gong

School of Computer Science, Guangdong University of Science and Technology, Dongguan,  
Guangdong 523668, China

## Abstract

With the deep integration of informatization and digitization, data has become a core strategic resource in modern society. Python crawling technology, as a key tool for data acquisition and processing, plays an important role in multiple industries. However, traditional course teaching modes suffer from issues such as a disconnection between theory and practice, insufficient practical opportunities for students, and a lack of team collaboration abilities, making it difficult to meet the demand for crawling technology talent in modern enterprises. To address this, this paper proposes a project-driven reform of the "Data Acquisition and Web Crawling" course teaching mode. By introducing real enterprise project cases, such as data crawling from e-commerce websites and public opinion analysis on social media, students can master technology application in real business scenarios, enhancing their practical and problem-solving abilities. In addition, through group collaboration, students can develop team collaboration and communication skills during project implementation. To further optimize teaching effectiveness, this paper also suggests establishing a diversified evaluation system that covers classroom performance, homework completion, project practice outcomes, and other aspects, to comprehensively reflect students' learning situations. At the same time, modern information technology is utilized to build an online learning and exchange platform, providing students with more learning resources and practical opportunities, and promoting interaction and cooperation among students.

## Keywords

Project-driven; data acquisition and web crawling; practical teaching; curriculum reform.

## 1. Introduction

In the era of deep integration of informatization and digitization, data has become one of the most core strategic resources in modern society. Python crawling technology, as a key tool for acquiring and processing these data, plays an indispensable role in multiple industries such as e-commerce price monitoring and competitor analysis, financial investment decision support, market research and user profiling, smart city and traffic management, medical and health research, environmental monitoring and carbon neutrality, and news aggregation.

"Data Acquisition and Web Crawling" is an advanced professional course for big data-related majors, covering web data scraping, data parsing, image recognition and text processing, crawler data storage, crawler frameworks, and other content. Through this course, students can not only master the basic knowledge of Python crawling but also independently develop Python crawler project programs using frameworks. The course aims to cultivate students' abilities to solve practical problems using crawler technology, laying a solid foundation for students to engage in data acquisition and analysis-related work in the future.

## 2. Current Situation Analysis

The "Data Acquisition and Web Crawling" course is highly practical. However, traditional course teaching not only focuses on imparting theoretical knowledge and lacks the cultivation of students' practical abilities but also often adopts a too singular teaching mode, unable to meet students' diversified learning needs and adapt to the rapidly changing technological environment. Therefore, reforming the course teaching mode, optimizing course settings, improving teaching quality, and cultivating high-quality talent with crawler technology application abilities have become important tasks in the current educational field.

### (I) Issues of Updating and Optimizing Course Content

With the continuous development of crawling technology, its application scenarios and demands are constantly changing. Traditional teaching content may be outdated and unable to meet the actual needs of modern enterprises for talent. Therefore, it is necessary to update and optimize course content based on the latest technological trends and industry demands to ensure that students can learn the most cutting-edge and practical knowledge and skills.

### (II) Issues of Combining Practical Teaching with Theoretical Teaching

The traditional teaching of web crawling courses usually adopts the methods of lecture and demonstration, where students passively receive knowledge and lack opportunities for active thinking and exploration. There is often a serious disconnection between theoretical teaching and practical teaching. Theoretical teaching is too abstract and boring to stimulate students' interest and enthusiasm in learning, nor can it cultivate their innovative thinking and problem-solving abilities; while practical teaching lacks systematic guidance and sufficient practical opportunities. Therefore, there is a need to effectively combine theoretical teaching with practical teaching, so that students can master theoretical knowledge while also deepening their understanding and application of knowledge through practical operation.

### (III) Cultivation and Enhancement of Students' Practical Abilities

Web crawling technology is a highly practical discipline, and students' practical abilities are crucial for their future career development. However, traditional web crawling courses often focus too much on imparting theoretical knowledge while neglecting the cultivation of students' practical abilities. Teacher-student interaction is often limited, making it difficult for teachers to timely understand students' learning situations and problems, and also unable to reasonably adjust teaching strategies based on feedback, resulting in students' learning problems not being quickly resolved. In addition, practicing skills through only some simple and limited classroom experiments or after-school assignments is far from sufficient for students to proficiently master web crawling-related technologies. Therefore, it is necessary to adopt a project-driven teaching approach to allow students to learn and grow through practice, enhancing their practical abilities.

### (IV) Cultivation of Team Collaboration and Communication Abilities

Team collaboration and communication abilities are essential in the implementation of web crawling projects. However, traditional teaching methods often neglect this aspect of cultivation. Therefore, it is necessary to set up team collaboration projects to guide students on how to collaborate and communicate within a team, cultivating their team collaboration and communication abilities.

### (V) Construction and Improvement of Teaching Resources and Environment

Implementing project-driven teaching requires adequate teaching resources and a good teaching environment for support. Therefore, how to construct and improve teaching resources and the environment to provide strong support for project-driven teaching is also an important issue that this project needs to address.

### 3. The Significance of Introducing the Project-Driven Concept

To address the issues existing in the traditional teaching of the "Data Acquisition and Web Crawling" course, a reform based on a project-driven teaching mode is proposed, which introduces the project-driven concept into classroom teaching practice to stimulate students' interest and enthusiasm in learning. In addition, the proportion of practical courses is increased, and more actual cases and projects are introduced to allow students to learn and master web crawling technology through practical operations. The specific measures and their significance are as follows:

1. The project-driven teaching approach can promote cooperation and exchange between universities and enterprises, realizing resource sharing and complementary advantages. By cooperating with enterprises to carry out projects, universities can gain a deeper understanding of market demands and technology development trends, thereby more accurately adjusting teaching content and methods to cultivate high-quality talents with practical application abilities.
2. Traditional theoretical teaching often focuses on imparting knowledge points while neglecting the cultivation of students' practical abilities. Project-driven curriculum teaching introduces real project cases, allowing students to directly participate in actual projects and gain an in-depth understanding and mastery of the principles and applications of Python web crawling technology through hands-on practice. This teaching method not only helps students consolidate theoretical knowledge but also enhances their practical abilities and problem-solving abilities.
3. Compared with traditional lecture-based teaching, the project-driven teaching approach is more interesting and challenging. Students can feel a sense of accomplishment in the process of participating in projects, thereby stimulating their interest and enthusiasm for Python web crawling technology. This teaching method helps improve students' autonomous learning abilities and motivation, promoting enhanced learning effects.
4. The project-driven curriculum teaching mode establishes a diversified evaluation system, including classroom performance, homework completion, project practice achievements, and other aspects. Different evaluation methods and approaches are adopted according to the characteristics of different aspects, while introducing multiple evaluation methods such as student peer evaluation and teacher evaluation to more comprehensively reflect students' learning situations.

### 4. Reform of the Project-driven Curriculum Teaching Mode

#### (I) Reform of teaching content and methods

1. Replace traditional simulation or theoretical teaching tasks by introducing actual enterprise or industry web crawler project cases, such as crawling e-commerce website data and analyzing social media public opinion. In this way, students can be exposed to real business scenarios and data, thereby gaining a deeper understanding of the practical application of crawler technology.
2. Implement a practice-oriented teaching method to encourage students to learn and master crawler technology through hands-on operation. For example, teachers can set specific crawler tasks for students to learn and apply knowledge in the process of completing the tasks.
3. The project-driven course "Data Acquisition and Web Crawling" should not only cover the basic knowledge of web crawling but also related interdisciplinary knowledge, such as Python programming, data analysis, network security, data cleaning and storage. The integration of these knowledge areas helps students construct a complete knowledge system and improve their overall quality.

#### (II) Reform of teaching organization and management

1. Adopt a group collaboration approach for project-based learning to cultivate students' team collaboration and communication skills. Each group can be responsible for a specific crawler project. During the project implementation process, students need to jointly analyze requirements, formulate plans, divide labor and cooperate, and discuss and resolve issues when encountered, thereby enhancing the overall effectiveness of the team.

2. Implement and evaluate in stages, that is, divide the project into multiple stages, with clear learning objectives and evaluation criteria for each stage. This ensures that students have a clear direction throughout the learning process and can receive timely feedback and guidance.

3. Utilize online platforms to establish online communication spaces for students to learn and discuss anytime, anywhere. Teachers can regularly post learning resources, task requirements, and answers to questions on the platform to interact with students.

#### (III) Optimization of teaching resources and environment

1. Provide abundant teaching resources, including textbooks, courseware, video tutorials, online courses, etc. These resources can help students better understand and master crawler technology.

2. Provide a good practice environment, including laboratories, computer rooms, and other hardware facilities, as well as corresponding software tools and platforms. This ensures that students can smoothly carry out crawler practices.

3. Establish cooperative relationships with enterprises to jointly carry out practical teaching and project development. This not only provides students with more practical opportunities and real business scenarios but also strengthens the connection between schools and enterprises, promoting the integration of industry, academia, and research.

## 5. Project Innovation

#### (I) Introduce industry cases

Introduce actual industry cases into teaching to enable students to be exposed to real business scenarios and needs, thereby learning and applying crawler technology more targetedly. This helps to shorten the distance between students and actual work and improve their employment competitiveness.

#### (II) Strengthen technological innovation

Encourage students to try new technological methods and solutions during project implementation, such as using machine learning for automatic clustering of web content and rule-based engines for web content filtering. This innovative teaching method helps cultivate students' innovative thinking and problem-solving abilities.

#### (III) Build an online learning and exchange platform

Utilize modern information technology to construct an online learning and exchange platform, providing students with more learning resources and practical opportunities. Students can share experiences, discuss issues, and showcase achievements on the platform, thereby forming a good learning atmosphere and interactive mechanism.

## 6. Conclusion

With the rapid development of technology and changes in industry demands, relying solely on the imparting of theoretical knowledge in the teaching of the "Data Acquisition and Web Crawling" course can no longer meet the needs of modern society for talent in crawling technology. Therefore, more practical content needs to be introduced into teaching to cultivate students' practical operation and problem-solving abilities.

The project-driven teaching method for the "Data Acquisition and Web Crawling" course emphasizes taking actual projects as the core, allowing students to learn and master crawling technology in the process of completing projects. In this way, students can gain a deeper understanding of the practical application scenarios of crawling technology and continuously accumulate experience in practice. At the same time, the project-driven teaching method also helps stimulate students' interest and enthusiasm in learning, improving their learning outcomes.

## References

- [1] Rui Z , Yan-Huai Z , Yong-Mei Z ,et al.Teaching Model on Software Engineering Based on "Project-driven, Competition-promotion"[J].Computer Knowledge and Technology, 2019.
- [2] Hong-Mei S , Rui-Sheng J .Research on Case Teaching of Software Development Comprehensive Practice Based on Project Driven [J]. Procedia Engineering, 2012, 29:484-488. DOI:10.1016/j.proeng.2011.12.747.
- [3] Zhang Z .Software engineering Teaching Reform based on project-driven[J].Computer Education, 2009.DOI:10.1007/978-3-540-93824-8\_365.
- [4] Li Z , Li J M , Guan Y C ,et al.Practice of Actual Project Driven Teaching Model on Software Engineering Major[J].IEEE, 2009.DOI:10.1109/ICIECS.2009.5363170.
- [5] Huang L , Dai L , Guo B ,et al.Project-Driven Teaching Model for Software Project Management Course[C]//International Conference on Computer Science & Software Engineering. IEEE, 2008.DOI:10.1109/CSSE.2008.76.